Unit 1 ❑ Aims and scope of comparative physiology

Structure

1.1 General physiological functions and principles

The science of physiology is the analysis of function in living organisms. Physiology is a synthesizing science which applies physical and chemical methods to biology. The term physiology originated from the greek word 'Physiologikos' which means 'discourse on natural knowledge.' French physician Jean Fernel introduced the term in 1552. Physiology is thus a branch of science that deals with normal functions of the body.

In the science of physiology, there is nothing supernatural about life. All the living processes of an organism can be explained on the basis of physico-chemical changes and structural pecularity. The progress of ancillary subjects like biology, microscopic anatomy, physics, chemistry, have been very useful in understanding the working processes of the living organisms. All physiological processes are governed by basic laws of physics and chemistry. Accordingly to the changes in the environment, functional alterations also occur and thus the survival is made possible. Hence, it is held that Physiology is a tripod science, its three legs being Anatomy, Physics and Chemistry. To know physiology, one has to study the subject with a sound basic knowledge of these three subjects. Blood flow through cardio-vascular system is governed by the 'Laws of fluid dynamics', exchange of fluids between different body compartments is regulated by hydrostatic pressure and osmotic pressure ; transfer of information signal involving transaction and development of action potential ; exchange of gases, and solutes obey the laws of diffusion. A study of all these physical phenomenon in the body comes under the head Biophysics. Biochemistry is the branch of science that explains the functions of the body on a chemical basis. All biochemical reactions including energetics are all governed by chemical laws and principles of thermodynamics. A study of structure of an animal comes under the study of Anatomy. The study of gross structure is called macroscopic anatomy, while the study of fine structures is called microscopic anatomy. Histology, cytology, histochemistry and cytochemistry all come under the study of microscopic anatomy. Medical statistics is a branch of science that deals with evaluation of experimental

data. These are used to confirm the authenticity of any observed change in the functions of the body.

In short, the whole purpose of Physiology is to explore, with the help of modern techniques, the normal functioning of the living organisms—their principle, their mechanism and their control.

A short summary of physiological study

The structural and functional unit of a living body is a cell. A collection of similar cells having same origin and performing same function but held together by intercellular substance is called a tissue. There are unicellular and multicellular animals in the world. When a multicellular animal is analyzed in terms of cell, it is complicated and is seen to be made up of huge number of cells of various shape and size, and a large amount of intercellular substance. But if the animal body is analyzed in terms of tissue, it is found to be made up of basically four types of tissues like, epithelial, connective, muscular and nervous. The term 'organ' is used to denote a structure being made up of those four types of tissues and is responsible for a particular function. For example, liver is an organ made up of hepatic cells which are epithelial cells ; it has connective tissue fibres forming septa, walls of blood vessels contain smooth muscle cells i.e. muscular tissue ; it has nerve supply, so it possess nervous tissue. Liver has many functions including secretion, excretion and metabolism. Hepatectomised animal will not survive. Finally, many different organs join together to constitute a 'system' responsible for a particular function. For example, the excretory system, responsible for excretion is made up of organs like kidney, ureter, urinary bladder and urethra. A multicellular animal has many systems in it like skeletal, muscular, cardiovascular, respiratory, excretory, endocrine, reproductive, nervous etc. There is a question why so many systems are present in an animal body including humans. A critical analyses revealed that every-one of them is necessary for the living organism. The ultimate aim of any living organism is to survive in the world. For survival one has to do work. For doing work energy is necessary. Hence, every living organism must have a mechanism to yield energy for doing work. Food is the source of energy. The items of food are carbohydrate, lipid, protein (calorigenic food); vitamins, minerals and water. Of these, the catorigenic foods gives energy, hence the name. When these foods are oxidized energy is released. For example—

$$C_6H_{12}O_6 \xrightarrow{6O_2} 6CO_2 + 6H_2O + \text{energy.}$$

thus, a molecule of glucose, when oxidized to CO_2 and H_2O, energy is liberated. It is seen from the equation that, to achieve this, oxygen is supplied by respiratory system. So this system is required. Next about the food, monosacclaride, say glucose, is utilised by the cell ; but the carbohydrate food that we eat is not monosacharide. They are either polysacharide like starch or glycogen or disaccharide like sucrose or

lactose. Hence, in the body there should be a system to convert these poly- and di-saccharides to monosaccharides so that they can be utilized in the body for not only supplying energy but also for forming any carbohydrate like substances in the body. This conversion of complex food to simplest substances is called digestion, which is carried out by digestive system. Hence, this system is also necessary.

Next comes the distribution of food. Digestion occurs in the digestive tract, the end products of digestion are produced in the small intestine. From here the end-products are to be transported to the cells all over the body. This requires a transport system. This is done by cardio-vascular system. Hence, the animals cannot go without this system. When the foods are utilized by the body by way of metabolism, many intermediate products are formed which are harmful to the body and need elimination to outside the body, e.g. urea. Thus a system has to be developed to cause excretion i.e. excretory system. To maintain a structure, a skeletal system is required ; for movement, mascular system is necessary. All these different systems must be properly regulated. For such regulation, endocrine system and nervous system have been developed.

1.2 Validity of comparative approach

For practical purposes, physiology has been divided into three categories, like— i) Cellular physiology, (ii) Physiology of special groups, and (iii) Comparative physiology.

An understanding of comparative animal physiology requires some background knowledge in general zoology, animal morphology, biochemistry and cellular physiology.

Comparative physiology is the study of organ function in a wide range of groups of organisms. Comparative animal physiology integrates and co-ordinates functional relationships that occur in more than one group of animals. It is concerned with the ways in which diverse organisms perform similar functions. Study of the comparative physiology reveals that, generally dissimilar organisms may show striking similarities in functional characteristics as well as in responses to the same environmental stimulus. On the other hand, closely related animals frequently react differently to their surroundings. Light, temperature, O_2 tension and hormone balance are used or are considered as variables for each function. Comparative physiology uses in addition to these, species or animal type as a variable for each type of function. Comparative animal physiology use kind of animal as an experimental variable which is a unique kind of biological generalization. Comparative environmental and behavioral physiology constitutes a bridge between molecular and organismic biology.

The most important function of comparative physiology is to put man into perspective in biological history and phylogenetic relationships.

To achieve the goal, the plan for the study of comparative physiology is made to examine the relation between components of the environment and the whole organism and to analyze the interactions in terms of organs and cell physiology.

For the applied biologist, comparative physiology has practical application in describing the physiology of economically important animals. For the ecologists, it helps the understanding of restriction of plants and animals to particular habitats. For the medical physiologists, the comparative viewpoint places man in his proper biological perspective. For the general biologist, it provides meaning to natural variation, and general principles which can be reached only with kind of organism as a variable.

1.3 Organisms and cell physiology

1.3.1 Diffusion

Particles, molecules and ions have a tendency to spread uniformly in the entire available space by their incessant random movements. This tendency to spread is called diffusion. Though they are moving at random in all directions, a greater number of particles move from a region of higher concentration to the area of lower concentration, than in reverse direction. Hence, there occur a net diffusion from higher to lower concentration and the two concentrations become equal. In this state equal number of particles move in all directions and the net diffusion is zero.

Fick's First law of diffusion states that the rate of diffusion (flux) of a solute particle is directly proportional to the magnitude of concentration gradient and it occurs down the concentration gradient.

Graham's law states that the diffusional flux (J) of a gas varies inversely with the square root of its density (P) and molecular weight (M).

Diffusion coefficient (D) denotes the diffusibility of the particle or gas. In case of solid the 'D' is the mass of solute diffusing across 1 sq cm area in 1 sec, down a concentration gradient of unity. Diffusion coefficient of a gas is the volume of gas that diffuses across 1 sq. cm. area per second, down a partial pressure difference (ΔP) of unity.

Diffusion is directly proportional to :

(i) Concentration gradient / Pressure gradient / electrical gradient.
(ii) Solubility in the medium
(iii) Temperature of the medium
(iv) Cross-sectional area through which diffusion is taking place.

Diffusion is inversely proportional to :

(i) Distance to be travelled / thickness of the membrane

(ii) Diameter of the diffusing particle

(iii) Viscosity of the medium

(iv) Shape of the particle—spherical particle diffuse easily than elongated particle.

(v) Charge of the particle and charge of the pore. If the charges are same, diffusion will be less, if the charges are opposite, diffusion will be more.

Biological application

1. Absorption of certain substances like pentoses, some minerals, some water soluble vitamins from intestine is carried out by diffusion. Same is the case of renal reabsorption of urea.
2. Water, water soluble substances pass largely by diffusion through water filled pores of the cell membrane. This is dependent on the pore size and size of the diffusing particle. Pore size is about 0.8 nm. Hydrated K^+ ion is 0.4 nm in diameter, hydrated Na^+ ion has 0.5 nm diameter. Hence, movement of K^+ is two times faster than the movement of Na^+. Glucose and galactose molecules have a diameter of 0.85 nm. Hence, they cannot pass via pores. They pass via membranes after combining with some carrier.
3. There are voltage gated ion channels in the membrane. When these channels open, ions move by diffusion along electrochemical gradient.
4. Exchange of respiratory gases occur by way of diffusion. The partial pressure of O_2 is higher in alveolar air and lower in the deoxygenated blood. So O_2 enters blood from the lungs. Alternatively, the partial pressure of CO_2 is higher in deoxygenated blood and lower in the alveolar air. So, CO_2 diffuses from blood to alveolar air. In the same way, due to difference of partial pressure, O_2 flows from the oxygenated blood to the tissue cell and CO_2 diffuses from the tissue cells to the blood.

   During hyperventillation, alveolar size is increased. Hence, the surface area is increased and wall thickness of the alveoli is decreased. So, rate of diffusion of O_2 and CO_2 is increased.
5. The alveolar surface area per unit body weight is larger in children than in adult humans. So, the resting O_2 uptake from the alveoli is higher in children than in adults.
6. There is a difference of partial pressure of O_2 and CO_2 between alveolar air and inspired air. So, O_2 and CO_2 diffuse down their respective pressure gradients between the terminal bronchioles and alveoli. This is how the alveolar air has its O_2 renewed and CO_2 partially removed.

7. When the cell size is increased, the surface volume ratio is decreased. This decreases diffusion of gases and solute across the plasma membrane, per unit volume of the cell. Enhanced cell-size also decreases rate of diffusion.
8. Replacement of cutaneous respiration across the general body surface by gill respiration and pulmonary respiration enhances the respiratory surface. So, respiratory exchange increases with the evolution of gills and lungs.

1.3.2 Osmosis

The movement of solvent from solvent side to the solution side at or from dilute solution to a concentrated solution, when the two are separated by a semipermeable membrane is called osmosis. Semi permeable membrane is that membrane which allows only the solvent molecules to pass through and not the solute molecules.

Osmotic pressure (OP) is the pressure which has to be applied on the solution side to stop the osmotic inflow into it from pure solvent.

Van't Hoff's laws of osmotic pressure

1. It is directly proportional to the molar concentraion of the solute so long as the temperature remains constant.

$$\pi = k_1 c$$

(where, π = Osmotic pressure, k_1 = constant, c = molar concentration.)

2. Osmotic pressure of a solution is directly proportional to the absolute temperature (T), so long as its concentration remains constant.

$$\pi = k_2 T \text{ (where, } k_2 \text{ = constant)}$$

3. Van't Hoff - Avogadro law :

Identical numbers of moles of different solutes produce an identical osmotic pressure, when dissolved in the same volume of the solvent at the same temperature.

Osmotic pressure (OP) is expressed as atmosphere, mm of Hg or dynes per sq. cm (dynes cm^{-2}).

1 mole of a nonionized solute is equivalent to 1 osm (osmole). Osmolarity of a solution is its solute concentration in osmoles per litre of the solution. A solution of one mOsm of any solute in a litre possess an osmotic pressure of 19.3 mm Hg at 38°C. Osmolatity of a solution is its solute concentration in Osm per kg of solvent.

It can be determined by—

1. Osmometer.
2. Barkley—Hartley method.
3. Freezing point method. This method depends on the direct proportionality between the osmotic pressure and the depression of freezing point of a solution. Depression of Freezing point is measured by Beckmann Thermometer.

$\Delta t = K_f M$ [where, M = Molal concentration of solute.

or, $M = \frac{\Delta t}{K_f}$ Δt = Depression of freezing point of the solution.

K_f = Cryoscopic constant]

The freezing point of one Molal solution is called cryoscopic constant (K_f) / molal freezing point.

The K_f for water is – 1.858°C

$\pi = CRT$

π = Osmotic pressure in atmosphere.

where C = Molal concentration

R = Molar gas constant (0.082 litre atmosphere)

T = Temperature in Absolute scale.

A sample of urine freezes at –0.56°C. Calculation of its OP at 37°C

$$\text{Molal conc. } C = \frac{-0.56}{-1.858} = 0.3$$

The sample will have OP at 37°C (π = CRT)

$0.3 \times 0.082 \times 310 = 7.6$ atms.

or, 7.6×760 mm Hg.

Osmotic work

When substances are transferred from lower concentration to higher concentration, osmotic work must be performed upon them. But when substances pass from higher to lower concentration, osmotic work is done by them.

Relation between osmotic work and concentration change is given by the equation—

$$W_{min} = NRT \ln \frac{C_2}{C_1} \quad \text{or, } 2.3\ NRT \log \frac{C_2}{C_1}$$

[W_{min} = Minimum osmotic work in small calories involved in the transfer of N moles of substance from a molal concentration of C_1 to a molal concentration C_2.

R = Gas constant = 1.987 cal/mole/degree

T = Absolute temperature

ln = Natural log

2.3 is the factor for converting natural log to the log to the base 10.

Calculation of osmotic work to be done to transfer or secrete 3.545 g of Cl^- from plasma urine at 37°C, when the Cl^- concentrations in plasma and urine are 0.1 and 0.2 Mol respectively.

Cl^- conc. in Plasma ..... 0.1 Molal C_1

Cl^- conc. in Urine ..... 0.2 Molal C_2

Osmotic work to be done to transfer / secrete 3.545g of Cl^- at 37°C in urine.

N = 0.1, R = 1.987, T = 310

$$W_{min} = NRT\ 2.3 \log \frac{C_2}{C_1}$$

$$= 0.1 \times 1.987 \times 310 \times 2.3 \log \frac{0.2}{0.1}$$

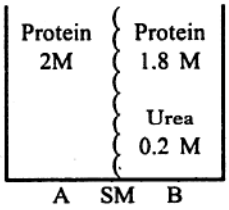
$$= 141.7 \log 2$$

$$= 141.7 \times 0.301$$

$$= 42.65 \text{ cal}$$

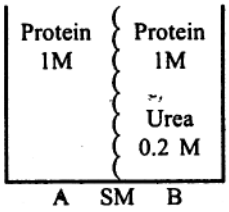
Osmoticity and tonicity

The osmoticity of a solution depends on the total solute concentration, both diffusible and nondiffusible. Whereas, tonicity of a solution depends on the concentration of nondiffusible solute only. Hence, two solutions may be isosmotic, but may notbe isotonic.



Solution A is isosmotic with solution B,
But solution A is hypertonic to solution B.

Similarly, two solutions may be isotonic but may not be isosmotic.



Solution A and B are isotonic, but the Solution B is hyperosmotic to Solution A.

SM = Semipermeable membrane.

Solvent always flows from by hypotonic to hypertonic solution. In biological system, tonicity is considered because biological membrane are not strictly semipermeable. They allow some solutes to pass.

Biological application of osmosis

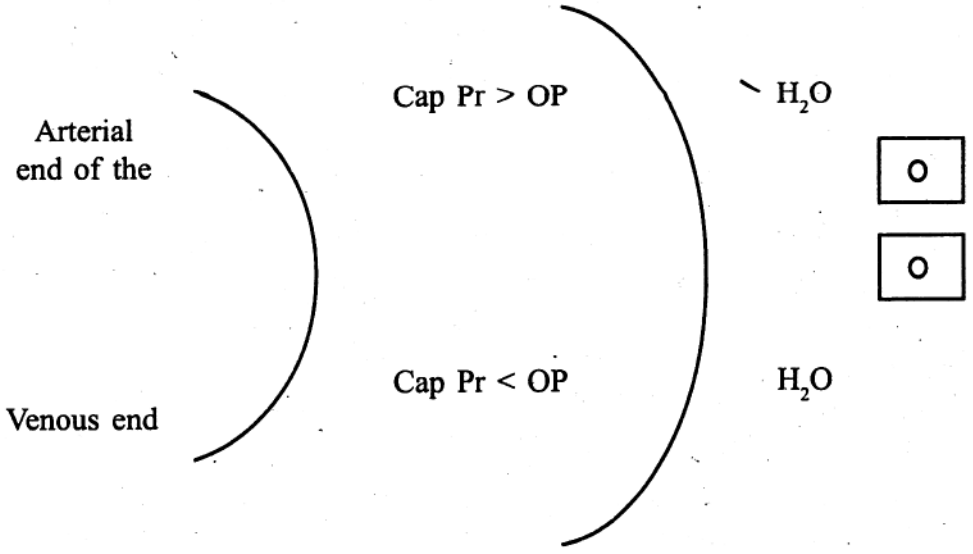
1. **Hemolysis, crenation, plasmolysis :** RBC is hemolysed or animal cell ruptures if placed in hypotonic solution. This happens because solvent flows from hypotonic solution to RBC or cell. Consequently, they swell and burst at a certain degree of swelling. Similarly, these cells or RBC will shrink (crenated)

if placed in hypertonic solution. It is due to osmotic outflow of water from the hypertonic solution. Plant cells lose water when placed in hypertonic solution. So, the cell membrane collapses and withdraws from the cell wall. This is called plasmolysis. Plant cells swell if placed in hypotonic solution, but they do not burst due to rigidity of cell wall.

2. **Osmotic distension of RBC :** Osmotic pressure of RBC fluid is 1.5 atm., higher than plasma, due to higher electrolyte concentration in RBC than plasma. So, RBC remain slightly distended, but they do not rupture. However, due to genetic disorder, Vit-E deficiency, or selenium deficiency and other defects, RBC may not be able to withstand such distension and rupture.

   RBC of camel are more resistant to osmotic distension. Camel can drink more than 100 litres of H_2O in 10 minutes. Blood becomes temporarily highly hypotonic, but RBCs do not rupture. It has been observed that RBC may be distended upto two times its volume, but hemolysis does not occur.

3. Osmotic pressure of plasma is higher than tissue fluid. It is due to plasma proteins. Total osmotic pressure of plasma is about 5453 mm Hg and that of ECF is 5430 mm Hg. The difference of 23 mm Hg is due to plasma proteins. It is called colloidal osmotic pressure of plasma. This osmotic pressure, hydrostatic pressure i.e. capillary pressure and pressure of the tissue fluid play an important role in the exchange of body fluid across the blood capillaries. Owing to pressure differences, fluid passes out from arterial end of blood



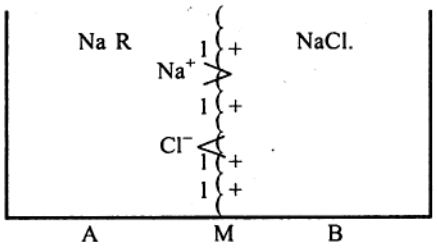
capillaries to the tissue fluid and body fluid enters from the tissue fluid into the capillaries at the venous end. These forces also govern the flow of fluid between any two compartments.

In Kwashiorkor, (a form of childhood malnutrition), hepatic cirrhosis and nephrosis, plasma protein concentration is decreased. So, colloidal osmotic pressure is decreased. This decreases water retention by the plasma and so edema develops.

4. Osmotic pressure plays a vital role in the absorption of water from intestine and kidney tubules.
5. Water absorption by plant roots is also governed by osmotic pressure. Root hair cells have higher osmotic pressure than surrounding soil-fluid. So, water enters into root hair cells. Water moves from one cell to the next cell by cell to cell osmosis and thus other cells are also distended. Hence, the cells become turgid and rigid and they stand erect on watering.

1.3.3 Donann membrane equilibrium

When a non diffusible ion in present in a solution, the distribution of diffusible ions across the membrane will be unequal. This was observed by Donann and it is called Donann membrane equilibrium.



In the compartments A and B, NaCl solution is present. The two compartments are separated by a semi-permeable membrane. Which allows NaCl to pass through but not the nondiffusible ion R. When R is not present the ion distribution in the two compartments is :

$Na^+ A = Na^+ B$

$Cl^- A = Cl^- B$

$Na^+ \times Cl^- A = Na^+ B \times Cl^- B.$

But when the nondiffusible ion (R) is added the distribution of diffusible ions become unequal :

$Na^+A > Na^+ B$

$Cl^- A < Cl^- B.$

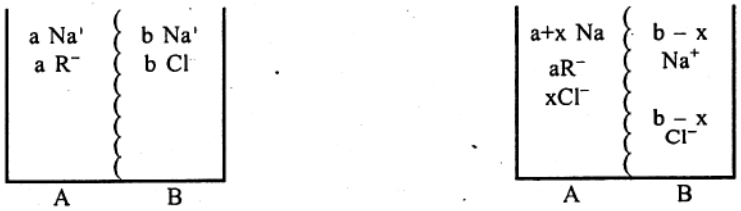
$Na^+ A \times Cl^- A = Na^+ B \times Cl^- B$

The unequal distribution depends on the nature of nondiffusible ion. In a situation, as described above, where the nondiffusible ion is negative there will be more positive diffusible ion in the compartments that contains the nondiffusible ion.

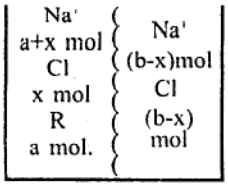
Because of such unequal distribution of diffusible ion, there will be a pH difference, a potential difference, and an osmotic pressure difference on the two sides

of the membrane. If we consider the distribution of H^+ and OH^- then there will be more H^+ in the compartment that contains negative nondiffusible ion. Hence, there will be a pH differnce in the two compartments. The fluid in the compartment where the nondiffusible negative ion is present will have a lower pH (acidic) compared to the other compartment. Because of the presence of negatively charged hemoglobin in RBC, the pH of RBC fluid is less than plasma. The membrane is permeable to Na and Cl. Concentration of Na^+ in A is more than Na^+ concentration in compartment B. The concentration of Cl^- is less in compartment A than in the compartment B. Thus Na ions in A will try to move from compartment A to compartment B and Cl^- will try to move from compartment B to compartment A. This can only be prevented by developing an opposite electrical gradient. Hence, the side of the membrane where the non-diffusible negative ion is present is negatively charged and the other side of the membrane is positively charged. Donnan Phenomenon plays an important role in the development of resting membrane potential. It can be shown by calculation that total solute concentration in the compartment that contains the non diffusible ion is more compared to the other compartment.

In the diagram of the concentration of NaR is taken as 'a' and the concentration of NaCl as 'b' then at equilibrium, the concentration of NaCl in the two compartments will be :



Concentration of Na^+A will be (a+x) and Cl^- A will be x. Whereas concentration of Na^+ b will be b-x and Cl^- will be b-x. Since the product of the concentration of Na^+ and Cl^- will be equal in the two compartments. The equation can be written as $(b-x)^2 = (a+x)x$. From the equation the value of x in terms of a and b will be $x = \frac{b^2}{a+2b}$. Thus the amount of NaCl (x) that has moved from compartment B to compartment A is inversely proportional to the concentration of non diffusible ion in compartment A. When the concentration of NaCl (b) will be large relative to the concentration of nondiffusible ion (a) the value of x will be more. If specific arithmetic number is assigned to a as 1 mole and b as 2 mole then specific value of x will be obtained.

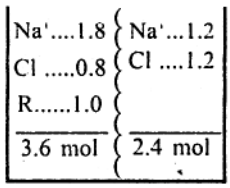


if a = 1 mol
and b = 2 moles

$$x = \frac{b^2}{a+2b}$$

$$= \frac{4}{1+4}$$

$$= \frac{4}{5} \text{ or } 0.8 \text{ mol.}$$



Thus the total solute concentration in compartments A is 3.6 mols and in compartment B is 2.4 mols. Hence the osmotic pressure of the solution in compartment A will be more than in compartment B.

The Donnan principle operates to regulate the distribution of electroyte ions across the membrane of living organism. However, the plasma membrane is not a strictly semipreamble membrane. It has selective permeability property. Thus the membrane may be impermeable to some diffusible ion and in that case such ions do not move following Donnan phenomenon. However this may produce Donnan effect. Thus Donnan effect is there but its quantitative evaluation is difficult. Deviation from Donnan Phenomenon has also been found. The distribution of cations, Na^+, K^+, Ca^{2+} an Mg^{2+} between plasma and lymph follows Donnam principle. They are high in plasma than in lymph. Plasma has a higher concentration of negatively charged nondiffusible ion protein. The difference in the composition of glomerular filtrate and plasma also follows Donnan principle. The glomerular filtrate contains less amount of positive diffusible ion and more amount of negative diffusible ion compared to plasma. This is also due to Donnan effect because of negatively charged plasma protein. Peculiarly, however, the concentration of Cl^- is higher in lymph than in plasma as required by Donnan principle, but the concentration of HCO_3^- in the plasma is little higher than in lymph which is contrary to Donnan principle.

1.3.4 pH and buffer

pH is negative logarithm of concentration of hydrogen ion (CH) to the base 10.

pH = –log CH

It is a measure of the H^+ concentration of a solution. It is determined by pH meter.

Buffer is a mixture of weak acid and its salt or weak base and its salt.

Henderson— Hasselbach equation of a buffer solution is given below.

$pH = pKa + \log \frac{Salt}{Acid}$. pKa is the dissociation constant of the buffer acid. In a buffer solution with equal amount of salt and acid. pH = pKa

Each buffer has a pH and a capacity.

The pH of a buffer depends on the pKa and salt ratio. The capacity of a buffer depends on the salt/acid ratio and on its amount. The buffers can not prevent the change of pH but it can resist the change of pH when any acid or base is added to it. For example when any strong acid is added to a buffer sol the salt of the buffer react with the strong acid added. As a result, equivalent amount of buffer acid is produced. Because buffer acid is a weak acid the pH is not changed much. However, as the salt/acid ratio is altered, there should be a new salt/acid ratio, so pH will definitely be changed.

$$\frac{CH_3COONa}{CH_3COOH} + HCl \longrightarrow CH_3COONa + HCl \rightarrow CH_3COOH + NaCl$$

Thus the buffer present in our body fluid help to resist the change of pH during production of metabolic intermediates. The buffer of our body has a high salt/acid ratio e.g. $\frac{HCO_3^-}{H_2CO_3} = \frac{20}{1}$. Such a buffer is good for handling acids. Important buffers present in our body fluid are bicarbonate, phosphate and protein buffers.

1.3.5 Poiseuille's Law

This law states the relationship between pressure gradient, resistance and volume flow.

Poiseuille and Hagen equation is $Q = \frac{P_1 - P_2 \pi (r)^4}{8\eta L}$

Where, Q = volume flow in ml/sec.

$P_1 - P_2$ = pressure difference between 2 points (dynes/cm^2.)

r = radius of the tube in cm.

L = length of the tube in cm.

η = viscasity of the fluid in poise.

It is revealed from the equation that

1) It pressure gradient is doubled, the flow will be doubled.

2) If the length of the tube or viscosity is doubled, the flow will become half.

3) If the radius of the tube is doubled. The flow will increase 16 times.

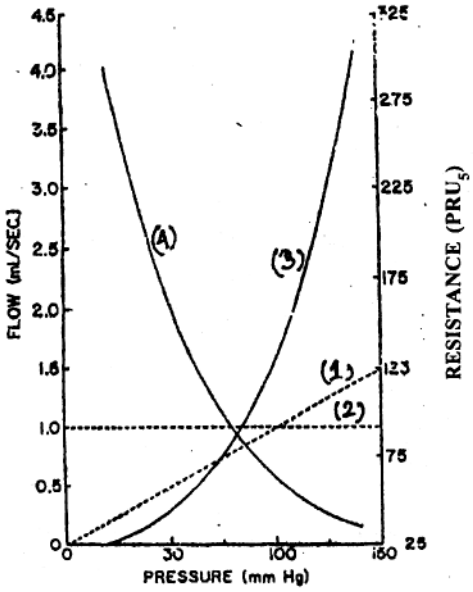


Fig 1.1 : The relationships of flow, pressure and resistance in nondistensible tubes using Newtoning fluids (dotted lines) at constant temperature (Poiseulle's law) and blood in vessels (solid lines).

In Poiseuille's experiment the fluid was a constant visocity fluid (Newtonian fluid) the tube was a rigid tube.

This law is applicable in closed vascular system but it is not obeyed in toto, because the vascular tubes are not rigid but distensible.

With a Newtonian fluid flowing through a rigid tube the pressure flow curve (Fig 1.1) is linear (1) because the resistance is not changed as pressure is increased (2). But with flow of blood in a vascular system it is not so (3) because the resistance is changed (4). It is very high at low pressure. When there is no flow (closing pressure). At high pressure the resistance is almost constant so flow becomes more linear.

1.4 Suggested Questions

1. Why respiratory exchange increased with the evolution of gills and lungs ?
2. How alveolar air gets its O_2 renewed and CO_2 partially removed ?
3. Why resting alveolar O_2 uptake is higher in children than the adults ?
4. Why rate of diffusion of respiratory gas is increased in hyperventillation ?
5. What is diffusion ? Name the factors which favour the rate of diffusion/ oppose the rate of diffusion.
6. State the Fick's law of diffusion and Graham's law of diffusion.
7. What is diffusion coefficient ?
8. Describe the role of diffusion in the transfer of respiratory gases.
9. Describe why evolution of gills and lungs is advantageous compared to cutaneous breathing.
10. Gill and lung respiration is better than cutaneous respiration—Justify.
11. What is Donnan membrane equilibrium ? State the role of Donnan phenomenon on osmotic pressure difference across the plasma membrane ?
12. Why pH of RBC fluid is less than plasma ?

13. Write the Henderson-Hasselbotch equation for a buffer solution.
14. Mention the factors on which the pH and capacity of a buffer depends.
15. Describe the role of pressure gradiant and resistance on the volume flow in a closed vascular system.
16. What is osmosis ?
17. What is osmotic pressure ?
18. State the laws of osmotic pressure.
19. State the role of osmotic pressure in the exchange of fluid across the blood capillaries.

Unit 2 ❑ Thermoregulation

Structure

2.1 Regulation of body temperature in homeothermic animals specially mammals

Animal organisms have been classified into two groups depending on their capability of regulation of body temperature namely, homeothermic and poikilothermic animals.

Homeothermic Animals : The animals which maintain their body temperature more or less constant inspite of changes in the temperature of the environment are called homeothermic animals. They are also called warm blooded animals. Birds and mammals belong to this category.

Poikilothermic animals : The animals which do not maintain a constant body temperature and it varies as the environmental temperature is changed are called poikilothermic animals. They are also called cold blooded animals. Fish, amphibia, reptiles and invertebrates belong to this group.

In homeothermic animals a balance is maintained between heat gain and heat loss mechanisms present in them and thus the body temperature is kept constant. This is achieved by the operation of a complex neuro-hormonal mechanism regulated by the temperature controlling centre located in the hypothalamus.

Birds and mammals have a normal body temperature ranging between 40^0 – 43^0 C and 36^0 – 39^0 C respectively.

Normal body temperature in human subject

The normal body temperature of man ranges between 35·8 – $37·3^0$ C when measured by introducing clinical thermometer into the mouth cavity (oral temperature). The temperature recorded from arm pit is called axillary temperature. It is slightly lower (37^0 C) than oral temperature. The rectal temperature is slightly higher than oral temperature (about 37^0 C). The superficial temperature i.e. skin or surface temperature ranges between $29·5^0$ C and $33·9^0$ C. The average temperature in deeper

tissue is called core temperature. It is always more than oral or rectal temperature. It is about 37·8^0 C (100^0 F).

The body temperature is slightly more in children than adults. It shows diurnal variation. It is about 1^0 C less in the morning and it reaches maximum value in the afternoon (1^0 C more than normal). The body temperature is increased in exercise, emotion and after meals. It is slightly less in females than males and the temperature rises sharply immediately after ovulation (0·5^0 C – 1^0 C). It is reduced 0·5^0 C in sleep.

A constant rectal temperature can be maintained in a nude human subject at air temperature or 0 – 1^0 C for 1–2 hours. In dry air maximum tolerance to high temperature has also been observed e.g. 200^0 C in a nude subject and 260^0 C in heavily dressed man.

Heat gain and heat loss mechanisms of the body

The body will gain heat from the environment if the body temperature is less than the environment and vice versa by way of conduction, convection and radiation. Thermogenesis and thermolysis : Heat production/gain is thermogenesis, heat loss is thermolysis in the body. However, the physiological processes of heat gain mechanism are shivering and nonshivering thermogenesis and vasoconstriction. On the other hand, physiological processes of heat loss from the body are vasodilation, sweating and fenting. The mechanism by which the body temperature is normally adjusted is known as thermotaxis.

Temperature regulating centre

A balance is maintained between thermogenesis and thermolysis and thus the body temperature is maintained. There is a temperature regulating centre in the hypothalamus which maintains this balance and thus help maintain homeothermy. If this part is destroyed the animals become poikilothermic.

On the basis of ablation and stimulation experiments it was shown that the anterior hypothalamus act as hot responsive centre, stimulation of which increases heat loss and dereases heat gain. The posterior hypothalamus acts as a cold responsive centre, activation of which stimulates heat gain mechanism and inhibits heat loss mechanisms.

However, subsequent studies were made with physiological stimuli like cooling and heating the hypothalamic regions by diathermy or by using thermodes. The studies revealed that the temperature detecting centre for both heat and cold are located in the anterior hypothalamus, whereas the centre initiating heat loss or heat gain mechanisms are present in the posterior hypothalamus.

Microelectrode studies on the hypothalamic neurons revealed the presence of different temperature sensitive neurons in the anterior hypothalamus.

(a) Some warm sensitive neurons respond on local heating of hypothalamus and they stimulale heat loss mechanism.

(b) Some respond to local cooling of hypothalamus and they activate heat gain machanism.

(c) There are some neurons whose rate of firing is changed in response to peripheral warming or cooling.

(d) There are neurons which respond to a rise in both hypothalamic and peripheral temperatures.

(e) There are neurons whose firing rate remain unchanged with temperature variation.

Serotonergic and cholinergic neurons are involved in temperature regulation, Mayers and Sharpe observed in monkey with push-pull technique using saline that when the donor is cooled a transmitter is released which can cause shivering in receipient. Similarlv heating the animal causes the recipient to lower body temperature. Subsequently it was observed in 'saline withdrawal' technique that saline withdrawn from donor monkey during cooling contained increased concentration of serotonin. Liberation of noradrenaline was augmented by peripheral warming.

Mayers *et al* reported the presence of a heat conservation and production pathway originating in the anterior hypothalamus is passing through the posterior hypothalamus and there is a heat dissipation pathway originating in the posterior hypothalamus.

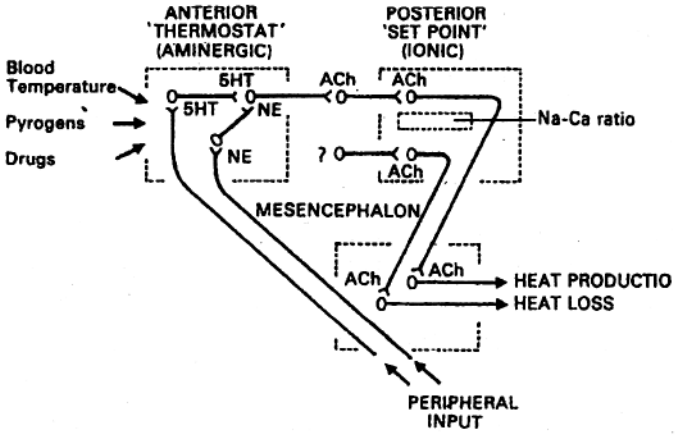


Fig 2.1 : Diagram of a model to account for temperature regulation under normal conditions as well as during a pyrogen-induced fever. Factor which affect the aminergic 'thermostat' in the anterior hypothalamus at given, and the outflow from the posterior hypothalamic 'set point' is mediated by a cholinergic system which passes through the mesencephalon. 5H = 5-hydroxytryptamine ; NE=noradrenalin ; ACh = acetylcholine. (Fro Myers, R. D. (1971). *Pyrogens and Fever.* CIBA Foundation Symposium Elsevier, Amsterdam.)

They proposed a mechanism of action of hypothalamus in the temperature regulation on the basis of experiments in monkey.

Serotonergic neurons in the preoptic region increase their firing in response to cooling. This causes activation of cholinergic pathway to the posterior hypothalamus and heat production is initiated (Fig. 2.1).

The noradrenergic neurons in the anterior hypothalamus are stimulated due to warming. They inhibit the serotonin cholinergic heat production

pathway by noradrenergic blockade of the synapse of serotonin—cholinenrgic junction. Such suppression of the heat production pathway permits the second cholinergic system in the posterior hypothalamus to activate the efferent heat loss pathway.

Set-point in the hypothalamus

Different experimental evidence indicate that there is a 'set point' in the hypothalamus which is like that of a tharmostat. When the temperature goes above the set point heat dissipation is increased and the temperature is lowered and brought back to set point. On the other hand if the temperature goes below the set point heat production occurs, the body temperature is raised and the set point in reached.

The presence of warm and cold sensitive neurons in the hypothalamus indicate that the combined action of these neurons play the role of hypothalamic thermostat. Thus temperature sensitive neurons in the anterior hypothalamus form the basis of a 'set point' mechanism.

It is held that over a physiological range of hypothalamic temperatures both warm and cold sensitive neurons or sensors are active. However a change in temperature will increase the discharge of one type of neuron and will decrease the firing of the other type. Hence, there will be a value of hypothalamic temperature at which the activity of the two sets of sensors will be balanced in terms of body temperature regulating responses which they produce. This will be the setpoint value of hypothalamic temperature. There are hypothalamic neurons whose firing rates remain unaltered by changes in hypothalmic temperature. These neurons or neural firings can provide a 'reference signal'. The difference between the firing rates of these temperature insensitive neurons and the temperature sensitive neurons provide an 'error signal'. The direction and magnitude of the 'error signal' will determine the extent of operation of heat gain and heat loss mechanisms to bring the temperature to the set point level (Fig 2.2).

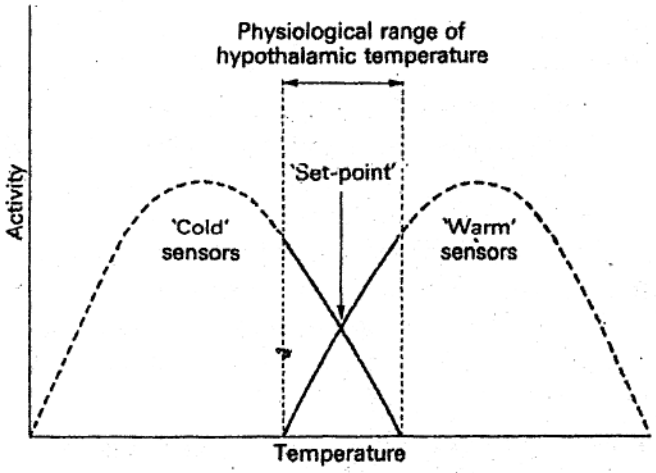


Fig 2.2 : Determination of hypothalamic 'set-point' temperature by the balance of activity in warm and cold-sensitive neurons. The response characteristics of these temperature sensors have been represented by bell-shaped curves. The interrupted portions of these curves are a hypothetical projection based on the suggestion that the temperature-sensitive characteristics of these sensors may be similar to those of the warm and cold receptors in the cat's tongue as described by Hensel and Zotterman.

Basic mechanism of temperature regulation

A change in body temperature, with reference to 'Set point', stimulates thermoreceptors present on the body surface as well as inside the body. These impulses (originate due to stimulation of receptors) reach the hypothalamus via neural pathways and stimulate it. Altered temperature of blood also activates it. It then modifies the heat gain and heat loss mechanisms via posterior hypothalamus, as required and the body temperature is kept constant. It is believed that in the cold the neural impulses are very important part in hot, the altered temperature of the blood is the most effective slimulus. It has been observed that when one hand is immersed in hot water, vasodilation is seen in the other arm. But this does not occur if the blood flow from the arm immersed in hot water is blocked. Similarly, if one hand is immersed in cold, there occur vasoconstriction in the other hand but this does not stop even if the blood flow for this region in blocked.

Temperature regulation in the cold

When the body temperature tends to decrease on exposure to cold environment, the following changes occur to keep the body temperature normal.

The cold receptors present in the body are stimulated, the neural discharge in the nerves attached to this is increased. These impulses and also the decreased temperature of blood stimulates the temperaure detecting centre in the anterior hypothalamus. This results in stimulation of thermogenesis and inhibition of thermolysis.

Thermogenesis is the primary motor centre for shivering and locatd in the dorsomedial portion of the posterior hypothalamus. Impulse discharge from here increase the tone of skeletal muscles throughout the body. The arrector pili muscles also contract. Thus heat production is increased. During maximum shivering body heat production can be increased four to five times normal. Intially there occur uncoordinated muscle twitches, the intensity of which increases until the rhythmic activity of visible shivering appears. It has been, shown that tensing the muscles can raise the heat production to 2–3 times the basal level, shivering starts when the body temperature goes below critical temperature.

Non-shivering thermogenesis also starts at the same time. Impulse discharge via sympathetic fibers reach the adrenal medulla and causes discharge of adreno medullary hormones adrenaline and noradrenaline. Neural impulse also cause release of TRH from hypothalamus which causes release of TSH from anterior pituitary. This hormone stimulates thyroid gland and causes secretion of T_3 and T_4 hormones. The calorigenic effect of these hormones increase heat production. Adrenaline increases glycolyis and fatty acid oxidation. It increases lipolysis specially from brown fat, and the oxidation of the released fatty acid is increased. Beside heat production it helps in heat conservation by preventing heat loss by way of vasoconstriction. Thyroid hormones

specially triiodothyronine (T_3) increases general metabolism and heat production. It is held that calorigenic effect of thyroid hormones is due to increased ATPase activity, oxidative phosphorylation, glycolysis as well as uncoupling of oxidation and phosphorytation. It also increases the calorigenic effect of adrenaline. Beside thermogenesis, there occur vasoconstriction in the body. This is caused by increased sympathetic discharge from posterior hypothalamus. This vasoconstriction decreases heat loss from the body and thus heat is conserved, and body heat production is increased. Along with this, there occur inhibition of sweating and this is another mechanism to prevent heat loss. In this way by stimulating shivering and nonshvering thermogenesis and vasoconstriction and inhibition of sweating, body heat production is increased and heat loss from the body is inhibited. As a result, fall of temperature is prevented and normal body temperature is maintained.

Different exprimental results indicate that :-

(1) The controlling centre for integrating and coordinating various temperature conservation function is contained in the posterior hypothalamus, although it is not thermosensitive.

(2) Critical temperature is the external temperature below which heat production has to be increased by shivering thermogenesis to maintain normal body temperature. It differs with species. In case of a tropical animal it is between 20^0–30^0 C, whereas in arctic animal like huskey dog the basal heat production may not increase even at enviromental temp, below –30^0C. In this respect it is 20^0 C in a nude human subject.

(3) In case of muscle shivering both flexors and extensors are stimulated.

(4) The nerve inpulse goes to the muscle via lower motor neuron being activated by tecto-spinal or rubro-spinal tract and not by pyramidal tract.

(5) Shivering begins when environmental temperature is below critical temperature (20^0 C in case of nude man). The degree of shivering increases as the ambient temperature is decreased.

(6) In respect of body temperature shivering begins when the core temperature goes below the set point (37^0 C in case of man). Shivering is increased as the body temperature goes much below the set point.

(7) Shivering is more if both air and core temperature are low.

(8) When air temperature goes below 27^0C, vesoconstriction occurs.

Temperature regulation in hot environment

When the body temperature tends to increase on exposure to hot environment the following changes occur to maintain the normal body temperature.

The hot receptors present in the body are stimulated. Nerve inpulse in generated in the nerve fibres attached to them. These impulses as well as increased temperature

of blood stimulate the temperature detecting centre located in the anterior hypothalamus. This results in inhibition of thermogenesis and stimulation of thermolysis.

Vasodilation and sweating cause thermolysis. It has been found that when the air temperature goes above 27^0 C the skin vessels dialate. This helps in heat loss from the body to outside.

Sweating : The sweat glands are innervated by sympathetic cholinergic fibres. When body temperature is increased impulse goes to the sweat glands from posterior hypothalamus via this pathway and cause secretion of sweat. Evaporation of sweat takes away the latent heat of vaporisation from the skin surface and thus the body is cooled.

It has been found that above 29^0 C environmental temperature, very little heat can be lost by convection. So, vasodialation does not help much in heat loss from the body. Above 35^0 C environmental temperature, the body loses heat entirely by evaporation. The air temperature at which sweating begins has been found to be about 31^0 C and for those lightly clad is 29^0 C in humans (the average skin temperature is about 34^0 C).

In respect of body temperature, when it goes above the setpoint, sweating begins. It increases as the body temperature is increased. It has been found that the degree of sweating is related to skin temperature and core temperature. When both core and air temperature high the degree of sweating is more. It has been observed that maximum rate of sweat secretion for thermolysis (thermal sweat) may be as high as 1.7 liters/ hour or more. When one liter of sweat is evaporated 580 Kcal of heat is lost from the body

As sweat coures from the blood rapid sweating demands a large cutaneous blood flow and so requires cutaneous vasodialation. This is brought about by :

(a) External heat acting directly on the blood vessels.

(b) Reflexly from cutaneous warm ending.

(c) By the rise of blood temperature acting directly on the hypothalamic centre.

(d) Activity of sweat gland secretion leads to formation of bradykinin which acts as vasodialator.

If the air is humid, sweat cannot be evaporated and so sweating cannot help in heat loss.

The secretion of sweat also show adaptation. If exposure to heat is continued, the sweat secretion is increated, it starts at a lower temperature, i.e. threshold for sweat secretion is decreased. The NaCl content of sweat is decreased by the action of aldosterone secreted in this condition. This prevents salt loss from the body. In case of excessive sweating, water and salt should be ingested to prevent dehydration and salt deficiency.

It has been shown that some birds and mammals allow their body temperatures to vary widely, either regionally in the body or in the whole body for sametime.

Temperature regulation in new born

The new born of all species, including human infants are of smaller size compared to their corresponding adults. Hence, they have greater surface area in relation to their body weight. This poses a great problem in the maintenance of normal body temperature.

A fall in body temperature can be prevented by decreasing heat loss and increasing heat production, cutaneous vasoconstriction has been found to occur in response to cold even in premature babies. The new born has a higher capacity to increase its heat production. A newborn animal or human infant, when exposed to cold hunch themselves and tucks their limbs to reduce effective surface area. Brown adipose tissue plays a very important role in heat production in infants. Fatty acid produced from lipolysis in brown adipose tissue are oxidized within these tissue as well as oxidized in other tissues after being carried there via circulation.

Tempurature regulation in non-sweating animals

In non-sweating animals, evaporative heat loss is achieved by other means. In case of birds panting and gular flutter help in evaporative heat loss from respiratory tract. Gular flutter, however, has been shown to be metabolically less expensive. In respect of heat gain in cold, these animals respond to cold by shivering. Muscular activity appear to be their only means of increasing heat production. Non-shivering thermogenesis has not been demonstrated in them. They lack brown fat and fail to show a thermogenic response to norepinephrine. However, they depend on white adipose tissue for compensation in cold.

When plenty of water is available, some birds may increase their cooling by urinating on their legs. Kangaroo rats have no sweat glands. Evaporation from the lungs is a great source for heat loss in them. However, due to scarcity of drinking water, they do not use much water for heat regulation, rather they are nocturnal and move for food only during night when the environment is cool.

Small animals such as rodents have no true sweat gland and also do not pant. They avoid heat by living in underground burrows.

Behavioral regulation of body temperature

Animals have been found to change their behaviour by regulating their body temperature. In a cold condition the animals including humans are found to curl themselves or remain close together to prevent heat loss. Alternatively, in hot weather the animal stretch themselves to increase the surface area to favour heat loss. The

rodents, specially in deserts have been found to enter into burrows in the daytime to avoid heat. Moreover, they become nocturnal so that this may be able to acquire food at night. This type of temperature regulation has been found to be much documented in several poikilotherms, both vertebrates and invertebrates.

Some use solar energy, while others utilize metabolic heat to raise the body temperature. The terrestrial environment is more prone than aquatic to sudden temperature changes. The most successful terrestrial poikilotherm insects and reptiles have made use of behavioral response to avoid extreme temperature or to elevate temperatures sufficiently for certain activities. Some reptiles have well developed sensory organs for this purpose. The infrared sense organs in the facial pit of rattle snakes can detect a temperature difference of the order of 0·001 to 0·005^0 C. This helps the animal to orient themselves to warm and cool environments. It also helps the animal in detecting a wam blooded and cold pray. The insects use solar or metabolic energy to warm up before flight. Social insects like ants, termites and bees may regulate their temperature in their nest or hives through varied activities.

Temperature regulation in aquatic animals

Water has high thermal conductance and a high heat capacity. Hence, the thermal loss to water is much higher than to air of the same temperature. The cooling power of water may become as high as 100 times as great as for air.

Many whales and seals live and swim in the near freezing water. However, regarding body temperature they are similar to other warm-blooded animal. It is around 36^0–38^0 C. There are three ways by which aquatic animals can cause heat balance in cold condition.

(1) They can live with a lowered body temperature.

(2) They can increase their metabolic rate to increase heat production.

(3) They can increase their body insulation to reduce heat loss.

It has been found that several species of seals and dolphins have resting metabolic heat production twice as high as would be expected from their body size. However, in the harp seal, their metabolic rate remained same in water even when the temperature goes down to the freezing point. (The critical temperature for harp seal in water is below freezing point, but it has not been determined). Thus the effective solution to the problem is to develop effective insulation and this is done in these animals.

They have a thick blubber under the skin that acts as an insulator. In seals, the temperature of the skin surface is identical to that of water but at the depth of about 50 mm (the thickness of the blubber), the temperature is nearly that of core temperature. The seals and whales being such well insulated feel difficulty in heat loss when the temperature of water in increased. In such a situation its skin temperature is increased

to eliminate heat. This is achieved by increasing blood flow through the blubber to the superficial layer of the skin which is well supplied with blood vessels. The cutaneous blood vascular system permits a precise regulation of the amount of heat that reaches the skin surface and thus is lost to the environment. Since the insulator is located internally to the surface of heat dissipation (skin), blood can bypass the insulator and heat loss during heavy exercise or in warm water can be independent of insulator.

In arctic land mammals, furs act as insulator. It is located outside the skin surface. The surface temperature of the body skin under the fur is close to core body temperature. Most of the insulation resides outside the skin surface. In case of polar bear fur is the insulator. But it has also a substantial layer of blubber under the skin. This is very important. When the polar bear swim in cold water, most of the furs get wet, most of the insulation value of furs is lost, the blubber plays an important role in heat conservation. Hence, blubber plays an inportant role in samiaquatic way of life.

In seals and whales, that lack blubber, have flippers are flakes. These appendages are well supplied with blood vessels and can lose substantial amount of heat if required. However, they require heat conservation in the cold which is achieved by developing heat exchanger system. It has been found that in the whale flipper, each artery is completely surrounded by veins. Thus, as warm arterial blood flows into the flipper, it is cooled by the cold venous blood that surrounds it in all sides. The arterial blood therefore reaches the periphery precooled and loses little heat to the water. The heat has been transferred to the venous blood, which is pre-warmed before it reaches the body. This kind of vascular heat exchange arrangement is called counter current heat exchanger, because blood flows in opposite directions in the two vessels.

2.2 Temperature regulation in poikilotherms

Poikilothermic animals attain the temperature of the environment where they live and do not maintain a constant body temperature like homeotherms. But some poikilotherms have some regulation of body temperature and thereby keep their body tempereture above the ambient temperature. Such temperature regulation has been developed to satisfy their requirement to combat thermal stress.

The insects are poikilotherms, but they face thermal problem because of high rate of metabolism during their flight. If the insect is too cold its muscles will not contract suitably for flight. In such a situation, it can increase the temperature of their flight- muscle by contraction similar to shivering in man. On a cold day a butterfly

or moth has been found to vibrate their spread cut wings for several minutes before take-off. At this stage, the temperature of the flight muscle has been found to be raised to about 35^0 C within 6 minutes in an ambient temperature of 20^0 C.

Snakes and lizards have been found to expose themselves to the sun in the early morning and raise their body temperature above the surrounding temperature by sunbasking. A lizard caught in the early morning at Peru (altitude—15,000 ft.), had a body temperature of 31^0 C. Later in the day it avoids sun and in hot days take shelter under rocks or cool burrows. A desert lizard save itself from overheating by behavioral changes because its lethal temperature is about 45^0 C and it will die if it is exposed to such temperature for 10 – 15 minutes.

Active temperature regulation has been observed in snakes. Large python has been found to coil their bodies around their eggs. At a room temperature of 25^0 C, the snakes maintain their body temperature about 5^0 C higher and this is achieved by spasmodic muscle cantraction like shivering in man.

2.3 Hibernation

Sometimes the animals pass through a state of dormancy to overcome adverse conditions. Dormancy is a general term for reduced body activities, including reduced metabolic rate. Dormancy has been variously classified according to its depth and duration e.g. sleep, torpor, hibernation, winter sleep and estivation. All these are specific physiological conditions, attained and maintained to pass over unfavorable conditions.

Torpor : It is a state of inactivity often with lowered body temperature and reduced metabolism that some homeotherms enter into so as to conserve energy stores. Small endotherms, because of their high metabolic rate are subject to starvation during periods of inactivity when they are not feeding. During these periods some animals enter into a state of torpor in which the temperature and mebabolic rate subsides. Daily torpor is practicised by many terrestrial birds. Several species of small mammals also undergo torpor e.g. Shrews.

Estivation : It is a state of dormancy in response to high ambient temperatures and/or danger of dehydration. It is also called summer sleep. Some species of both vertebrates and invertebrates exhibit estivation. Well known as estivators are African lungfish (Protopterus).

Hibernation : It is a state of deep torpor, or winter dormancy, in animals in cold climates, lasting weeks or months. Hibarnation is a well regulated physiological state that permits survival during most unfavourable part of the year. In this state the body temperature is greatly lowered and the metabolism, respiration, heart rate are greatly reduced.

As the ambient temperature goes below the critical temperature the heat production is increased to keep the body temperature normal. To have this, more food is required. In the winter, the food supply is short. So the animal cannot consume adequate food necessary to increase the metabolism so that the body temperature is kept constant and the animals will succumb to cold.

In such a situation, the animals give up homeothermy and pass into the state of hibernation and remains dormant. It remains in this state during the unfavourable condition. When the environmental condition becomes favourable it awakes and becomes active again.

Hibernating animals

It is seen in some mammals are birds. Pearson classified homeotherms into obligatory, stubborn and indifferent. Some members of the stubborn and indifferent homeotherms truly hibernate. Mammals that hibernate are some monotremes insectivores, rodents and bats.

Preparation for hibernation

Hibernation usually occurs annually and the animal prepares itself for hibernation by accumulating food. In hibernating bats circadian rhythm in body temperature and metabolism has been observed and this rhythm has been found to disappear as hibernation continues. Ultradian rhythm, however, has been found to persist for at least four years in hibernating golden-mantled squirrels. The animal gradually enter into hibernation when the ambient temperature is below a critical temperature.

Many mammals and a few birds regularly hibernate in each winter. The body temperature drops almost to the level of surroundings. Heart rate, respirations, metabolism and many other functions are greatly reduced. They show little response to external stimuli with the active life suspended, they can survive a long winter. Most animals that hibernate are of small size. Thus many rodents, hamstars, pocket mice, hibernate. insectivores at high lattitude, (e.g. hedge hog) bats hibernate. Humming birds, insect eating swifts and some mouse also hibernate. After the hibernation period is over, they revert back to prehibernation state. This is called arousal.

State of hibernation

In the state of hibernation the plysiological status of the animal is some what different.

1. The body temperature is very close to the ambient temperature and rises and falls with it. The colon and esophageal temperature may be 2–3^0 C above air temperatures. When air temp is 0^{0}C, animals tend to hold their body temperature at about 2^{0}C and the 0_2 consumption increases.

During hibernation, the thermostat activity of hypothalamus is reset at a low level as 20^{0}C or more below normal. At ambient temperature between 5^0 C – 15^0 C, many hibernators keep this body temperature as little as 1^{0}C above ambient temperature.

During hibernation, the thermoregulatory control operate at a low set point and with a reduced sensivity.

2. Prolonged period of suspended respiration develops, the rate of breathing may be reduced to even one or less per minute. Due to reduced respiratory exchange blood of many hibernators become acidic. This acidosis may further lower enzyme activity due to departure from optimal pH of metabolic enzymes.

3. Heart rate is markedly reduced. In case of active ground squirrel whose heart rate is 200 – 400/minute drops to 7–10/minute in the hibernating state.

The P – T interval of ECG in lengthened. Block of conduction may result in uncoordinated beats, vagal stimulation has been found to show no effect on heart. That hibernants can maintain there physiological functions at a low temperature is an adaptive phenomenon. For example, impulse conduction through nerve is blocked at 9^0C in rat (non-hibernant) but at $3{\cdot}4^0$ C in hamstar (hibernant). Na-pump mechanism is almost completely inhibited at 5^0 C in non hibernants, but in hibernants it still persists.

4. Blood flow is reduced, cardiac output shows a small decrease and that is accompanied with decreased heart rate. Stroke volume, however, remains unchanged, WBC count may be reduced and hematocrit slightly diminished. Clotting time has been prolonged it is due to a decrease in prothrombin. Serum Mg^{++} has been found to be high. Metabolism is reduced by 20 – 100 times. The RQ corresponds to fat.

5. Activity of nervous system does not entirely stop when the body cools. Citellus shows low amplitude cortical waves at 5^0 C and at a brain temperature of $6{\cdot}1^0$C it can still localize sound, erect pinae, vocalize and move.

6. Hibernation can last for weeks or for several months in cold climates. Many hibernators arouse periodically (once a week or every four to six weeks) to empty the bladder and defecation.

7. Some hibernators become temporarily resistant to X-irradiation.

8. It is a state in which dormany or torpor occurs that is much more pronounced than deep sleep.

Hibernating animals revert back to original normal state under favourable condition – the term called arousal. The time taken to go into hibernation is often much higher than the time taken for arousal. In ground squirrel, the time taken for attaining the peak torpid state is about 12 – 18 hours but arousal requires less than 3 hours. The hibernators are usually small. There are no large hibernators. They undergo winter sleep. This is because, they have less need to save fuel. This is again due to their normal BMR is low relative to their fuel stores owing to allometry of metabolism and fuel storage. Secondly, because of large mass and low BMR a prolonged metabolic effort would be required to raise the body temperature to normal level for a very low ambient temperature. For example, a large bear would require

at least 24 – 48 hours to warm up to 37^0 C from a hibernating temp of 5^0C. Warming up of such a large mass would also be energetically very expensive.

Arousal from hibernation

The hibernant retains its ability to arouse when the ambient temperature increases appreciably. The animal warms up again and returns to the warm blooded state. Arousal from hibernation is a rapid awakening with warming starting from thoracic region. It is a process of self rewarming and does not require external heating. It is a rather rapid process. The rectal temperature can increase by nearly 20^0 C in an hour. Awakening starts with an explosive outburst of heat production and the peak metabolism is maintained for sometime (e.g. 1/2 hr for marmot) after which the heat production subside gradualy into its basal state.

The rewarning during arousal from hibernation is most expensive.

During arousal, heat is produced by violent muscular shivering and oxidation of fat in the brown adipose tissue. Brown fat or brown adipose tissue are present in smaller or larger patches along the neck and between the shoulders. These cells are filled with fat and with large mitochondria (responsible for brown color— hence the name brown adipose tissue or brown fat.) The tissue has rich blood supply and connected by sympathetic nerves. Here oxidation of fat yields high amount of energy because in these cells uncoupling of oxidation and phosphorylation occurs. It is caused by a protein called thermogenin (MW 32 kDa). Hence, heat is produced at a very high rate. Besides fat it also oxidizes other substrates supplied with blood. The heat production is turned on by noradenaline or by nervous stimulation. This brown fat is found in all hibernating mammals, but it also occurs in many new born mammals including man. It is rather peculiar that they are not seen in a number of birds which regularly hibernate. In most mammals, the brown fat has been found to be lost but it remains in hibernators.

The rewarming occurs differently in different parts of the body. It has been seen that the anterior part of the body that contain vital organs like heart and brain warms much faster than the posterior part. In this respect, rewarming of the heart at the begining is not only essential but must be an initial step because the proper functions of the heart is needed to supply circulation of oxygen for all other organs. The major mass of brown fat are also located in the anterior part. It has been observed that the reheating process in the posterior part begins only when the anterior part has reached near normal temperature.

Control of hibernation

Hibernation – arousal is the cyclic phenomenon operating in hibernating animals. It is regularly controlled. Neuroendocrine involvement has been well documented in this process.

A circadian rhythm has been noticed. The begining of hibernation cycle is usually

associated with a particular time of the year, but it is not necessarily induced by low temperature or adequate food.

Yearly cycle of hibernation is influenced by the duration of the daily light cycle and is also associated with a dark cycle.

A well regulated heat production require a well coordinated Central Nervous System (CNS). e.g. European Hedgehog keep the body temp. at $+5^0$ to $+6^0$ C as the ambient temperature goes below freezing temperature. Aes imilar well regulated torpor has been observed in West Indian humming bird *Eulampis jugularis* which like other humming birds readily becomes torpid. The body temp. of *Eulampis*, when torpid, approaches air temperature but if air temperature drops to below 18^0 C it results for the fall (in body temperature) and keep the body temp at 18 – 20^0C.

It is seen in hamstars that in case of lesion in post-hypothalamus, they fail to enter into hibernation. However in case of lesion in anterior hypothalamus, the ground squirrel may enter into hibernation but they fail to arouse.

Adrenalectomized animals do not hibernate and in hibernation adrenal cortex is depleted. Hibernating animals do not respond by thermogenesis to injection of norepinephrine.

2.4 Suggested questions

1. Explain homeothermic and poikilothermic animals. Discuss the role of hypothalamus in the regulation of body temperature.
2. Write briefly on the regulation of body temperature in poikilotermic animals.
3. How body temperature in regulated in aquatic animals.
4. Describe how body temperature is regulated in hot/cold environment.
5. Comment on temperature regulation in infants.
6. Write a note on behavioral regulation of body temperature.
7. How body temperature is regulated in non sweating animals
8. What is hibernation ? Write a note on the preparation for hibernation.
9. What is arousal ? Describe the process.
10. Write briefly on the neuroendocrine involvement during hibernation.

Unit 3 ❑ Communication among animals

Structure

3.1 Bioluminescence

The emission of light of visible spectrum by living organism is called bioluminescence.

It is a process in which living organisms convert chemical energy into light.

Luminescent organisms have been observed throughout the ages. Christopher Columbus during his historic voyage across the Atlantic noticed mysterious patches of luminescent light around the water of his ships. Aristotle observed that the flesh of dead fish and damp wood appear to luminescence. Subsequently it was shown that glow in dead flesh was dependent on oxygen. Raphael Dubois was the first to isolate the light producing chemicals from 'clams'. This paved the way for the characterization of the molecular chemical and physiological mechanisms behind the process. Bioluminescence has been observed in thousands of species including bacteria, fungi, and marine animals. Bioluminescence can be defined as the emission of ecologically functional light by living organisms. It is primarily a marine phenomenen with a few exceptions seen in freshwater and terrestrial organisms, light emission plays an important role in the life of bioluminescent animals.

Types of bioluminescent animals

Many organisms produce light with the help of photogenic organs, tissues or cells. This is called self-bioluminescence e.g. fire-flies. Some other organisms emit light which is actually produced by some symbiont bacteria present in their bodies. This is called hetero-bioluminescence e.g. Loligo and some fishes.

Bioluminescent organisms

Although it is basically a marine phenomenon many animals exhibit this. It is seen in Protozoa, Cnidaria, Annelida, Arthropoda, Mollusca, Echinodermata, Protochordata and Fishes. It is not seen in terrestrial vertebrates.

Photogenic devices

In some cases, e.g. protozoa the chemicals and enzyme required for light

production are diffused throughout the cell; in ophiuroids the light producing apparatus is present in photogenic cells present in body surface. In higher organisms definite photogenic organs are present which produces light. These organs have been best studied in insects and fishes. These are paired organs in insects being located in the head and abdominal segments. Typically, each photogenic organ has a layer of photocytes containing mitochondria, required chemicals and enzymes for chemical reaction. These cells are arranged cylindrically at right angles to the translucent cuticle which permits light to pass through it. Behind the photocytes, there is a reflecting surface chiefly consisting of urate granules. They receive oxygen through air tubes or tracheoles (Fig. 3.1).

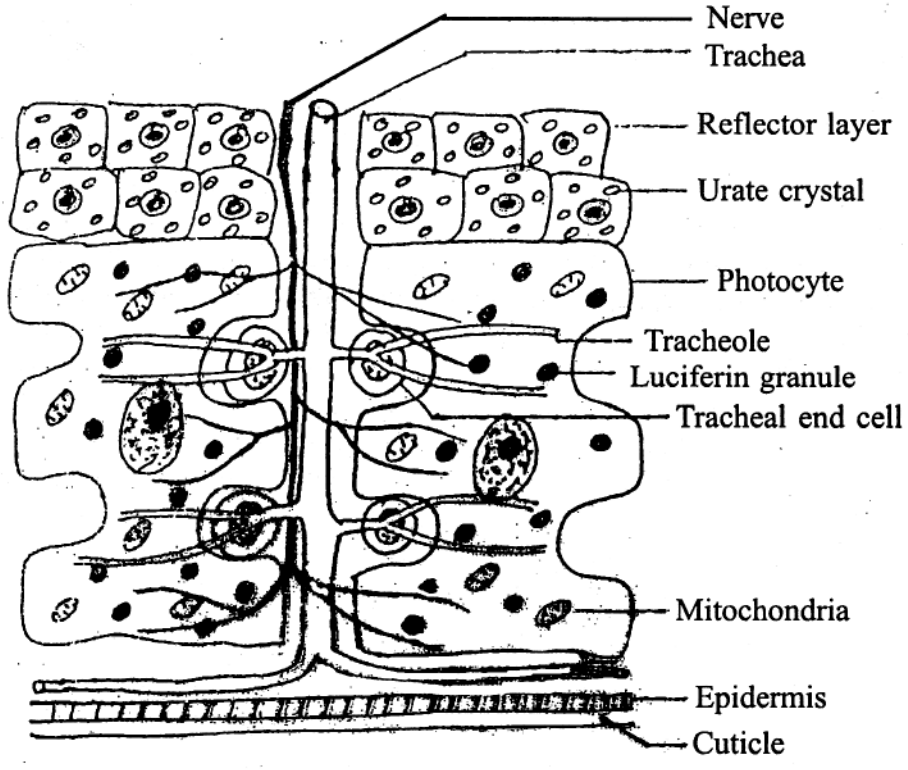


Fig 3.1 : Photogenic organ of a fire-fly. (After Chapman)

A number of fishes, mostly deep sea forms, possess charasteristic luminescent organs called photophores. These organs are probably specialized gland cells of the epidermis. They show considerable variation in their number and mode of distribution on the body. These glanduler structures generally occur along the lateral and ventral sides of the body and head. They may be arranged in one or two rows extending on the sides from head to tail as in *Scopelus* and *Halosauropnis* or they may be located in some limited parts of the body. Besides these organs, a few large and complex

organs may be present on different parts of the body e.g. suborbital organ in *Opostomias micripnuhs, Scopelus benoitii and Pahystomias microdon*. Large photophores present on the elongated first finrays of pectoral, dorsal fins of Angler fishes. In toad fish a large number of photophores are present along the lateral line.

Fishes with luminescent organs have world wide distribution. Majority of them are bathypelagic living at a median depth (500 – 2500 meters) and sometimes migrate to the surface at night. Some species are also found in deeper waters.

Luminescent organs are of two main types. Some species are self luminous. In others the light is produced by symbiotic bacteria. Self luminous photophores are found in some elasmobranchs (Squalidae and Torpidinidae) and in the teleost belonging to the families Stomiatidae and Myctophidae (or Scopelidae). These photophores are simple and complex. The simple photophores consist of a series of radially arranged glandular tubules that receive branches from the adjoining cranial and spinal nerves. In more complex photophores additional structures like a reflecting layer and a lens like structure also develop. The suborbital organ of *Pachystomias microdon* is quite complex in structure. It is a cup like structure and its wall is composed of several concertric layers. Externally there is a layer of black pigment and numerous glandular tubles are present in the cup. A thick layer of light reflecting spicules is present in the cup where axial part is full of a number of radially arranged glandular tubules. The mouth of the cup is occupied by a lens-like structure and the skin forms a covering like an iris diaphragm over it. The organ is supplied by a branch of fifth cranial nerve. Both simple and complex photophores may be present in the same species.

Luminescent organs in which light is produced by symbiotic bacteria are found in a large number of species belonging to different families like *Malacocephalus laevis, Monocentris japonicus, Photoblepharon, Anomalops, Leiognathus* and the Angler fishes.

Structurally these photophores consist of a large number of glandular tubules that secrete luminous bacteria. In some genera e.g. *Malacocephalus*, the highly vascular gland opens by a duct on the ventral surface of the fish in front of the anus *(Malacocephalus)*, species of some genera (*Photoblepharon* and *Anomalops*) possess an elongated luminescent organ below each eye. It consists of numerous long parallel glandular type with rich blood supply. The organ has pores opening to the exterior at the anterior end a 'reflector layer' at the hind end.

Since these luminescent organs produces light for long periods due to bacterial luminescence, mechanical devices have developed to turn the light on and off. In *Anomalops*, the light is cut off by moving the luminescent organ downward by a hinge, so that it comes in contact with a black pigmented tissue. In *Photoblepharon* a fold of black tissue is drawn up like an eye lid over the organ to cut off the light.

Bioluminescence reaction

Reaction occurs between the substrate luciferin and the enzyme luciferase in presence of oxygen. Bioluminesence is an enzymatically catalyzed chemiluminescence. In chemiluminiscence the reaction releases energy but instead of being lost as heat or coupled to some synthetic reaction the energy is used for the specific excitation of a molecule capable of releasing the energy as a photon. The energy of the photon is not fixed, but depends upon the color of the light. The energy, E, is given by the fundamental equation $E = h\nu$. when h is the planck's constant and ν is the frequency. For bioluninescence, where the light is in the frequency range correspoding to wave length between 450-600 nm. the energy involved ranges from about 65 – 45 k cal per mole of photons (an einstein).

French scientist Dubois coined the term luciferin (color peaning) which is the substrate and the enzyme called luciferase that catalyzes the reaction.

Luciferin and lucifeurase are now used an generic terms to refer to the substrate and the enzyme involved in any bioluminescent reaction. Although different specific molecular species are involved in each different class or group of organisms, the general reaction mechanisms have certain close similarities and thus can be represented by a generalized equation. Luciferin is a heat stable, heterocyclic phenolic compound with a molecular weight of 280 and an emperical formula $C_{15} H_8 N_2 S_2 O_3$. There are different types of luciferins like bacterial luciferin, Dinoflagellate luciferin, Vargulin, Coelenterazines, Firefly luciferin.

Luciferases are heat-sensitive enzyme. Firefly luciferase belongs to acyl adenylate thioester forming super family. The enzyme of different organisms have different amino acid composition and structurs. For exemple pure bacterial luciferase is a simple protein with MW 79 k Da. Fire fly luciferase has a MW of approximately 100 k Dal. with two different subunits and one active centre per molecule.

Giese (1973) proposed the following scheme of reaction within photocytes in firefly.

1. At first luciferase catalyzes reaction between luciferin (LH_2) and ATP, the products formed are Adenyl-luciferin and pyrophosphate.

$$LH_2 + ATP \xrightarrow[\text{luciferase}]{} LH_2 - AMP + PPi$$

2. Next, the enzyme catalyzes oxidation of adenyl-luciferin by atmospheric oxygen with the production of high-energy adenyl-oxyluciferin and water

$$LH_2 - AMP + \tfrac{1}{2} O_2 \xrightarrow[\text{luciferase}]{} L - AMP^* + H_2O$$

(High energy adenyl-oxyluciferin)

3. In the next step, high-energy adenyl - oxyluciferin spontaneously returns to ground state liberating its energy in the form of light.

$$L\text{–}AMP^{*} \longrightarrow L - AMP + \text{light.}$$

(Ground state)

4. Later, adenyloxy-luciferin breaks down into oxyluciferin + AMP

$$LAMP \longrightarrow L + AMP$$

(Oxyluciferin)

important characteristics of the reaction :

(a) Luciferase acts best at a pH of 7·8

(b) The activity is highest at 23° C.

(c) Atmospheric oxygen and ATP are required for their reaction.

(d) Divalent cations like Mg^{++}, Mn^{++} or Co^{++} stimulate luciferase action and the ionic environment probably also controls color of the light produced.

(e) Different types of luciferin – luciferase control the colour.

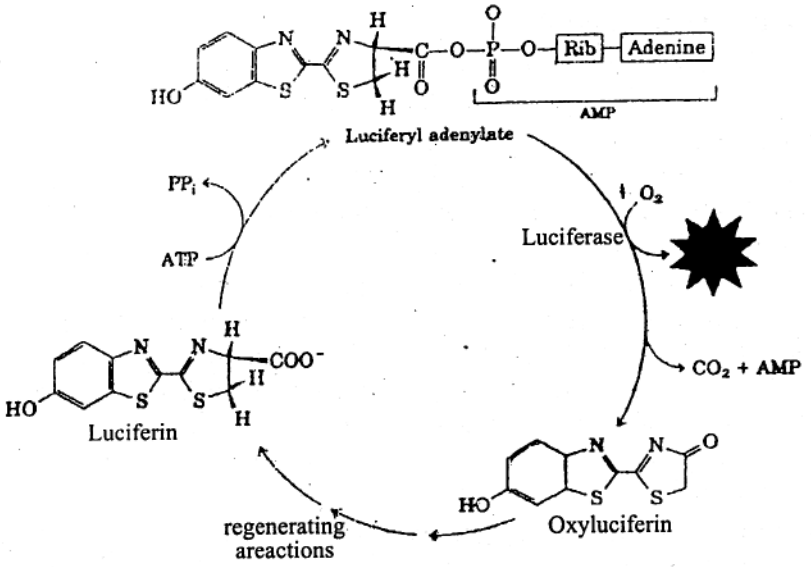


Fig 3.2 : Important components in the firefly bioluminescence cycle.

Physical properties of colours :

(a) The light produced is not hot, it has a temperature below 0·001° C.

(b) The light is free of UV and infra red rays.

(c) Intensity of light is low.

(d) Wavelength of light ranges from 5000 – 6000 Å.

(e) The light may be of different colours.

The annelids give blue light. Amongst fireflies, *Fulgora* gives white light, many species produce yellowish green but *Pyrophorus* gives out green light from thorax, and red light from abdomen. Fishes produce blue or bluish-green light. The Dragon fish, however, emits red light. It is known as dragon light. It is peculier that only dragon fish can perceive red light. This allows them to haunt unseen and find prospective mates without alerting their own enemies (Fig. 3.2).

Some organisms emit light continuously but some emit flashes that range in duration from about 1/10th of a second to 10 seconds.

Regulation of bioluminescence

Mechanical, nervous and hormonal processes control light production. In protozoa, mechanical stimulation by waves in the sea stimulates light production. In higher organism nervous regulation exist. In insects, when the nerves supplying photogenic organs are stimulated acetylcholine is produced. It reacts with ATP and CoA and forms acetyl CoA which in turn stimulates luciferase action. In fishes, photophores are under the control of Vth cranial nerve and spinal nerves.

Adrenaline has also been found to stimulate photophore activity.

Importance of bioluminiscence

1. Light flashes help the organism to find out prey, to escape from danger to puzzle and frighten their enemies, to attract opposite sex.

Many deep sea fishes illuminate the surroundings are thus help to find out preys. Many organisms attract their preys by light flashes e.g. deep sea fish, fireflies. Deep sea fishes can frighten their enemies by sudden flashes of light. Flashes of light may act as a mating signal.

2. There are several well known instances of bioluminescence during sexual behaviour. In some cases the light plays a part in the timing of reproduction and in synchronising the activities of males and males. The mating of fireflies often depends on a very accurate signaling system, the males flash their lanterns as they fly approximately 50 cm above the ground if a female sees one of these flashes within 3–4 meters, she may be expected to flash back after an exact interval of 2 sec at 25^0C. This attracts the male in her direction.

3. Luciferin – luciferase system is a highly sensitive test for detection and determination of ATP. In the laboratory pure firefly luciferin and luciferase are used to measure minute quantities of ATP by the intensity of light flash produced. As little as few picomoles (10^{-12} mole) of ATP can be measured by this processes.

In can also be used as a sensor for the determination of intracellular calcium concentration in the micromolar to nanomolar range.

4. Bioluninescence can help in detecting energy problems in human cells. This technique is now used to study abnormalities like ailmens of heart, muscular dystrophy, urological problems etc. It has been found that injection of luciferin and luciferase exhibit different reactions in normal and cancerous cells.

5. In pest management it is used as a tool for mapping the distribution pattern of organisms.

6. The bright luminescence generated from luciferase assay made it ideal for sensitive non-radioactive assay.

7. It helps in the detection of life in an unknown planet. Presence of ATP an indication of life, is ditected by bioluminescence technique in the soil of an unknown planet.

8. This technique is used in the study of genetics. It is used to detect the presence of some gene in a cell and to determine whether the gene is turned on or off. It has been shown that taking a single gene from the jelly fish aequorea and attaching it to the gene of another organism make the cell glow green when that gene is turned on. The jelly fish gene encode a protein called green fluorescent protein (GFP). When the jelly fish is disturbed, Ca^{++} bind to aequorin and produces blue light in the absence of GFP, but in presence of GFP green light is produced. The scientist Charles was able to get GFP to shine green in the absence of Ca^{++} and aequorin by simply shining a blue light on it. This discovery has a broad application in the area of genetics.

3.2 Pheromones and other semiochemicals

The term 'pheromone' is originated from the greek words 'pherin', to transfer, and 'hormone', to excite. It describes a class of chemicals that are communicated between animals of same species and that elicit stereotyped behavioral or neuroendocrine responses. Some pheromones, called releaser pheromones, elicit an immediate response. While others are termed as primer pheromones, which induce long-term changes in behavioral endocrine state.

Pheromone induced responses are mediated primarily through vomeronasal organ (VNO). It is also known as 'Jacobsons's Organ'. It is part of the olfactory system. It is present in a variety of non-human vertebrates as well as in humans. VNO was first discovered by Ruysch (1703). He was a military doctor. He found it in a soldier with a facial wound. It was named as Jacobson's organ after the name of Jacobson, who discovered it in animals. Since 1985, many scientists were able to demonstrate VNO in most humans. VNO is located bilaterally on the anterior third of the floor of the nasal septum. It opens into the nasal cavity by a pit which varies in size from 0.2 to 2 mm. It is situated 1-2 cm. from the posterior margin of the nostril. It is lined by a pseudo-stratified columnar epithelium, 60 μm in depth that lie on a thick basement membrane.

The VNO neuro-epithelium contains three cell types—

(i) Small polygonal dark staining basal cells measuring about 6 μm in diameter, called Basal cell.

(ii) Tall cylinder/columner cells with densely staining cytoplasm called Dark cell.

(iii) Tall columner calls like the dark cells, but lightly stained, hence called Light cells.

All these cells differ from surrounding respiratory epithelium in the nose by having no cilia.

Compounds occurring naturally on the human skin were found to cause a local depolarisation when applied directly to the VNO. This depolarisation had the characteristic of a receptor potential. Subsequently it was shown that these compounds were 16-androstenes and estrenes. These compounds did not produce response from olfactory epithlium. Moreover, olfactory stimulant (e.g. cineole) has no effect on VNO. A pheromone called vomeropherin (pregna-4, 20-diene-3, 6-dione (PDD), caused evoked potential in VNO and also changed gonadotropin pulsatility in males, resulting in a reduced level of LH and testosterone. PDD also decreased respiratory frequency, increased heart rate and also caused event related changes of EES pattern.

Stern and McClintock (1998) have shown that odourless axillary compound from the armpits of women in the late follicular phase of their menstrual cycle accelerated the preovulatory surge of LH of recepient women and shortened the menstrual cycles. Axillary compounds of the same donor collected later in the menstrual cycle (at ovulation) showed the opposite effects—they delayed the LH surge of the recepients and lengthened their menstrual cycle. Savic et al (2001) showed that 'androstadienone' (a human, in particular male secretion) caused activation of hypothalamus of women (gender-specific action), but not in men. It also activates anterior part of the Inferior Prefrontral Cortex (PFC) and the Superior Temporal Cortex (STP) (in addition to olfactory area). The PFC and STP have been implicated in aspects of attention, visual perception and recognition and social congnition.

A pheromone can act as a reinforcing agent or a one-trial conditioning agent in which the presence of a pheromone converts a second odour (that of the partner or infant) in conditioned stimulus. Mice can distinguish one another by odour. This odour is genetically determined and partly specified by the H-2 major histocompatibility complex (MHC) gene located on chromosome 17. Genes located on chromosome Y also regulate production of some odours. The human equivalent of MHC locus is HLA (Human Leucocyte Antigen). There are odourous substances secreted in the mouse urine. These odours many play a part in pregnancy block (Bruce effect), aggression and other mouse social behaviours. It is held that, in human, axillary odours have chemical differences which makes the discrimination possible. Some of these individual specific odours may be under the control of HLA genes. Studies have shown that women prefer those male odours that have HLA types different from their own. However, this preference is reversed if they use oral

contraceptives. Schaal (1980), reported that, mother could recognise their own newborn infant from the smell of a previously worn T-shirt. Infant also prefer breast or axillary pads from their own mothers, distinguishing the odour from other kin.

Receptors or at least their own RNA, for phenomones have been found in the human olfactory epithelium. However, presence of pheromonen receptor protein expressed in the surface membrane of an olfactory receptor neurone and the response of this receptor protein to a ligand (a potential pheromone), have not been convincingly observed as it is in case of olfactory receptors.

Androstenone is the male human pheromone, that helps to attract women. In practice, it is blended with favourite cologne to produce a cologne odour that will attract women. Similarly, androstenol is the female human pehromone, that helps to attract men. It has been observed that once men detect the female human pheromone scent, they subconsciously become more attracted, more receptive an more willing to offer attention. Regarding the mechanism of action, it is held that the chemical scent triggers the part of the brain, where the sexual attraction feeling starts. Once, the powerful feeling of sexual attraction sets in, it moves a person generally more attentive and responsive to the person, who is the source of that sexual attraction.

In humans, pheromone production is primarily linked with secretion of apocrine glands of skin, other glandular secretions and the moist areas of the body like axillae, mouth, feet and genitals. Freshly produced apocrine secretions are odourless. They are transformed to odourous products by microorganisms. The type and density of cutaneous microorganisms on different areas of the body interacting with skin and other glandular secretions give rise to a variety of odours from various body sites.

Types of pheromones

The pheromones are of different types like sex attractant pheromone, fear pheromone, aggresive pheromone, marker pheromoen etc.

1. Sex attractant pheromone : In case of Gypsy moth, the female secretes pheromones which is perceived by thousands of males with the help of their antenae which bear olfactory epithelium. Sensation followed the releaser pathway to manifest behavioral changes. This pheromone can attract thousands of males from a distance of half a mile. This has also been found in humans. It has been possible to synthesize it in the laboratory for use in pest control. Similar female sex attractant pheromones are also secreted by male animals, e.g. a secretion from salivary gland of male boar, a secretion from preputial gland of musk deer.

2. Fear pheromone : It has been observed that if an animal is frightened, fear pheromones are produced. If such a 'frightened animal' is put in a cage having normal animals, the latter are also frightened. These pheromones are secreted in the urine of frightened animals. If a smaple of such urine is applied to a normal individual in a small quantity, the normal animal get frightened.

3. Aggressive pheromone : It is particularly found in male mammals. If some adult and sexually mature bandikoots put in a single cage, they become aggressive to each other. They start fighting and eventually might die. This pheromone is secreted from prepuital gland.

4. Marker pheromone : Some animals secrete marker pheromones which help them to identify their territory. It is secreted from sebaceous gland of female blacktail deer and from anal gland of tigress.

3.3 Audio signal

The production and perception of sound is not only a means for communication it is also important for detection of environment and navigation.

Some animals have both sound production and sound detection systems, some may have either. Many animals can hear but cannot communicate with sound. A source of sound can serve as a warning even in those animals which cannot produce sound but is capable of hearing. In cases where sound is used for communication the production is often restricted to one sex, usually male and the principal function of sound is related to male attraction and maintenance of teritory. Animals use various types of sounds for various purposes but all are related to better survival. Insects can produce and perceive a wide variety of sounds which are used for communication often with the opposite sex.

Among invertebrates, communication through sounds are limited to a few groups of insects, especially Orthoptera, Hymenoptera, Cicadidae and perhaps some Crustaceans. The first two groups produce sound by stridulation, rubbing a toothed structure accross a ribbed plate. This results in the production of a burst of pulses extending upto 40 – 50 KC / sec. Cicadas produce sound by vibrating a thin section of cuticle, some decapod crustaceans make sound by thumping the substrate with their pinchers. Hearing is more widespread and is accomplished most commonly by a modified tracheal structure, the tympanic membrane. This is called tympanic organ. This may be located in the legs in Orthoptera, Arachnida and Crustacea in the thorax or abdomen as in Lepidoptera and Homoptera or in the antennae as in Diptera. The number of receptor neurons also vary for 2 in case of Moth to 70 or more in the Locust, but they all respond to the same frequencies determined by the properties of the tympanic membrane.

In these cases there may not be any frequency discrimination but sensitivity vaires and intensity discrimination is possible. The principal function of these organs is to detect a sound, its recognition apperently by the number of bursts per secound and its localization. Some insects respond to sounds in human audio range. They also respond to ultra high frequency sound as are produced by bats. Sound reception and

its processing help a moth to escape its predator bat. Cicadas sing species-specific songs conisting of complex pattern of clicks and rasps. Courtships and rivalry calls of male crickets can be identified as species specific, each produced by a patterned sequence of muscle contractions. Male lycosid spiders produce courtship and threat sounds by scraping the palps, the frequency of the courtship sounds increases when a female displays legwaving. Some species of bees signal in the hive, the location of feeding places; honeybees also produce sound as a part of communication from foragers to other workers; the sounds are produced by skeletal movements caused by contraction of wing muscles at 250 Hz. Duration of each sound has been found to be related to the distance between the hive and the food – duration is short if the distance is less and long if the distance is more. Sound vibrations can be received by Johnshon's organ present in flies. Vibrations in substrate water and air are important to insects as warning signals, aids to locomotor and posture control and communication. Sounds serve many orthopterans for social communication. Coding of sound in insects is mostly dependent on temporal patterning of pulses by amplitude and not much by frequency and harmonies as happens in case of vertebrates.

Vertebrates

Among the vertebrates, the most sophisticated use of sound for communication is performed by birds and mammals. However, other vertebrates also use sound in different functions.

In case of aquatic animals, fish have a variety of sound receptors like skin, lateral line, and three labyrinthine chambers. They use the first two to detect displacement and low frequencies. Several teleost fishes produce drumming granting or scraping sounds others, produce sound by vibrating there air bladder schooling calls and the sound possibly serve simple communicative functions specially in intraspecific alarm and in reproductive behaviour.

The sounds are mostly produced by contraction of muscles along the swim bladder or between the pectoral girdle and the swimbladder. A few kinds of fish produce sound by resonating the swim bladder. Elasmobranchs lack swim bladder, yet they respond to low frequency sound even after their lateral line is damaged. However, usefulness of sound is limited because most fishes have poor hearing capability. They respond only to low frequency upto about 1000 cps at high intensity. This is because they do not have a specialized cochlea or any means to concentrate acoustic energy on the otolith.

Some fishes have Weberian ossicles (Weberian apparatus) linking saccules and airbladder. Airbladder acts as a tympanic membrane, pick up vibrations in water. These fishes hear sounds upto several thousand cycles/sec and are several thousand times more sensitive than fish without weberian ossicles. Frogs emit mating calls which are species specific. Cetaceans can detect direction of sound in water. A

porprise (Tursiops) with its eyes covered has been found to emit sound pulses 1 – 1·5 msec duration, repeated about 16/sec while cruising ; and accelerating to 190/sec when near a reflecting object such as a fragments of fish. The whales emit sound which may constitute songs lasting several minutes and which are repeated many times. An individual whale may be distinguished by its song, the intensities are high. In the layer of water frequented by whales, the songs may carry for many miles.

Terrestrial vertebrates

The terrestrial vertebrates have found much more complicated uses of sound. Sound is produced by devices which are used in passing air via respiratory passages. The nature of the sound ranges from simple hiss of reptiles to the resonating vocal sacs of amplihia, the syrinx of birds and the larynx and vocal cords of mammals (with lips and tongue helping in man). However most of the anurans in amplibia do not make sounds, among the reptiles only the lizard of the family gekkonidae utilize sound for communication. Even in birds and mammals, nearly all can hear well, though there are some which are usually mute. Animals show maximum sensitivity within a restricted range of frequency. Frogs show maximum sensitivity between 3,000 to 4,000 cps. In the bull-frog response is greatest in the range of 100 – 200 cps, which is used as a mating call. Snakes and most lizards have less well developed ears and respond only to local sounds of a few hundred cps.

The hearing mechanism in birds and mammals are highly developed. Behavioral and electrophysiological studies revealed that they have same absolute sensitivity, differential frequency and intensity determination and dynamic range. The principal mechanism of sound perception in birds and mammals have been described. When the sound wave strikes the tympanic membrane, it vibrates. These vibrations are amplified and transmitted through the ear ossicles in the middle ear to the oval window of the internal ear. This causes vibrations of the basilar membrane. The auditory receptors present here are excited (hair cells in the organ of corti). As a result, nerve impulse is generated in the auditory nerve (cochlear division of VIIIth nerve) by auditory transduction. These impulses reach the auditory cortex and thus perception of hearing occurs. The perception of sound involves perception of pitch, intensity and direction of sound.

In mammals (except bats) cochlea is coiled. It has 2·75 turns in man, 3 turns in cat, 4 turns in pig and guineapig. Frequency perception range in man is 16 Hz to 20 KHz with maximum sensitivity between 1KHz – 3KHz. Some dogs can hear upto 35KHz and rats and guineapigs can hear upto 40 KHz. Primates hear better at low frequency than mammals such as opossum and hedgehog.

Significance of audio signal

In general, sound is a basis for species recognition, communication, predator detection and echolocation. Insects produce and perceive a wide variety of sounds

which are used for communication often with the opposite sex. Many fish produce sounds that are used for communication. Frogs emit mating calls which are species-specific. Baby mice and other rodents call their mother particularly in cold using ultrasonic calls of 60-90 Kc/sec.

A general function of calls in song birds is for the establishment of territory and for the attraction of a mate. Song-patterns have species characteristics that are genetically determined. Crows are gulls, have a variety of calls like— alarm, distress, assembly, chorusing are others. Cave dwelling birds use echolocation. Vocal signal is the best way for communication in humans. Vocal signals are used by them not only for communication, but also for scolding, quarrelling etc. Humans also use sound in the field of industry and medical diagnostic centres. Ultrasound is used not only for diagnosis, but also for treatment.

Infrasound

Recently it has been discovered that homing pigeons can detect very low frequency sound as low as 0.05 Hz (1Hz is 1 cycle per sec ; 0.05 Hz corresponds to 1 cycle per 20 seconds). This response to infrasound is lost or reduced if middle or internal ear is damaged. Such responsiveness is very important in birds because infrasound is produced during thunderstorms, earthquakes, and wind over mountain ranges. Since infrasound are attenuated much less in air (attenuation of sound is inversely related to the square of the wavelength), it travels over a long distance and thus can be detected hundred or even thousand kilometres away. This helps the birds to determine the direction.

The elephant can communicate with each other using infrasound. Because the infrasound undergoes very little attenuation in air, these may be audible for communication specilly in a forest habitat. Some animals show abnormal behaviour prior to earthwake and this may be related to perception to infrasound.

Perception of infrasound and its use for correcting direction and to make communication has been observed in birds and elephants.

Echolocation

The power to localize an object using the echo of a sound is called Echolocation. It is a complicated process. It is comparable with audiolocation devices and rader device invented by man. Animals that echolocate use echoes of sound they produce to locate the objects in its path. Some animals both aquatic and terrestrial are capable for echolocation. It is particularly well developed in bats but exists also in other animals notably whales, dolphins, shrews and a few birds.

Origins of echolocation

The ability to detect and understand the environment is essential for survival. Special sensory devices like vision, hearing, taste and smell can be used to know the environment. For locating an object in the environment vision/light and hearing

sound are very much helpful and in the dark, echolocation is the only correct device available for this purpose. Hence, animals which have to move in the dark have developed echolocation devices to locate an object in its pathway and this helps in capturing a prey in navigation.

Basic principles of echolocation

The animal produces a sound which strikes an object in its pathway and produces an echo. The echosound reflected from the object is detected by the auditory apparatus, analyzed and thus the position of the object is ascertained.

Echolocation in different animals

Sound waves travel at a slower rate in air compared to water. The intensity of sound is more quickly attenuated in air compared to water. High frequency sounds are more rapidly attenuated. The higher the frequency of a sound the shorter is its wave length. There is an inverse relationship between frequency of sound waves and the size of the object that can be detected by its echo. Hence, detection of small-sized object requires high frequency sound. Frequncy also influences directionality of hearing. High frequency sound is capable of detecting the direction of sound more accurately.

In view of these different properties of sound the members of animal kingdom adapted different mechanisms for echolocation depending on their habitat and requirement. All animals do not echolocate and different types of echolocating mechanisms have been observed. Echolocation has been found to be well developed and studied extensively in some animals including bats in terrestrial habitat and cetacians in aquatic habitat.

The bats emit echolocation signals through mouth (oral emitters) or nose (nasal emitters). Most microchiropteran are oral emitters. Only members of Nycteridae, Megadermatidae Rich Rhinolophidle and Phyllostomidae are nasal emitters. Usually high frequency and high intensity sounds are emitted. The detailed mechanisms of echolocation show variation in different genera of bats. Most bats emit 20 – 100 KHz frequency sound and the an intensity of 110 decibel (in bronze bat). Narrow (CF) and broad band (FM) signals are used. CF helps in detecting the object but cannot exactly localize it, this is done with the help of FM.

Insectivorous bats capture an insect by three phases of acoustical orientation. In the first phase sound pulses (100 – 20 KHz) with an interval of 50 msec is sent. When the prey is detected pulses are sent with shorter intervals and this is the second phase. In third on final phase a buzz like sound is emitted, the frequency is decreased duratian of pulse is decreased (0.5 msec) and the interval between two signals is also reduced. Finally the bat scoops up the insect with its wings or in the webbing between its hind legs guiding the insect to its mouth.

The quality, frequency, duration and number of occurrence of sound impulse per second show much variation in different genera of bats. The nature of the sound may

be 'click' type or 'buzz' type. In vespertiliomidaes the duration of pulse ranges from 1– 4 msec. The frequency may be upto 120 kc/sec but drops slowly. Rate of pulse discharge is less than 10/sec at rest but rises above 100/sec while hunting or avoiding obstacles. In horseshoe bats, the frequency of pulse is 85 – 100 kc/s, duration is high 40 – 100 msec. The pulse repetition rate is slow and it is less than 10/sec. If the size of the object is close to the wave length of sound wave, the reflection of sound is better. Thus the frequency of sound wave emitted by bats depends on the size of the object it has to be detected. To locate small size object higher frequency is required (the wave length of a sound of 30 KHz is about 11·5 mm which is roughly the size of a small moth)

There are bats which can avoid or dodge wires with a diameter of 0·5mm. This happens when the wavelength of sound is 5 mm. It has been observed that bats will not be able to avoid hitting a wire when its diameter is less than 1/10th of the wave length of sound.

Regarding detection of distance several theories have been proposed. It is proposed that distance is detected from the time difference between the transmission of ultrasound and the reception of its echo. Mohra proposed that distance is understood from the loudness of the object pye opined that the cochlea can understand the difference between two types of notes— the emitted note and the received echo note. Auditory nerves carry the information about the difference between these two notes to the brain. From this difference between these two notes, bats estimate the distance of object. It has been observed that vespertilioned bats can delect insects at a distance of 50 cm – 1 metre in the dark by echolocation. On the other hand the horseshoe bats can detect insects even at a distance of 6 meters. That echolocation is effected through auditory apparatus has been observed. Spallenzam did experiments with blind folded bats and observed that they can capture insects in the dark and the can return to their place of residence for a distance. Griffin noted that blind bats can capture insects which do not produce any sound and they can detect and chase pebbles and cotton spitballs thrown into the air.

The intensity of sound emitted for echolocation is very high and can be compared to the sound of jet engine which can damage the ear, but this is necessary because the intensity of sound drops faster in the air than water. It can damage the ear. A deafening mechanism has been observed in bats to handle the situation. When the sound is emitted the tensor tympani muscles attacted to tympanic membrane and the stapedius muscle attached to stapes contract. As a result the malleus is pulled inward away from the tympanic membrane and the stapes is pulled outward away from the oval window. In this way transmission of sound wave to the internal ear is reduced. However if this reduced phase is continued the reception of reflected sound will be disturbed. But this does not happen because these two muscles immediately after contraction relax and thus the ear osscicles get back to their specific position and sound transmission standard becomes normal.

The bats change their sound frequency during capturing prey of different sizes. They emit sound to 10 kc/sec (wave length 34 mm) to capture large sized prey and sound over 130 kc/sec with wave length of 5 mm to capture small-sized prey. By using echolocation, a bat can capture two seperate mosquitoes or fireflies in about 0·5 sec. The fish eating bats of Trinidad can even use echolocation to find and capture their under-water prey by detecting the ripples that are produced on the surface of water when a fish swims just under the surface of water.

Different experimental evidence indicate that a number of morphological and neuronal modifications help the bats in detecting the echoes. The snout in covered by couplex folds and the nostrils are spaced to produce a megaphone effect. The ear develop large pinnae which help to capture echoes. The eardrum and ear ossicles are small and light which provide high fidelity at high sound frequencies. Contraction of the muscles of ear ossicles reduce the sensitivity of the ear during the emission of high intensity sound; immediately the muscles relax and the capability of sound wave transmission reverts back to normal. (This is common in the ear of mammals). The bones housing the middle and inner ear are insulated from the rest of the skull by blood sinuses, connective tissue and fatty tissue. This reduces direct transmission of sound from the mouth to the inner ear. The auditory centres of the brain occupy a very large portion of the brain to receive the auditory signals and through the process of neural computation, construct from the auditory cues, a spatial representation of the external world.

Bat-moth predator—prey interaction

Experimental results indicate a remarkable series of adaptations by certain moths in response to predation by bats. Nocturnal of some families e.g. Nocteridae have developed sensitive auditory system to receive the frequencies emitted by bats. When these moths detect the approach of a bat, they alter their levels of flight, when the bats become nearer, the moths fly irregularly. When the bats become very close, the sound is very loud and repitition rate is also high. The moths dive directly on the ground. The moths themselves produce ultrasonic sound to detect the attacking predator. Some moths develop a noise-making organ on each side of the thorax. When the moths are disturbed, these organs produce trails of clicks with prominent ultrasonic components. It has been shown in laboratory experiments that flying bats turn away from their targets when confronted with moth-produced pulses. Hence, these pulses protect moths from the bats. The bats sometimes, in case of capturing moths, abandon echolocation for detecting the prey ; but instead listen to sounds produced by the prey. Echolocation signals of bat helps the moth to detect the foraging bats. Moths have been shown to be able to detect the cries of bats at a distance of 30 meters.

Echolocation in birds

At least two species of birds have been shown to use echolocation. The oil birds

(Steatornis) of South America and the cave swiftlets *(Collocalia)* of South-east Asia are not closely related but both live and nest in deep caves. The best known are the oilbirds or guacharos. They have been found to fly freely in the dark caves without hitting the walls or other obstacles. They use sound of 7000 Hz., which are audible by man. The nature of the sound they produce is like that of ticking of a typewriter. It has been observed that if ear of these birds are plugged, they fail to orient themselves in the dark. However, they are able to fly in a lighted room using their eyes.

Echolocation in aquatic animals

Echolocation has been observed in marine species. Tony shrews use a variety of sounds (30-60 kcps) to explore strange pieces of unfamiliar objects. Both dolphins and whales use echoes to avoid colliding with objects and with the ocean bottoms as well as for finding food. It is very useful in locating food deep under-water when the visibility is very low or zero. A trained dolplin has been found to locate a dead fish in the tank which cannot be seen in man. The dolphins produce sound by vibrations in the nasal sac system near the nasal plugs. That the source of sound is not larynx but nasal has been confirmed using imaging tools such as CAT, MRI, and RET scans. The nasal sac system consists of a series of muscular valves and compliant sacs associated with the blow holes. The muscles associated with these air sacs contract synchronously with the echolocation clicks, while the muscles around the larynx do not. A pair of small, dorsal fatty projection with a lip-like structure, called 'museau de singe' (also called monkey lips) control the passage of air through this system. Cranford hypothesized that the passage of pressurized air past the liplike structure (museau de singe) produce sounds in much the same manner as the glotis in man. It has been suggested that sound waves produced by the nasal sacs are focussed in the forward direction through a structure called melon (which is situated anterior to the monkey lips). It is a lens shaped fatty structure that gives a dome shaped profile to the forehead of many odontocetes. The lipid composition of melon has been analyzed and its acoustic properties suggest that it may serve as an 'acoustic lens' to focus outgoing energy. The echolocation capability of sperm whale has been inferred on the basis of data from other odontocetes. The click of a sperm whale consists of pulses. Clicks lasts for roughly 10 – 20 msecs, and the clicks are rapeated from less than one click per second to 40 per second. The sonar clicks are produced in the front of the sperm whale's had by pnematic action of the 'museau de singe' like that of dolphins. The sperm whale head acts as a sound reflector. Experimental evidence indicate that diving sperm whales use trains of clicks for echolocation of prey.

Some small cetaeceans that inhabit turbid water have tiny eyes and presumably are dependent on echolocation.

Ambient noise imaging (ANI)

It refers to the use of sound to see underwater. Taylor et al., using powerful computers and models inferred that dolphins might be able to detect using ANI. These models predict that dolphins use ANI to see useful images for tens of meters underwater. It has been suggested that dolphins and perhaps other marine mammals, have a whole new way of seeing with sound.

Multiple uses of Echolocation calls

Echolocation helps the animal to detect the direction, distance, size and texture of the objects in its environment. It is seen in birds and mammals. It is mostly developed in two groups of mammals. The microchiropteran bats are some cetaceans like porpoises and dolphins. Two groups of birds also use it.

Echolocation signals are used by bats for communication, as well as for orientation and locating prey. By modifying signal design some bats exploit habitats. Echolocation calls display considerable geographic variation within a species. *L. cinereus* in Arizona had calls with a minimum frequency 53·8% higher than the minimum frequency of calls in Menitoba (26·0 kHz versus 16·9 kHz). Eight of 12 species studied showed a difference of 3 kHz in different geographic location. Under certain conditions same bats add an extra pulse to the echolocation calls. It has been observed that when one bat is very close to other bats, the former lowers its call frequency and add a warning 'houk' to the signal.

Echolocation calls communicate several kinds of information. Same rely on signals of conspecifics to locate day roosts, mating sites, hibernation sites and feeding areas. The solitary forager *E. maculatum* in contrast reacted aggresively to the playbacks of calls of a conspecific individual. This bat either attacks the speaker or abruptly moves away.

Echolocation calls provide vocal signatures. This is the basis for mother-young recognition in a number of species. In a captive colony individuals recognize each other by listening to echolocation calls. Vocal signatures are individual interactions among foraging bats. A number of species are known to alter their signal design when changing from foraging in open areas to these closer to obstactes.

3.4 Suggested questions

1. What is audio signal ? Write briefly on its different uses.
2. Give an account of transmission and reception of audio signal.
3. What is echolocation ? Write briefly on the echolocation in aquatic animals.
4. Why high frequency sound is used for echolocation ?
5. Comment on the utility of CF and FM band in the echolocations.
6. Write briefly on bat-moth predator-prey interaction.

7. What is Ambient Noise Imaging (ANI)?
8. What is biioluminescence ? Write a note on its regulation .
9. Describe the reactions of bioluminescence with a note on the enzymes involved in the process.
10. Write a short account on the importance of bioluminescence.
11. What are phenomones ? Write a note on vomeronasal organ.
12. Write briefly on the different types of phenomenes.

Unit 4 ❑ Contractile elements, cell and tissues among different phylogenic groups

Structure

4.1 Muscle structure and function correlation

Muscle cells are specialized cells. They use ATP energy to generate force or to do work. Because work can take many forms like locomotion, pumping of blood or peristalsis, several types of muscles have evolved. The three basic types of muscles are skeletal muscle, cardiac muscle and smooth muscle.

Skeletal muscles are attached to bones, spans a joint and thus helps in joint movement and locomotion. They are voluntary. Their activity is controlled by nervous system. They play a key role in numerous activities like maintenance of posture, locomotion, speech, respiration etc.

Structure

Individual cells are of variable length. They may be as long as 25 cm, but their diameter ranges from 10-80 μm. Each muscle cell is covered thinly by a connective tissue layer called endomysium. Individual muscle fibres are then grouped together and get a covering of another connective tissue called perimysium. These are called fesciculus. Finally fasciculi are grouped together and they are covered by another connective tissue sheath, known as eipmysium. This epimysium covered structure is the muscle. At the ends of the muscle, the connective tissue fibres join to form tendon which are attached to the bones. All the muscle fibres are not of uniform length and extend the entire stretch. The fibres of shorter length become attached to the connective tissue inter lacing the muscle fibres. More than 600 muscles have been identified in the human body. Some are small and consists of few hundred fibres, but large muscles may contain several hundred fibres.

Each muscle fiber is multinuclated, bounded by the plasma membrane called sarcolemma. The protoplasm within is called sarcoplasm or myoplasm. Besides common organelles, the sarcoplasm contain myofibrils and sarcoplasmic reticulum. The myofibrils run along the length of the cell. They show alternate dark (Anisotropic) and light (Isotropic) bands. Proper alignment of the bands give rise to transverse bands and the longitudinal disposition of the myofibrils that gives rise to longitudinal

striations. These two together gives the muscle fibre a cross-sectional appearance. At the middle of the light band there is a dark line called Z-line. Thus a myofibril becomes made up of Z-line bonded units. These are called sarcomere. Average length is 2 μm. Each sarcomere is bonded at both sides by Z-line. At the centre there is A-band and on either side there is I-band. The A-band is made up of thick filament, called myosin. The I-band contains the filament called actin filament. At the middle of the H-zone, there is a fine dark line called M-line. The actin filaments extend from the Z-line to the border of the H-Zone. Thus there is an area where both A-band and I-band overlap, or myosin-actin overlap. This overlapping zone is called O-band. Besides actin many other proteins are associated with thin filament. These

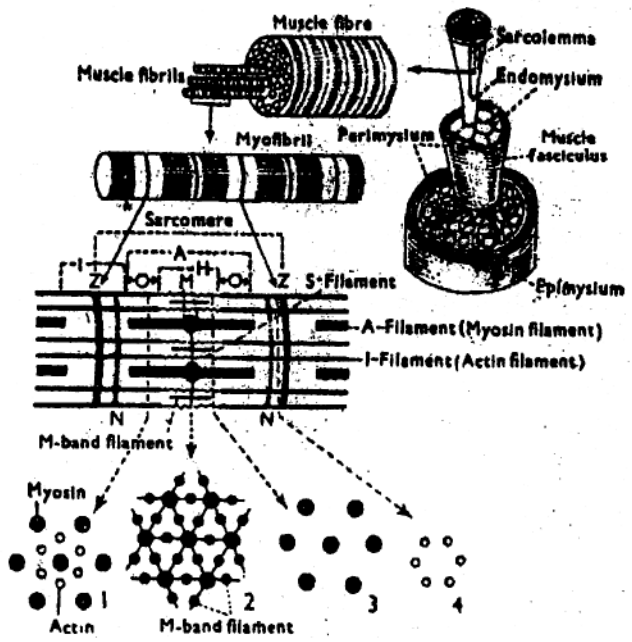


Fig 4.1 (a) : Anatomical organisation of skeletal muscle from gross to moelcular level (diagrammatic).

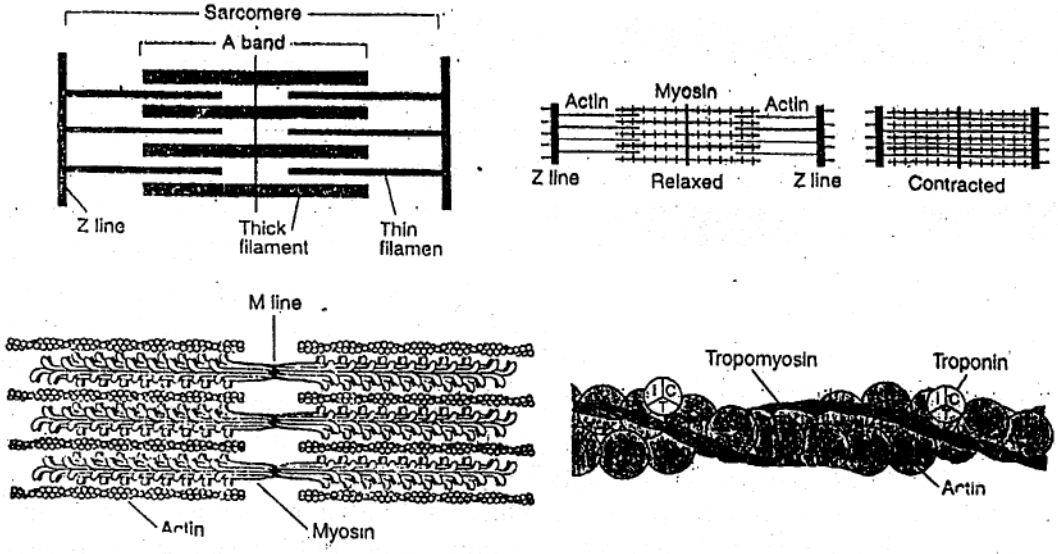


Fig 4.1 (b) : Top left : Arrangement of thin (actin) and thick (myosin) filaments in skeletal muscle. **Top right :** Sliding of actin on myosin during contraction so that Z lines move closer together. **Bottom left :** Detail of relation of myosin to actin. Note that myosin thick filaments reverse polarity at the M line in the middle of the sarcomere (Modified from Alberts B et al : *Molecular Biology of the Cell, 2nd 3d. Garland, 1989)* **Bottom right :** Diagrammatic representation of the arrangement of actin, tropomyosin, and the three subunits of troponin (I, C, and T).

include (a) Troponin, complex made up of 3 subunits, called troponin I, attached to actin, troponin T attached to tropomyosin and troponin C that attaches with Ca^{2+}. (b) tropomyosin, (c) tropomodulin, (d) a-actinin, (e) cap Z protein, (f) titin. Titin extends from Z-line to middle of the sarcomere and attached with myosin. It helps in proper alignmentof actin and myosin. Tropomodulin is located at the end of this filament towards the centre of the sarcomere and may participate in the setting of the length of this filament. α-actinin and capZ protein serve to anchor thin filament to the Z-line. Additional proteins present in the thick filament are myomesin and C-protein. These participate in the bipolar organisation or packing of the thick filament or both. Desmin binds Z-line to plasma membrane. Nebulin extends along the length of thin filament and may participate in the regulation of the length of thin filament (Fig-4.1).

Muscle contraction involves participation of (i) actin, myosin, troponin, tropomyosin system and (ii) sarcotubular system.

(i) The contractile proteins actin and myosin as well as other regulator proteins like tropomyosin and troponin are specifically arranged in a sarcomere and take part in muscular contraction.

Actin exists in two forms G-actin and F-actin. F-actin is arranged as a two-stranded hotical filament. Tropomyosin covers the actin filament. The actin molecule has myosin binding site. After every seven actin molecule, troponin is present. Troponin-I remains bound with actin, and troponin T with tropomyosin, and troponin C can bind with Ca^{2+}. The head of the cross bridge remains very close to the actin molecule. It has ATP-ase activity, ATP binding site and actin binding site. But it cannot bind with actin because myosin binding site of actin remains covered by tropomyosin.

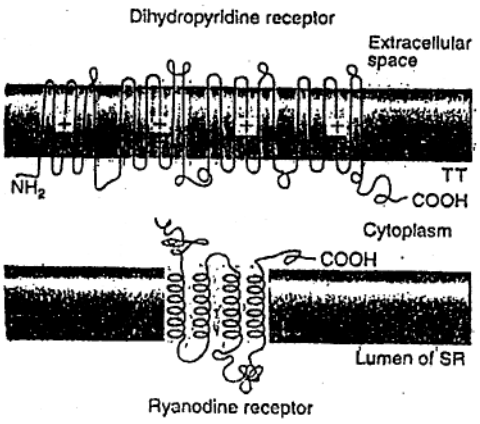


Fig 4.2 : Showing DHP and RYR receptors.

(ii) The Sarcotubular system essentially consists of T-tubule which is an invagination of sarcolemma, and longitudinal tubules running parallel to myofibrils. These tubules end in a dialated structure near the T-tubule, called cistern. One T-tubule and two longitudinal tubules with two cisterne form a triad system. In the T-tubule there is dihydorpyridine receptor (DHP receptor) and on the cisterni there is ryanodine receptors which is a calcium channel. Ca2+ remain stored in the cisterni by combining with calsequestrene. In each sarcomere there are two triads each located at the A-I junction (Fig. 4.2).

In the resting state intracellular Ca^{2+} concentration is low, about 10^{-7} moles/litre. ADP+P ; are attached to myosin head (cross bridge). On stimulation, muscle action

potential develops. It passes along the sarcolemma and goes deep into the cell via T-tubule, and it excites DHP receptor. When it is stimulated, the RYR receptor is excited through protein-protein interaction (Fig 4.3).

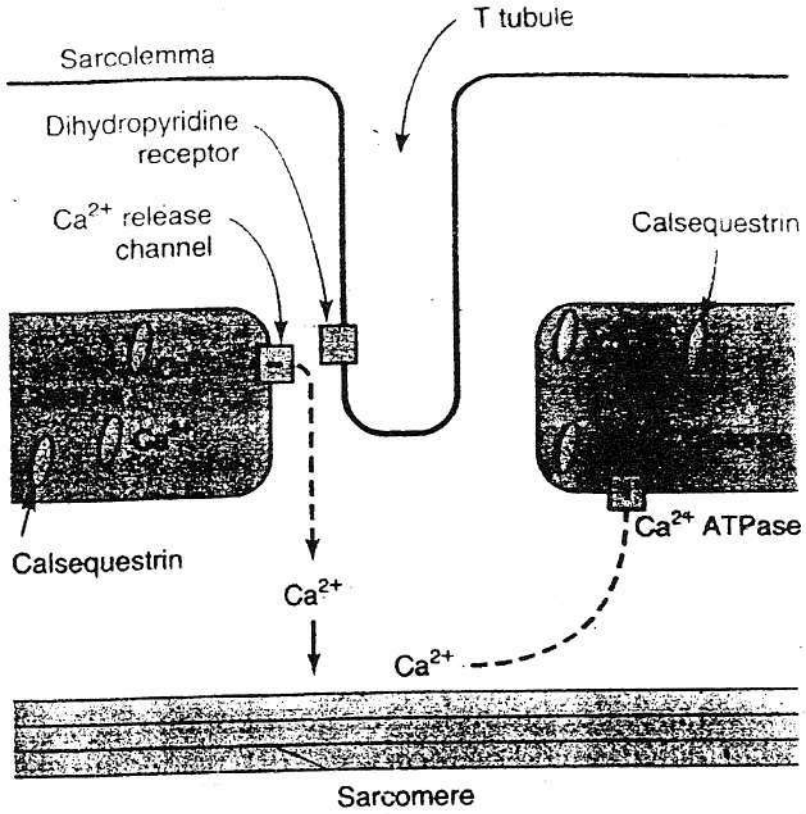


Fig 4.3 : Diagram of the relationship among the sarcolemma (plasma membrane), a T tubule, and two cisternae of the sarcoplasmic reticulum of skeletal muscle (not to scale). The T tubule extends inward from the sarcolemma. A wave of depolarization, initiated by a nerve impulse, is transmitted from the sarcolemma down the T tubule. It is then conveyed to the Ca^{2+} release channel (ryanodine receptor), perhaps by interaction between it and the dihydropyridine receptor (slow Ca^{2+} voltage channel), which are shown in close proximity. Release of Ca^{2+} from the Ca^{2+} release channel into the cytosol initiates contraction. Subsequently, Ca^{2+} is pumped back into the cisternae of the sarcoplasmic reticulum by the Ca^{2+} ATPase (Ca^{2+} pump) and stored there, in part bound to calsequestrin.

Excitation of RYR receptor causes opening of Ca^{2+} channel and Ca^{2+} is released. It binds with troponin C, and causes a conformational change of it. This leads to a movement of tropomyosin. As a result, the myosin binding site of actin molecule is exposed. Next, myosin head gets attached to actin. When actin attaches to myosin, the myosin ATPase activity is increased. So ADP and Pi leaves myosin and actin-myosin interaction occurs. The myosin head binds and so the actin molecule is pulled towards the centre of the sarcomere. ATP then gets bond with myosin and due to this binding actin can no longer remain bound with myosin. Actin is removed and myosin head again goes back to its previous state and gets ready to bind with another actin

molecule. Meantime, ATP is split into ADP+Pi by the ATPase action of myosin-head. But the ATpase action is low, so ADD+Pi do not leave the myosin head, rather remain attached to it. Actin-myosin recycling occurs as muscle contraction occurs.

Intracellular concentration of Ca^{2+} does not remain high for long time. Immediately after its release, it is taken back into longitudinal tubule by active process. As a result, intracellular Ca^{2+} concentration is decreased. As it goes to normal level, 10^{-7} moles/lit, Ca^{2+} leaves troponin C and goes back to tubule. When troponin C becomes Ca^{2+} free, again a conformational change occurs in troponin molecule. As a result, tropomyosin moves and covers the myosin touching site of actin molecule. Here, actin-myosin interaction does not occur and muscle contraction stops and the muscle relax.

During muscular contraction, the muscle fibre is shrotened. The actin filaments slide into and in between myosin filaments, and during relaxation the actin filaments slide out of the myosin filaments. The structure of sareomere has been observed during rest and during contraction.

Cardiac muscle

The cardiac muscle cells are short cylindrical in shape. Typically they measure about 100 μm long and 10 μm in diameter, they are uninucleated. They are cross striated like skeletal muscle because of specific alignment of thick and thin filaments. Especiality of cardiac muscle is that, the cells are conneted with each other through intercalated discs. These are low resistance partitions like gap-junctions, Junction adherens. Through these, signal is transmitted from one cell to the next, and thus a functional syncitium is established.

The sarcotubular system is well developed like skeletal muscle. It has triad and diad systems. But each sarcomere has one, because it is locked at the Z-line. The ryanodine receptor has a special property leading to a situation called calcium induced calcium release. When action potential passes down the T-tubules, the DHP receptor is activated. As a result voltage gated Ca^{2+} channels present here open, and ECF Ca^{2+} enters into the cell. This calcium binds with RYR receptor. Such Ca^{2+} binding opens the calcium channel present here. This is called calcium induced calcium release. Thus intracellular Ca^{2+} concentration is increased and goes above 10^{-5} moles/lit. Now, Ca2+ binds with troponin C and finally muscle contraction occurs in the same way as happens in skeletal muscle. It has been shown experimentally that if cardiac muscle fibres are placed in a medium muscle contraction will occur if the bathing medium contains Ca^{2+}. But this is not required to cause contraction of skeletal muscle. The mechanism of muscular contraction is also similar to skeletal muscle that is actin regulated or actin based muscular contraction.

Smooth muscle

Smooth muscle cells are uninucleated, fusiform in shape with tapering ends.

These are 400-600 μm long. Their diameter ranges from 2-10 μm. The thick and thin filaments are 10,000 times longer than their diameter and are tightly packed.

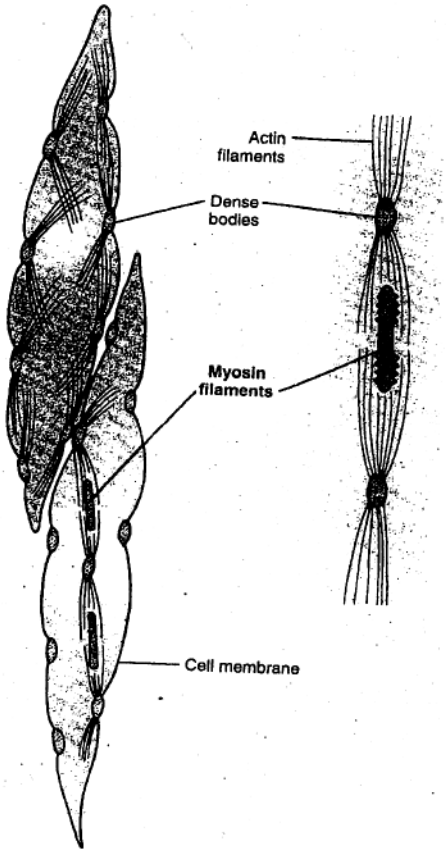


Fig 4.4 : Physical structure of smooth muscle. The upper left-hand fiber shows actin filaments radiating from dense bodies. The lower fiber and the right-hand insert demonstrate the relation of myosin filaments to the actin filaments.

The smooth muscle has been divided into two groups : single unit and multi-unit. In single unit muscle, the muscle cells are electrically coupled, through special junctions. These are gap junctions and adherens junctions, connecting the adjacent cells. Through these, signals (action potential) can pass from one cell to the next. A wave of electrical activity and a wave of contraction occurs and it can be initiated by pacemaker cell (a cell that exhibit spontaneous depolarisation).

Because of syncitial arrangement, they are called cardiac muscle type. In case of multi unit smooth muscle, the individual cells are not electrically coupled, and each cell can contract independently.

Scrcotubular system is very weakly developed. The sarcolemma shows small invaginations known as caveoli. These represent the rudimentary form of T-tubules of skeletal muscle cells. Just behind caveoli there are small fragments of longitudinal tubules. All these represents rudimentary form of sarcotubular system.

The dispcsition of actin and myosin are also not well organised as seen in skeletal and cardiac muscles. There are dense bodies attached to sarcolemma, some are also present in the sarcoplasm. They represent Z-line of the sarcomere. Actin filaments are seen radiating from dense bodies. Myosin filaments are seen between two dense bodies (Fig 4.4).

Mechanism of contraction of smooth muscle

In the smooth muscle cells troponin is absent. A number of proteins and enzymes are involved in its contraction. In the resting state, actin-myosin interaction does not occur and the muscle remain in relaxed state. A protein 'caldesmon' present in the

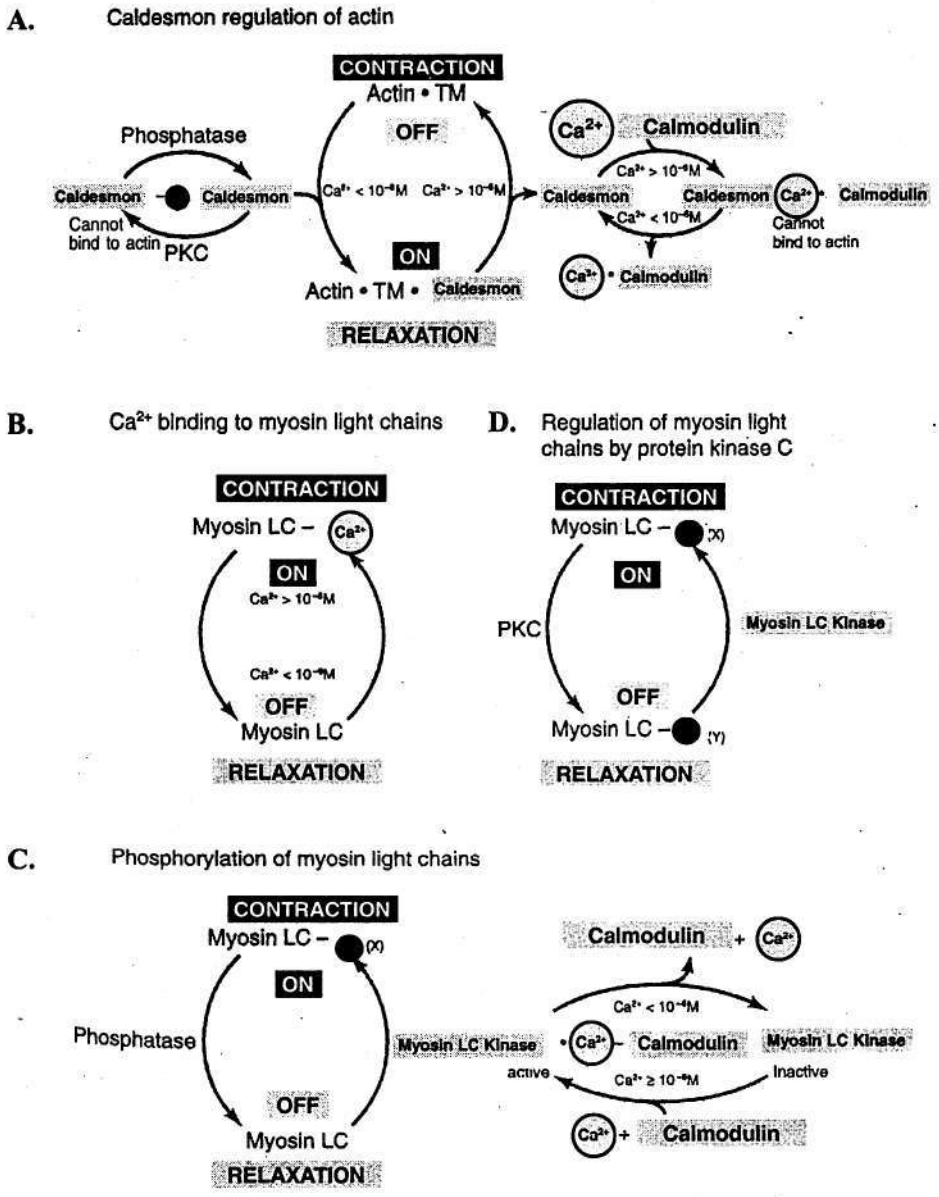


Fig 4.5 : Both actin-and myosin-dependent mechanisms control smooth-muscle contraction and relaxation. (A) Binding of caldesmon to the actin and tropomyosin (TM) of thin filaments prevents contraction. At cytosolic Ca^{2+} levels above 10^{-6} M, formation of the Ca^{2+}/calmodulin complex occurs. Binding of this complex to caldesmon releases it from thin filaments, allowing the muscle to contract. Phosphorylation of caldesmon by protein kinase C (PKC) also prevents it from binding to thin filaments and promotes contraction. (B) Binding of Ca^{2+} to the regulatory light chains of myosin allows actin-myosin interactions and promotes contraction. (C) Phosphorylation of the regulatory light chains by myosin LCkinase, which is activated by Ca^{2+} calmodulin, also promotes muscle contraction. (D) Phosphorylation of the regulatory light chains by protein kinase C, at a site other than that acted upon by myosin LC kinase, inhibits myosin-actin interactions and causes smooth-muscle relaxation. [Adapted from Lodish et al., 1995.]

sarcoplasm binds with the actin and thus actin-myosin interaction is kept prevented.

When intracellular Ca^{2+} concentration is increased and goes above 10^{-5} moles/lit, Ca^{2+} binds with a protein in the sarcoplasm called calmodulin. This Ca^{2+}— calmodulin complex binds with caldesmon and it is released from actin. So, actin-myosin interaction occurs and the muscle contracts. A fall in intracellular Ca^{2+} causes decreased formation of calmodulin-Ca^{2+} complex. So caldesmon becomes free. This now again binds with action and prevents actin-myosin interaction. This results in relaxation of the smooth muscles (Fig 4.5).

Besides this, there are other mechanisms also, that operates or influence the process. There is a regulatory light chain of myosin which remain in two form— phosphorylated and dephosphorylated forms. The dephosphorylated prevents actin-myocin interaction and so the muscle remain in relaxed state.

Calmodulin Ca^{2+} complex activates an enzymes, called myosin light chain kinase (MLCK). This enzyme phosphorylates myosin light chains. The phosphorylated form of myosin light chain cannot prevent actin-myosin interaction and so muscles contract. There is a myosin light chain phosphatase that dephosphory lates myosin light chain. But, it is again phosphorylated immediately by the kinase, so long calmodulin—Ca^{2+} complex is there. However, when Ca^{2+} concentration is decreased, Ca^{2+} ion are removed from calmodulin Ca^{2+} complex. So, kinase is no longer activated. So, further phosphorylation of myosin light chain does not occur and the muscles relax.

PKC can cause phosphorylation of caldesmon and thereby can also prevent actin-myosin interaction. However, it can also phosphorylate the regulatory light chain of myosin and this also inhibits actin-myosin interaction and the muscle relaxes. Both PKC and MLCK phosphory lates myosin light chain (the regulatory light chain of myosin), but the phosphorylation sites are different and the action is also different. MLCK phosphorylation leads to contraction and PKC phosphorylation causes relaxation of the smooth muscle.

Sequence of events in muscular contraction

The general sequence of events in the contraction of three different types of muscles. The skeletal and cardiac muscles undergo action based muscular contraction, whereas the contraction of smooth muscles is myosin regulated.

Sequence :

1. When the muscle fibre is stimulated via neurons (natural stimulus) or otherwise action potential develops.
2. The muscle action potential passes via sarcolemma and enter deep into the muscle cell via T-tubules (in skeletal, and cardiac muscle), and via DHP receptor and RYR receptor escitation causes release of calcium. In smooth muscle, mostly Ca^{2+} comes from ECF due to opening of voltage gated Ca^{2+}

channel present in the plasma membrane of smooth muscle cell.

3. Intracellular Ca^{2+} concentration is increased. It binds with troponin C (in case of skeletal and cardiac muscle cells). This causes a conformational change in troponin molecule. As a result, tropomyosin moves and the myosin binding site of actin is exposed. Actin myosin binding occurs and the muscle contracts. In smooth muscle cells, troponin is absent. Here, Ca^{2+}-Calmodulin complex is formed, that initiates muscles contraction involving caldesmon, MLCK, PKC, muscle phosphatase. (Fig 4.6)
4. As intracellular Ca^{2+} concentration is decreased, Ca^{2+} dissociates from troponin

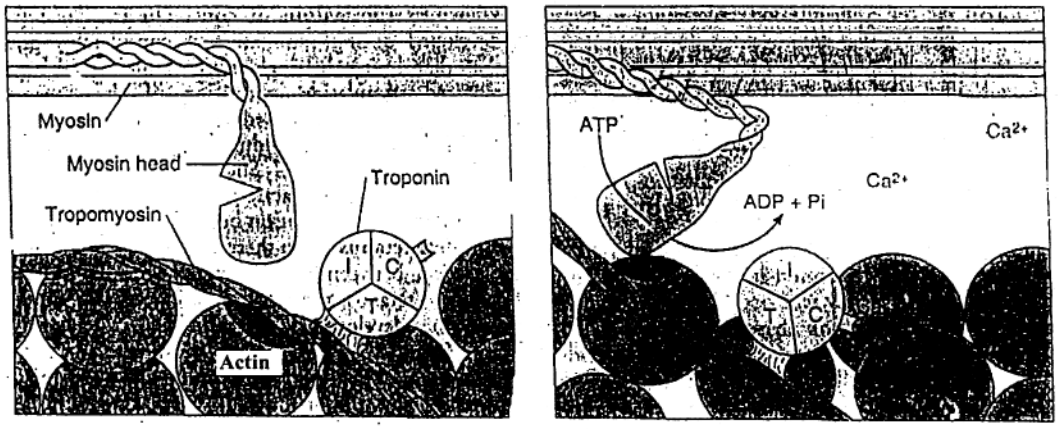


Fig 4.6 : Initiation of muscle contraction by Ca^{2+}. When Ca^{2+} binds to troponin C, tropomyosin is displaced laterally, exposing the binding site for myosin on actin (dark area). The myosin head then binds, ATP is hydrolyzed, and the configuration of the head and neck region of myosin changes. For simplicity, only one of the two heads of the myosin-II molecule is shown.

C and Ca^{2+} calmodulin. This leads to relaxation.

Basically the (i) development of action potential, that is, degree and duration of depolarization and repolarisation, (ii) functioning of the sarcotubular system, that is sensitivity of calcium-associated with degree and duration of calcium release and reuptake, (iii) actin-myosin interaction that is degree and duration of action and myosin attachment and detachment cycle, (iv) energy yielding system in the muscle cells, are the determinants of muscle function. These muscles have been selected based on their structure and physiological propertiesfor placement in specific areas of the body for performing specific function. (Fig. 4.7)

In case of skeletal muscle, the muscle fibers are not inter-connected. So, all or none law is applicable to individual fibere. So, when it contracts, the strength of stimulus can produce summation or tetanus. But in cardiac muscle, the muscle cells are electrically connected resulting functional syncitium. Here all or none law is applicable to whole of heart. The refractory period is short in skeletal muscle. So

summation and tetanus are possible but this does not happen in case of heart, because it is absolutely refractory during systoleand relatively refractory during diastole. The heart cannot afford summation, tetanus or fatigue. Hence, its muscle fibres have such property. The single unit smooth muscle is called cardiac muscle type and multi unit smooth muscle is termed as skeletal muscle type, and their function is also different.

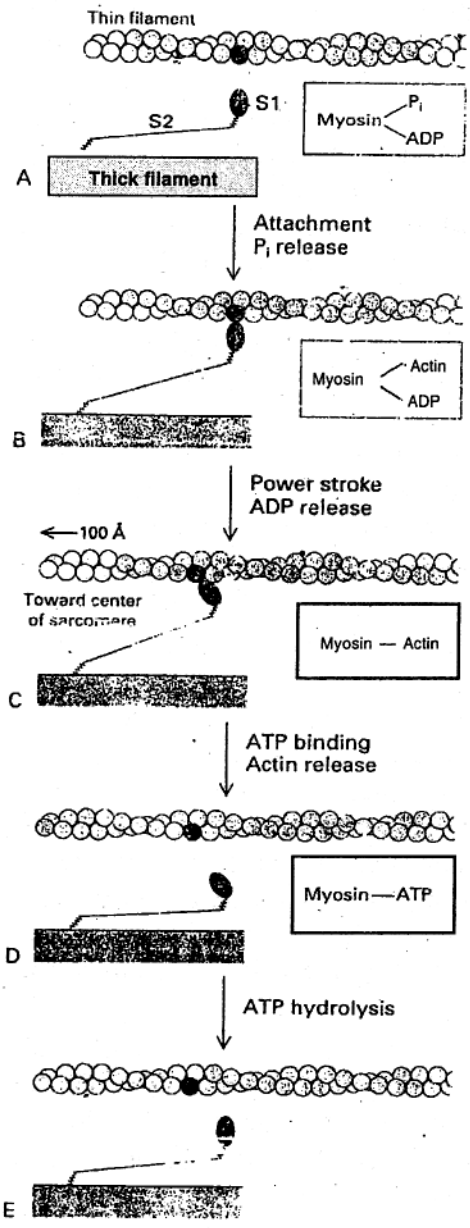


Fig 4.7 : Proposed mechanism for the generation of force by the interaction of an S1 unit of a myosin filament with an actin filament. In the power stroke, the thin filament moves relative to the thick filament when S1 undergoes conformational changes accompaying the release of ADP.

The action potential duration is short in skeleted muscle, but prolonged in cardiac and smooth muscle. Mechanical change is related to electrical change. Hence, the twitch duration is short in skeletal muscle but is is more in caridac muscle and more prolonged in smooth muscle.

Skeletal muscles are attached to skeleton and joints. These muscles take part in movement and locomotion. The muscles are so arranged that they can provide optimum output. Each muscle has origin and insertion and as they contract, the insertion part moves towards the origin. Depending on this property the muscles are arranged accordingly so that the purpose is served. For example, the biceps muscle has origin in the scapula and insertion in the radius, thus in contraction causes flexion of the elbow joint. The muscle of the rib are so arranged that contraction of external intercostal causes elevation of the ribs, thus increases the antero-posterior diameter of the thoracic cavity. Similarly, contraction of diaphragon muscle causes its downward movement towards the abdominal cavity, and thus superior-inferior diameter of the thoracic cavity is increased. In this way, as the diameter of thoracic cavity is increased, the intra-thoracic pressure is decreased. This is turn decreases intra-pulmonary pressure and so air enters into

the lungs and inspiration occurs.

Force is generated during musculer contraction and this forece is transmitted via tendon to the bones, within the muscle, there are elastic fibres in two forms— (i) parallel elastic component—the fibres lie or run in between muscle fibres. (ii) Series elastic components—these represent the elastic fibres that connect the contractile element with the tendon that is attached to the bone. These are involved in producing isometric and isotonic contractions. Isometric contraction is that where tension is developed but the muscle is not shortened. eg. holding a weight on the palm of a stretched hand. Similarly, an isotonic contraction is that in which muscle stretches but tension is unchanged, eg. lifting a load by flexing the elbow joint. The biceps muscle contract isotonoically as the load is lifted. Because of series elastic element, the force generated during mascular contraction can be transmitted to the bones via the tendon.

Sometimes, the muscles are attached to a long tendon that in turn is attached to the bone. eg. the muscles that cause movement of the finger-joints are present in the lower arm. Such arrangement definitely help in smooth movement of the fingers.

The force generated during muscle contraction is directly related to actin-myosin overlap. It has been shown that the more the overlap, the more the tension development. A study of length-tension relationship revealed that as the length of the fiber is increased the tension developed is dependent of actin-myosin overlap. At longer length, the tension is small and it is zero, when there is no overlap (Fig 4.8). When the length is gradually decreased, the overlap is slowly increased, and the tension gradually rises. It becomes maximum (plateau phase) when there is maximum actin-myosin overlap. A further decrease in length of the fiber again causes a decrease in tension, because overlap of actin retands tension development. Frogs use hip flexor muscles during jumping. It has been found that during jumping the sarcomere length changes from 2.3 μm (at rest) to 1.82 μm at the point of take off, and this length has been found to correspond to the plateau-phase of the length tension diagram of the hip-flexor muscle.

The red and white muscles are used according to the type of contraction required. For quick movement the white type is used and for sustained movement red muscle are used. It has been shown in fish that during steady swimming the sarcomere length of the red muscles coincide with the plateau phase of length-tension curve for this muscle, whereas in case of escape response, this is seen in case of white muscles. This proves which muscles are used for what kind of movement. Biophysical and biochemical characteristics confirm their suitability for doing sustained or quick movement. Thus animal body uses specific muscles for specific purpose.

It has been shown that as the length of the muscle fibre is increased, the tension development during contraction also increases progressively and becomes maximum at a particular length. It is called resting length of the muscle fibre. At this length the

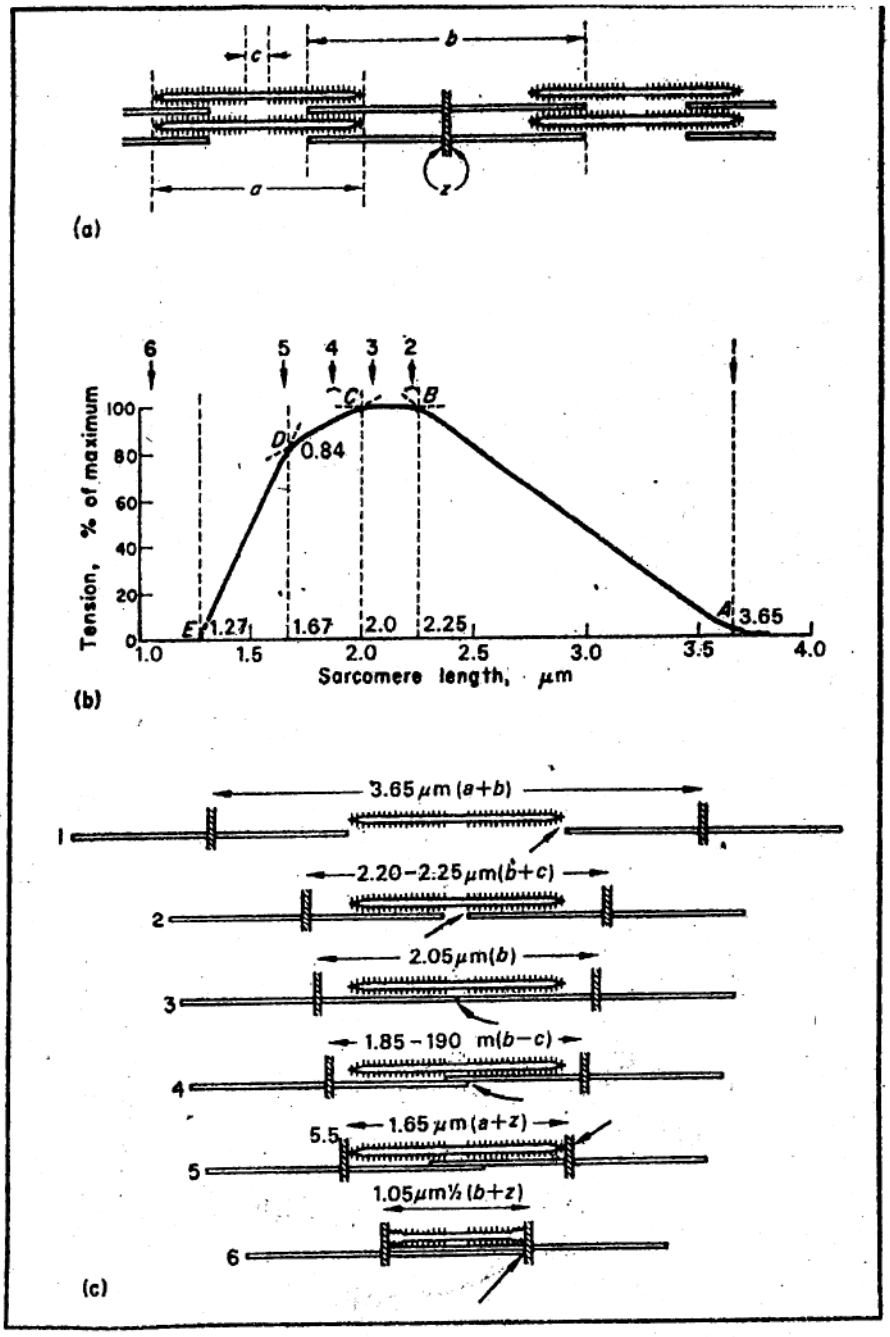


Fig 4.8 : Standard filaments lengths. a = 1.60 μm; b = 2.05 μm; c = 0.15—2 μm; z = 0.05 μm. (b) Tension-length curve from part of a single muscle fibre (schematic summary of results). The arrows along the top show the various critical stages of overlap that are portayed in (c). (c) Critical stages in the increase of overlap between thick and thin filaments as a sarcomere shortens. (Gordon, A.M., Huxley, A.F., and Julian, F. J. (1966). *J. Physiol., Lond.* 184, 170.)

actin-myosin overlap is maximum. If the length is further increased the tension development is decreased. Survey report revealed that the muscles present in the body during rest are at the resting length.

The sound producing muscles (sonic fibres) contract 10-100 times faster than those used for locomotion. Experimental results indicate that this has been made possible by a number of adaptations. These are— (i) an increase in Ca^{2+} kinetics ; (ii) cells are well equipped for aerobic metabolism and the space required for such assembly is made by reducing the number of myofilaments. The animal can afford such reduction in myofilaments in sonic muscles, because the muscles require faster rate and not increased force.

The force or power production during muscular contraction is denoted as V/Vmax, where V is the velocity of muscular contraction at any particular condition and Vmax is the maximum velocity of shortening possible. The power production is considered to be maximum, when the value of V/Vmax is 0.15—0.40. For frog hip flexor muscle during jump V/Vmax is about 0.32. Hence it is said the muscles while lifting the body they try to do it with maximum power.

The skeletal muscles of the vertebrates consist of muscle fibrets of more than one type. Some contain a high proportion of tonic fibers which show steady contraction other muscles contain a high percentage of twitch fibres which are specialized for rapid movements. Such muscle fibres having different properties have been found in the animal body and these are due to biochemical, metabolic and other structural adaptations.

Tonic muscle fibers contract very slowly and do not produce twitches. The motor neuron make contact with the muscle fiber at several points. Hence, action potential is not produced, and in fact they do not require AP to spread excitation. The actin-myosin detach slowly, hence the velocity of shortening is slow. So they are able to generate isometric tension very effectively. They are capable for slow steady contraction. Hence they are used for posture maintenance, where a slow sustained contraction is required. They are found in the postural muscles of amphibians, reptiles, and birds as well as the muscle spindles of all extra ocular muscles of mammals. Slow twitch or type I fibere are characterized by slow to moderate Vmax. Slow Ca^{2+} kinetics. They contract slowly and fatigue slowly. They generate 'All or none' AP. It has one of few end plates. It contains myoglobin. Muscles that contain more of these fibers are also called red muscles. The slow-fatigue is due to presence of large number of mitochondria and rich blood supply. These fibres are suitable for maintaining posture and for moderately fast repititive movements.

Fast twitch oxidative (type IIa) fibres have high Vmax. They acn be activated quickly. They fatigue slowly because they have more mitochondria, can produce ATP quickly by oxidative phosphorylation. They are specialized for rapid repetitive

movements as required in sustained streneous location. For these reasons, they are used for making the flight muscles of wild birds.

Fast twitch glycolytic (type IIb) fibres. These fibers have a high Vmax, very rapid Ca^{2+} kinetics and for the this they get activated and relax quickly. They have few mitochondria. Hence they are dependent on anaerobic glycolysis to genetate ATP. Hence, they fatigue quickly. These are found in breast muscle of domestic fowl, which are never used for flying and cannot produce ssustained activity. Ectothermic vertebrates, such as amphibians and reptiles also make extensive use of glycolytic muscle fibers.

Thus, the animal make up their muscle by different types of muscle fibers according to their requirement. Body not only can improve their muscle mass by growth of individual muscle cell, but they can also improve their power for performance by forming oxidative enzymes to improve aerobic power and by improving glycolytic enzymes to improve anaerobic power. Moreover, by changing to fast twitch fibers and vice versa. This is mediated via changes at genetic level.

Cardiac muscle cells are interconnected functionally. Here, the whole heart obeys all or none law. However, it has been observed that in a quiscent heart repeated stimulation increase the force of contraction for few strokes.This is known as stair case phenomenon or treppe. It is due to increased accumulation of Ca^{2+} that increases the force of contraction of cardiac muscle. At every situation Ca^{2+} is released but in case of repeated stimulation the rate at which it is released can not cope with the removal. So Ca^{2+} accumulates and produce the effect.

It has also been observed that when the cardiac muscle is stretched its force of contraction is increased, as happens in case of skeletal muscle. But the mechanism in cardiac muscle is different at the level of actin-myosin interaction. It has been shown that at saturating Ca^{2+} ion concentration stretched cardiac muscle exhibit greater force of contraction compared to control cardiac muscle. It is not due to difference of overlap of thick and thin filament as happens in skeletal muscle. Evidence suggest that stretch reduces the space between thick and thin filaments (ie., interfilament spacing) as this is associated with the ability of more myosin molecules to interact with actin.

Titin plays an important role in it. Titin binds to both actin and myosin in such a way that when the muscle is stretched, it brings actin more close to the myosin head and they increase the number of myosin heads that interact with actin. It has been found that protolysis of titin attenuated length dependent increase in force. Increased sensitivity to Ca^{2+} is possibly due to decreased interfilament space caused by titin (Fig 4.9).

Increased venous return increases cardiac output. This is possible because of the property of cardiac muscle. As more blood enters the heart (ventricle), the volume of the heart (ventricle) is increased, that is the heart muscle is stretched and this increases the force of contraction of heart muscle. So the blood goes out of the heart.

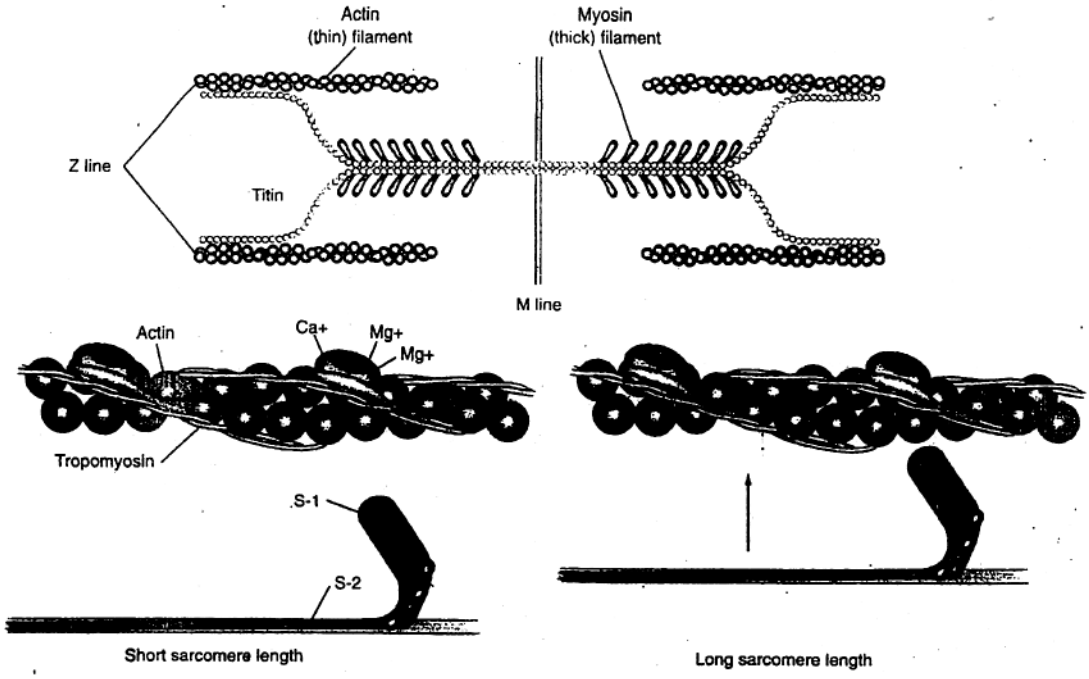


Fig 4.9 : Titin may contribute to the ability of stretch to increase the force of contraction of the heart. Titin binds to both myosin and actin such that stretch of the cardiac muscle may bring the actin filament closer to the myosin head and thus increase the number of myosin heads that interact with actin at a given intracellular [Ca++]. (Redrawn from Moss RL, Fitzsimons DP : Circ Res 90:11-13, 2002.)

In smooth muscle cell actin-myosin recycling occurs. Myosin cross-bridge head attaches itself to the actin, then it is released from actin and then reattaches with next actin in the next cycle. This rate of cycling is very slow. This is because ATPase activity of myosin is very low compared to skeletal muscle. Moreover, in case of tonic contraction, it has been found that, when full contraction has developed, it can be maintained with less ATP utilization. During actin-myosin attachment a great force is generated and this allow the smooth muscle cells to contract as much as 80% of its length (instead of 30% as seen in skeletal muscle). This can cause a great collapsing of the hollow tube.

The response to a sustained or tonic stimulation is a rapid contraction followed by a sustained maintenance of force with reduced cross-bridge cycling rates and ATP consumption. This behaviour is called latch state. It is advantageous for muscles that may need to withstand continuous external force, such as blood vessels, which must be able to withstand blood-pressure. ATP consumption during latch state is less than 1/300 the rate that is necessary for skeletal muscle to maintain the same force. Another important characteristic of smooth muscle cells is length adaptation. It is able to adjust length-tension relationship when chronically stretched or shortened.

This is very helped during filling of urinary bladder. Hence bladder wall is made up of such smooth muscle cells.

It has been observed that a sudden increase in bladder volume causes stretching of the smooth muscle and the pressureof uninary bladder is increased but subsequently within 15 sec to one minute despite continuous stretch the pressure becomes almost back to normal. This has been named as stress relaxation. Because of this, a hollow organ can maintain almost the same pressure in spite of changes in the length of smooth muscle.

The single unit smooth muscles are present in the GI tract. The multi unit smooth muscles can be found in vas deferens and iris muscle.

4.2 Movements—ameboid, ciliary and flageller

Amoeboid movement

This terms derives its name from motion of Amoeba. There are two terms, 'movement' and 'locomotion'. An animal being attached to a site can show a great variety of movement, but the animal does not shift to another place. It is called movement eg, corals and sponges. But when the animal shifts its position from one place to another by movement it is called locomotion. There are animals which have cilia and flagella. These also show movement and movement of these can help the animal to move from one place to another. Thus, ciliary movement and flageller movement are not only related to 'movement', but also to 'locomotion'.

A typical Amoeba consists of an outer layer, the plasmalemma. It is not water wettable and has adhessive property. It slides freely over the next inner layer called ectoplasm (gel-like). The plasmalemma is made up of outer filamentous coat and an inner membrane. The filaments are about 80Å in diameter, and extend outward 0·1 to 1·0 μm. The filamentous coat is made up of 35% lipid, 26% protein and 16% polysacharide.

Beneath the plasmalemma there is a hyaline layer. It is fluid in nature as judged by brownian movement. This layer is very thin in the region of attachment to the substrate. It is often thickened as an hyaline cap at the front of an advancing pscudopodium.

Types of cells that exhibit amoeboid movements

Types of cells that exhibit amoeboid movement are white blood cells, fibroblast, germinal cells in skin, and embryonic cells. Embryonic cells often must migrate long distances from their site of origin to new areas during development of special structures.

Process of amoeboid movement

The amoeboid movement involves cytoplasmic streaming, changes in cell shape and extension of pseudopodia. When an amoeba moves, its cytoplasm flows into

newly formed arm like extensions of the cell (pseudopodia). The pseudopodia gradually extend and enlarge so that the entire cell occupies the space where previously only a small psendopodium began to form. As the cell moves, new pseudopodia are formed in the direction of the movement, while the posterior parts are withdrawn.

The outer layer of the Amoeba is the stiff gel-like layer called ectoplasm. As the pseudopodium is formed, the more liquid endoplasm streams into it, and a new ectoplasm is formed on the surface. In the rear part of the advancing cell, the ectoplasmic gel is converted to a more liquid endoplasmic sol by a sol-gel transformation.

'Rolling and walking' are the two types of movements that have been described. The rate of movement also show much variation. It may be as slow as 1350 μm/sec. (e.g. *Plasmodium* of acellular slime mold) 5-6 cm/hr in migrating Plasmodium. Freely crawling amoebae move at the rate of 0.5 to 4.5 μm/sec.

In feeding, those amoeba, which travel by small pseudopodia, form food cups. These cups are also motile, encircle the food particle, and their distal ends join and the food is taken into a vacuole.

Mechanism of amoeboid movement

The exact mechanism is not fully understand, but involvement of certain processes have been known from different experiments. Involvement of membrane, nucleus, microtubules of microfilament system, and contractile proteins have been studied.

Total surface area of the membrane is greatly increased when an amoeboid cell changes from a nearly spherical shape to a multipodal form, it then decreases as pseudopodia are retracted. Three theories were proposed for this, but experimental evidence suggest that it is because of the membrane is a fluid or plastic surface that slides freely over the ectoplasm. Essentiality of the nucleaus has been established. Enucleated amoeba soon lose the organised progressive cytoplasmic flow that results in movement.

Involvement of microtubules are microfilaments in amoeboid movement has been well documented. They play a role in movement either by processes like sliding mechanism in mascular contraction or development of shear forces.

The birefringece seen in amoeba suggest an orderly assay of macromolecular structure. The presence of structures like actin and myosin have been well documented in amoeba.

Amoeboid cells have been found to respond to an electrical field and show biopotentials. In an electrical field, *Amoeba proteins* shows solation on the cathodal side and pseudopodia advance in that direction.

Amoeboid movement is caused by contraction of cytoplasm. It is held that at the point of attachment there is less space between the plasmalemma and the ectoplasm than elsewhere. It has been suggested that contact between plasmalemma and ectoplasm

in the *Amoeba proteus* initiates contraction. According to Allen such contraction takes place in the 'fountain zone' and the Amoeba is 'pulled along'. However, Jaha and Bovae proposed that the site of contraction is 'tail process', and the Amoeba is pushed forward.

Actin is largely conserved. It is of similar structure and occur as the filaments in the cell. Myosin from various sources are more diverse, but all bind to actin and causes ATP—hydrolysis with liberation of energy. On this basis, it has been suggested that both cytoplasmic steaming and the formation of pseudopodia may depend on the interaction between actin and myosin.

Control of amoeboid movement

Chemotaxis is the most important initiator of amoeboid movement. Movement occurs either towards the source of chemotactic agent (positive chemotaxis) or away from it (negative chemotaxis). How chemotaxis control the movement is not clear. But it has been observed that the cell surface exposed to the chemotactic agent, develop membrane changes that cause pseudopodial protrusion.

Cilia and flagella

Each cilium has the appearance of a sharp pointed straight or curved hair that projects from the surface of a cell. It is an outgrowth of a structure that lies immediately beneath the cell membrane called the basal body of the cilium. It is of shorter length.

Flagella on the other hand are membrane bound extensions of the cell. A basal body called kinetosome anchor the flagellum with the cytoplasm.

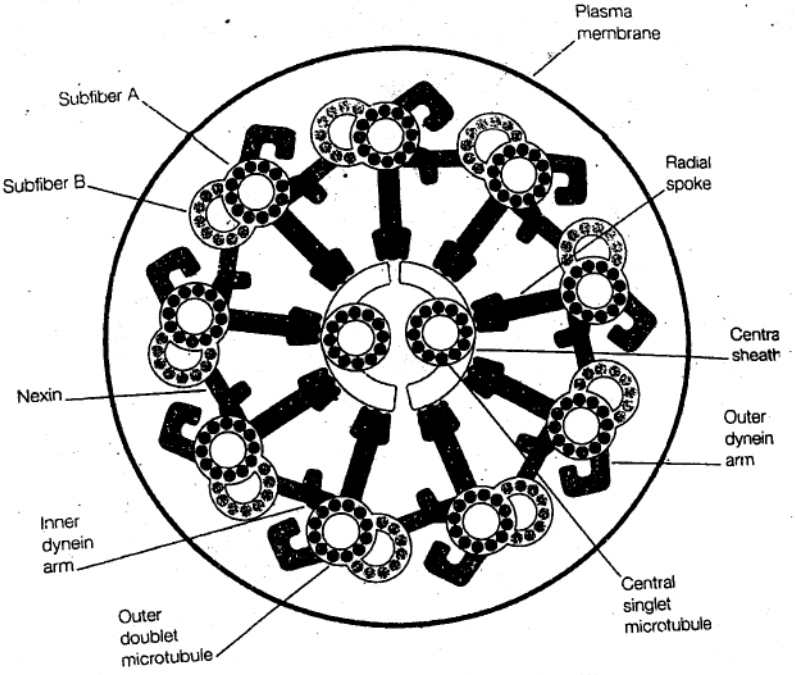


Fig 4.10 : Cross-sectional diagram of a cilium.

Basic structure of cilia and flagella

Flagella are of longer length compared to cilia. In eukaryotic cells, the cilia and flagella are of similar internal structure. Cilia and flagella contain microtubules. The microtubules are arranged as a bundle of nine doublets around the periphery with a pair of single microtubules running within them. This structure is called *axoneme.* In an axoneme, the 9+12 array has specific arrangement (Fig 4.10).

Each of the nine outer doubtlets appears like the figure eight. The smallar circle of the figure is termed subfiber A. The larger circle, subfiber B.

Subfiber A is joined to a central sheath by radial spokes. The neighbouring doubtlets are held together by nexin links. Two dynein arms emerge from each subfiber A with all the arms in a molecular cilium pointing in the same direction.

Mechanism of movement

The dynein in a large protein (MW 1000-2000 kDa). It consists of one, two or three heads depending on the source. The heads of dynein form cross bridges with the subfibres B and its has ATP binding site as well as ATPase activity (Fig. 4.11)

Binding of ATP to dynein causes it to dissociate from B subfiber. The ATPase activity of dynein splits ATP to ADP and Pi. On hydrolysis of ATP, dynein again binds with subfiber B with subsequeat release of ADP and Pi. This ATPase cycle leads to the movement of the cilium as the outer doublets of the axoneme slide past each other. Dynein is involved in converting the energy released from ATP hydrolysis into mechanical energy for movement. Movement is produced by the interaction of

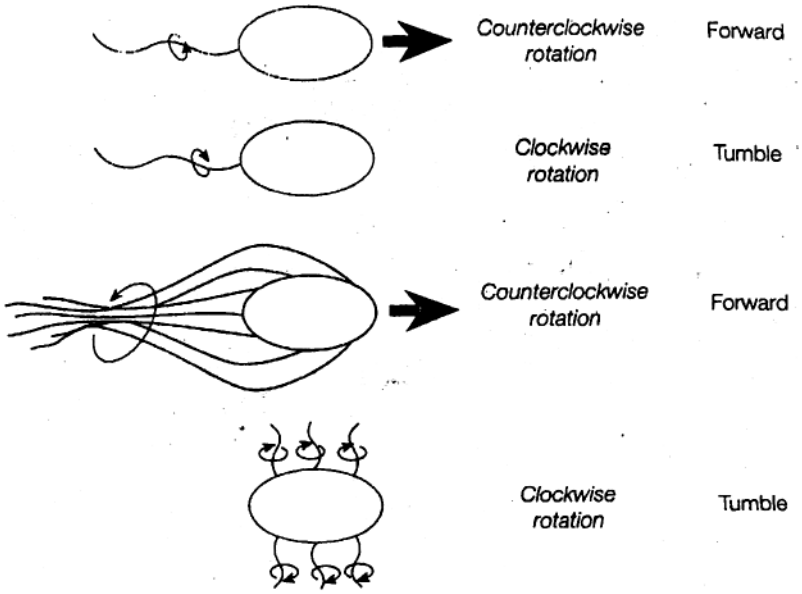


Fig 4.11 : Effects of flagellar rotation on prokaryotc movement.

the dynein arms with one of the microtubules of adjacent doublets. The force between adjacent doublets is generated by the dynein cross bridges. Thus the dynein arms on subfibre A of one doublet walk along subfiber B of the adjacent doublet. The radial spokes prevent sliding motion as happens in the muscle, and so the motion is converted to a local bending. The highly flexible protein, nexin, keeps adjacent doublets together during this process. Bending of the flagellum occurs when the extending dynein arms attached to the neighbouring B tubule, inducing active sliding movements out the expense of ATP.

Movements and its importance

1. A flagellum, like the tail of a sperm, beats with a symmetrical undulation that is propagaled as a wave along the flagellum. A cilium beats asymmetrically with a fast or dash like stroke in all direction followed by a slower recovery motion in which a bending cilium returns to its original position.
2. The number of cilia or flagella present in cell show much variation. A paramoecium may have several thousand cilia, cliliated cell in the respiratory passage may have 250 cilia. Few cilia are present in the epithelial lining of fallopian lube. Cilia in the respiratory passage help in the removal of mucous by mucous cell of the respiratory lining, along with trapped particles toward the nasal opening. Ciliary movement in the reproductive tract help in propalsion of ovum from the ovary to the uterine tube as this is necessary for fertilization and implantation of fertilized ovum.

   A flagellalled cell carries usually one or a few flagella. The sperm of a vast number of animals swim by means of flagella. Flagella exhibit typically sinusoidal motion in propelling fluid/water parrallel to their axis. The undulating action of the flagellum either propels water away from the surface of the cell body or draws water towards the cell body. The cilia exhibit an oar like motion, propelling water parallel to the cell surface.
3. Cilia and flagella are found in many protozoans and mainly related to locomotion, cilia are found in all animal phyla. Modified ciliary structure are present in insect eyes as well as in the majority of other sense organs.
4. Small animals use cilia and flagella for locomotion and muscles are used for the purpose in large animals.

Bacterial flagella : (Chemotaxis) by rotating their flagella. The flagella of bacteria are quite different. They are thinner (about 0.2 μm in diameter, against 0.25 μm for true flagella and cilia), short and relatively rigid. They are related by forces at the base where the are attached to the cell. They differ for eukaryotic cilia and flagella in two ways— (i) each bacterial flagellum is made up of if flagellin (53 kDa sub unit) as opposed to tubulin as (ii) it rotates rather than kinds. The rotary motion of the flagellum is driven by the basal body which acts like a motor. The direction of

flagellar rotation determines the type of movement. Prokaryotes with a single flagellum move forward during counter clockwise rotation and tumble where the flagellum rotates clockwise. Where there are more than one flagellum they behave as a single bundle during counter clockwise rotation and thus move forward, however, during clockwise rotation the flagella act independently and the organisms tumbles. Bacteria can move through the extra cellular medium towards attractants and away from repellents.

4.3 Specialized organs (eg : electric organs and tissues)

Fishes are unique in the animal kingdom in being capable of producing electric current. These are produced from an organ in the tail region and called electric organ. About 250 species of fishes, both chondrichthyes and osteichthyes are reported to possess electric organs. These species have evolved the electric organs independent of each other.

The following are the most important fishes that are known to possess the electric organ.

Elasmobranchs : (i) Electric Rays; (ii) Skates.

Teleosts : (i) Mormyridae; (ii) Gymnotidae, (iii) Siluridae.

It has been found that some fishes produce strong current, some produce weak current and there are some which can only sense electricity but cannot produce electricity. On this basis, electric fishes have been divided into three main categories :

A. Strongly electric fish :
   i) electric eel.
   ii) electric catfish.
   iii) electric rays.

B. Weakly electric fish :
   i) Knife fishes.
   ii) elephant nose.

C. Fishes that can only sense electricity :
   i) Sharks.
   ii) Rays.
   iii) Skate.
   iv) Catfish.
   v) Paddle fish.

Structure of electric organ

The shape and position of the electric organs differ greatly in different species of fishes but all of them have more or less a similar microscopic structure. It is made up of disc like cells called electroplates or electrocytes which are modifed muscle fibers.

These are embedded in a jelly like material and are bound together by connective tissue into an elongated compartment. One face of each electroplate is supplied by nerve fibers and the jelly receive blood capillaries. Each electroplate is a multinucleated cell with nearly transparent cytoplasm. Hence, the elecetric organ looks like a clear gelatinous mass as compared to the muscles.

Stimulation of electric organ

Normally a resting potential exists across both innervated and noninnervated surfaces. In the *Electrophorus* it is about 90 mv. the inside being negative. It is due to difference in the distribution of electrolyte concentration within and outside its cell. At the peak of the discharge, the membrance potential across the nervous face of the electroplate is reversed (60 mv) but the potential across the non nervous face remains unchanged. As a result a potential difference of 150 mv (60 + 90) develops

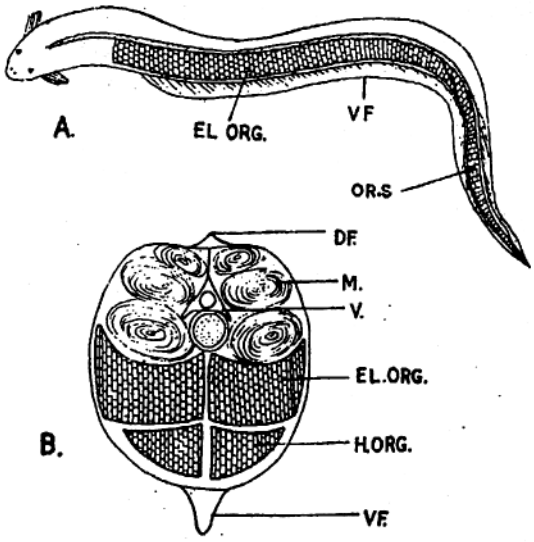


Fig 4.12 : (A) *Electrophorus electricus*, dissected to show electric organs. (B) T.S. *Electrophorus*. DF., Dorsal fin ; EL. ORG., Electric organs ; H. ORG., Hunter's organ ; M., Myotome ; OR. S., Organ of Sach's ; V., Vertebra ; VF., Ventral fin.

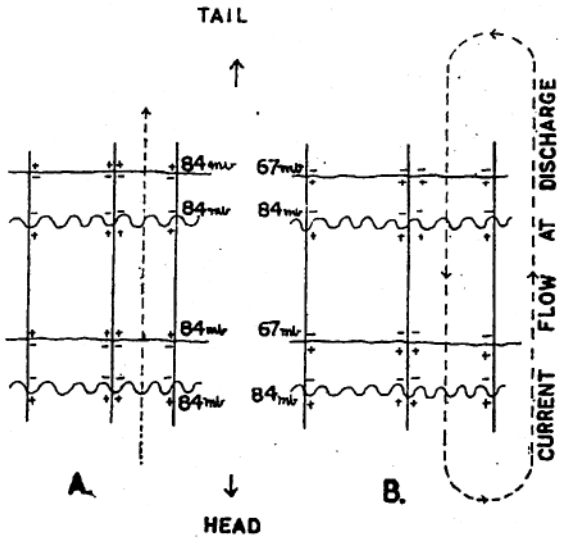


Fig 4.13 : Diagramatic representation of current flow in an electric fish. (A) at rest ; (B) at discharge.

between the two surface of a cell. The voltages of successive electroplates are added up in series and a strong current flows. According to ionic hypothesis, at first sodium ion and them potassium ion enter into the cell due to alteration of membrane permeability and thus results in potential difference. (inside positive and outside negative) (Fig 4.12, 4.13).

These electrocytes or electroplates receive simultaneous command signals from the brain to 'fire'. At the moment of 'firing' the electrocytes are asymmetrically polarized acting as serially connected batteries. The simultaneous firing of electrocytes results in the electric organ discharges (EODs) which are emitted in the surrounding water. In strongly electric fishes, such as electric eel, electric catfish and electric rays, the electric organ is huge containing numerous electrocytes and so their discharge voltage can reach as high as 600 volts. However, in weekly electric fishes, it is small and often less than a volt.

The electric discharge of the electric organs are of two types. In all strongly electric fishes and some weakly electric fishes the discharge is of pulse type. They discharge short electrical pulses intermittently. However, some weakly electric fishes produce wave-type discharge, they produce wave like continuous AC electricity. Some fishes like *Electrophorus* produce both strong and weak current according to need. Most of the electric fishes can produce electricity as well as can sense electricity. However, there are fishes which can only sense electricity but cannot produce it.

Strength of current generated :

The values of electric current produced by some electric fishes are given below :

Skate :		4 volts.
Electric rays :	*Torpedo* :	40 Volts.
	Narcine :	37 Volts.
Electric eel :	*Electrophorus* :	370 – 550 Volts
Electric Cat fish :	*Malapterurus* :	350 – 450 Volts.
Star-gazar :	*Astroscopus* :	50 Volts.

Functions of electric organ

Electrolocation :

The ability to locate an object with the help of electricity is called electrolocation. Fishes have a very sensitive sensory organ which can receive electricity. These are called electroreceptors which remain embedded in the skin.

There are two types of electroreceptors. An **ampullary receptor** contains

supportive cells that lie at the bottom of a narrow channel filled with gelatinous mucopolysaccharide. Afferent nerves embrace the receptor cells.

The tuberous receptor lies buried under the skin in an invagination beneath a loose layer of epithelial cells. This loose layer may differentiate into covering cells that cover the sensory cells are a superficial set of plug cells. These receptor respond to higher frequency than the ampullary receptors. It also can sense the electrical discharge of the fishes own electric organ.

Electroreceptors are used to detect a slight change of electric field caused by nearby objects. Hence, they can see objects electrically. The electrolocation may be active or passive. When the source of electricity used for electrolocation is their own electric organ it is called active electrolocation. The fishes, which cannot produce electricity but can sense it, can sense very weak electricity produced by prey animals. For example, a shark can find a small fish buried in sand by the weak electricity given off by the prey. This type of electrolocation is passive electrolocation because here the electricity is produced by the prey and not the predator which help in locating the object.

Self protection :

Electric fishes are well protected against their own current discharge and against the discharge of each other. This is achieved possibly by developing high insulation around nervous system. As in *Electrophorus*, the spinal cord and the swimming muscles are embedded in a thick layer of fat. Their nerves and muscles may also develop unusually high threshold of excitation.

Importance of electric organs

The form, position are strength of the electric organs show much variation and this is related to there functions.

Fishes with powerful electric organs use these both offensively in their hunt for food and defensively against their enemy. The primary function of the electric organs in all fishes is possibly defensive. Both *Torpedo* and *Electrophorus* have been observed to paralyse small fishes before eating them. Fishes with weak electric discharge create an electric field around themselves and if any object is close to it the electric field breaks up and thus the fish become aware of the presence of the object. Hence it may act as an warning device and it is of considerable value in muddy water where vision is not possible.

Such weak discharge also helps the fish to find direction in the dark water.

Electric organ discharge also helps in maintaining the territoriality by individual fish. It may also be useful for speices or even sex recognition.

Fish can use electricity to feel the environment and also can communicate with each other using electrical singal.

Jamming avoidance response

The wave type electric fish normally discharge at a fixed frequency and each individual has its own frequency. When two individuals having similar frequency meet, however, their EODs interfere with each other causing problem in electrolocation. This is called jamming. To avoid jamming the two fishes shift their frequencies until there frequencies differ in such a way so that operation of electrolocation becomes possible.

4.4 Suggested questions

1. Describe the basic structure of an electric organ and mention its importance.
2. What is jamming avoidance response.
3. Write briefly on electrolocation.
4. Describe the basic structure of cilia and flagella.
5. Write a note on the mechanism of movement of cilia and flagella.
6. Write briefly on the mechanism of ameoboid movement.
7. What is titin ? How it helps in muscle contraction ?
8. How the muscle shortens during muscular contraction ?
9. Describe the sequence of events in muscular contraction with a diagram.
10. Comment on length tension relationship in muscle contraction.
11. How collaping of hollow viscera is related to contraction of smooth muscle cells.
12. Comment a differential distribution of skeletal muscle fiber ?
13. What is caldesmon ? How it helps in muscle contraction ?

Unit 5 ❑ Adaptation

Structure

5.1 Introduction

There are two basic concepts in natural selection. These are fitness and environment. Environment means the kind of habitat in which the living organism or animal lives. Whereas fitness is strongly linked with adaptation. Environments are largely variable in relation to stress that they impose on their inhabitants. This stress may be both abiotic and biotic.

Whenever there is a change in the environment, compensatory changes occur in the living organisms to cope with the environmental modification. Short term compensatory changes in respense to environmental disturbance is called acclimatization or acclimation whereas long-term compensatory changes in known as adaptation. Adaptation is defined as a process; the means by which natural selection adjusts frequency of genes that codes for traits affecting fitness. For example, increasing haemoglobin concentration ; *in taxa* might be seen as an adaptation to potentially low oxygen environment. Adaptation in this sense, is a process that normally occurs very slowly, over hundred or thousand of generations and is usually not reversible. However, in extreme environment or where selective pressure from human interference are strong it can sometimes occurs very quickly. Adaptation is often used as a term for the characters or traits observed in animals that are the results of selection. For example presence of hemoglobin can be said to be an adaptation to increase the oxygen carriage in the blood. The processes of adaptation is usually a slow one that occurs over generations and is rarely reversible. Acclimatization on the other hand, is more rapid phenomenon whereby a biochemical or physiological change occur within the life of an individual animal resulting from exposure to a new condition in the environment. Thus an ascent to high altitude (mountain) may lead to acclimatization to low oxygen and low pressure; movement from arctic areas to southwards will lead to acclimatization to warmer temperature. When similar processes are allowed to occur in the laboratory it is called acclimation. A polar bear is said to be adapted to

polar temperature. The shape of the chest of sherpas are adapted for respiration at low oxygen pressure at high altitude. The human polar explorer or mountaneer get acclimatized to the environment. But they revert back gradually to normal when the environmental condition is changed. Mouse forcefully subjected to cold exposure in the laboratory, get acclimatized to cold. It reverts back to normal when the laboratory temperature is reverted to normal.

The avoiders have or develop mechanisms for getting away from or to avoid an environmental problem either in spece or in time. Avoidance in space is brought about by behavioural change for a small animal it may involve a search for an appropriate habitat using phototactic or chemotactic responses. They look for less stressful microhabitat in crevices or burrows. Larger migratory species try for large scale migration with the help of physiological adjustments. e.g. accumulation of food reserves. Avoidance in time require more complex responses at all levels. An animal entering into torpor accumulate food, construct or find a refuge; huddle in a ball to reduce its surface area for preventing heat loss in a cold environment; it may reduce its core temperature and metabolic rate, it may acquire a thick insulating layer to cope with temperature extrem; it may mobilize or generate new form of enzymes and new forms of membrane components.

The conformers change their internal states, similar to changes in the external environment. They do not try to maintain a homeostatic condition. This involves biochemical and physiological modifications. If the internal environment varies in terms of osmoticity, salinity or temperature, the cells must have a biochemical-physiological system that can function at diverse conditions especially enzymes and membrane stabilization. Their status is such that the animal is kept functioning at extreme conditions avoiding damaging effects of freezing on hypoxia etc. However in general the physiological and biochemical changes are small and cheap, hence are more economic.

The regulators on the other hand maintain their internal environment almost constant irrespective of changes in the external environment. This involves substantial and expensive biochemical and physiological adjustments. For example, to maintain temperature homeostasis, even in the best endothermic mammals, behavioral changes like basking, burrowing, wallowing, huddling, erecting or concealing appendages etc are adapted as a first line of defence. Along with this substantial physiological and biochemical adjustments, occur to have optimal effect. These include changes of blood flow or respiratory rate or nonshivering modification of thermogenesis and thermolysis or the production of heat-shock proteins or anti-freeze molecules.

5.2 Levels of adaptation

When an animal is confronted with changes in its environment it normally exhibit three types of responses - avoidance, conformity or regulation. These responses are the final outcome of adaptations at four levels. Biochemical, Physiological, Morphological and Behavioral. Adapted responses may appear at different spatial levels. Some responses are essentially subcellular, some affect the morphology or activity of whole cells; others manifest as effects on entire tissue or organs like changes in muscle size heart volume or arrangements of vascularisation.

However, there is a limit of conforming and regulating. For example, the osmoconformers show some regulation in extreme low salinity to avoid cell damage due to excessive swelling. Similarly the osmoregulators become unable to regulate at lower salinities and turn to conformers, homeothermic animal allow their extremities to become poikilothermic at extreme cold.

These strategies adapted to counteract enviromental changes are associated with different costs and benefits thus the 'avoidance' by shutting down is cheap but this causes the animal to remain out of the 'race' for some time without any growth and reproductive output. In case of avoidance by way of migration may be expensive but it allows the animal to continue with growth in another environment. Avoidance in poor physical environment by way of shut down or migration give additional benefits like avoiding predation or competition.

Conformity at the extremes of temperature, salinity or hypoxia may allow a minimal lifestyle but over a broad range it can maintain a reasonably productive life style at a cheaper cost. Regulation is rather definitely expensive; osmoregulation takes about substantial amount of energy but thermal regulation draws up of total energy budget. However, it can gather food all through and the conformers and avoiders become prey. With extra food they can grow and reproduce faster. Despite high cost they become dominant in many ecosystems.

Adaptation at different spatial levels

Adaptive responses may occur fandamentally at the molecular level, but they are manifested at different spatial levels in the whole animal. Animals are made up of several distinct compartments. Each of them may show different adaptive responses that leads to a change in the animal as a whole. The individual cells contain intracellular fluid (ICF) which is its own fluid environment. These cells are directly bathed in tissue fluid or extracellular fluid (ECF). In may species ECF is distinct from blood (ECF enclosed in specialized channel known as blood vessels) or lymph (ECF present in lymph vessels) or hemolymph. These fluids may be different in composition from the classical ECF. The relation between fluids of these three compartments and their

homeostasis is highly complex. Adaptation to environment may require modification at all these levels. For many invertebrate marine animals, the ECF and blood are identical with sea water and so adaptation occurs at cell–ECF level and within cells i.e. across intracellutar compartments. But in case of nonmarine animals there occur extensive regulation at the boundary between external environment and blood i.e. skin, as well as at the cellular lavel.

In terms of the whole animal, adaptations occur at different spatial levels or site.

1. To maintain difference between the outside world and the circulating blood, adaptation occurs at the outside surface e.g. skin. This surface (skin) may be made up of relatively unspecialized epidermal cells or a complete multilayered structure with chitinous, keratinous or lipid containing elements. Sometimes, variability in structure is confined to a particular area of the skin such as gill surface and other areas of skin are relatively inert or impermeable.

2. Adaptations can occur at the boundary between ECF and the circulating fluid. This mainly occurs in vertebrates when some constituents of blood pass out from blood capillaries into ECF and vice versa; some constituents of ECF goes back to blood via lymphatic system.

3. Adaptation at the boundary between ECF & ICF. It involves the cell membrane itself that control exchanges between ECF and ICF. The exact make up of ICF is very different from that of ECF. Total concentration of the fluid of the two compartments, however, should be similar to prevent osmotic swelling or shrinkage of the cells.

4. Adaptation within the cell : Cells themselves are strongly compartmentalized and the organelle menbranes are involved in regulating the exchanges between the cytoplasm and the nucleus, mitochondria or endoplasmic reticulum.

As the environmental condition is changed the animals show adaptation at different functional and spatial levels to cope with the environment which is essential for its survival.

5.3 Machanism of adaptation

Structural and functional modification of living organism occur as the environment undergoes changes. This is essential for survival of the living objects. Temporary changes occur to counteract the environmental changes and it reverts back to the previous form when the environmental changes are over. This type of change is called acclimatizion e.g. cardiospiratory changes in a mountaineer during ascent to high altitude.

But sometimes permanent changes occur to cope with the environmental changes involving genes (actually in DNA). This is called adaptation. The specific shape of the chest of sherpas living at high altitude is an adaptive respiratory change.

Proteins are basically involved in cellular processes. Cell division is the result of DNA duplicatian. Protiens are the products of genetic code of DNA. Alteration of DNA will alter protein structure, hence will alter cellular functions and thus will lead to phenotypic changes at the organism level. Adaptation occurs when any change in the DNA level becomes expressed via protein changes, as a trait that is beneficial in a particular environment and so persists due to selective advantage, eventually spreading through population. Hence anything that controls proteins is at the core of adaptation.

A change in the shape of the protein is associated with many cell-functions like switching effects via kinase-phosphetase system (e.g. phosphorylation of glycogen synthase makes it inactive — glycogen synthesis is inhibited, dephosphorylation of the enzyme makes it active – glycogen synthesis is stimulated. Former is caused by glucagon and the latter by insulin ; motor effects like ATP driven shape change in myosin causing its movement along actin during muscle contraction; change in shape of protein pump causing chanelling of ions (Na^+–K^+ pump) and other molecules across membranes and between cell compertments. These functions are again subjected to extremely complex regulation and amplification to produce controlled effects on whole organism.

DNA, RNA and protein synthesis

Most of the chromosomal DNA in any animal cell does not code for RNA or for protein. Only 1% of DNA sequence is transcribed into functional RNA sequence. DNA is duplicated during cell division. The transposons (Short pieces of DNA consisting of a few hundred to 10,000 nucleotide pairs) present in DNA organisation are involved in mutation. There is some evidence that they undergo long periods of quescence in the genome and then exhibit sudden bursts of activity (transposition bursts) being triggered by environmental change. These bursts form a link between environment and adaptation. They cause increased biological diversity because transpositions will potentially bring together two or more new traits which (of little value alone) become very useful by working together. Thus these transposition bursts can produce randomely modified progeny. This happens at times of environmental stress. Hence transposons act as useful symbiotic factors, generating diversity just when it is most needed.

A change in DNA is the fundamental mechanism of evolution. DNA is *insulted* in a normal course by various factors like thermal degradation. UV radiation metabolite action as oxygen radical. DNA repair enzymes like DNA repair nudclease. DNA polymerase & DNA ligase are induced by DNA - damage. The DNA is repaired and normal function is left undisturbed. This is a safety mechanism of the cell. Such repair is crucial in genetic recombination in meiosis where chromosomes undergo crossing-over. It has been found that minor mis-matches are corrected by such repair

but where there is a particularly poor match between the recombining strands, the pairing is usually aborted and no recombination occurs; excessive scrambling of the genome which might be detrimental to physiological functions is thus averted.

DNA forms RNA by transcription. The RNAs undergoes posttranscriptional modification and active RNAs are formed. There are three types of RNA – mRNA, rRNA and tRNA. mRNA with the help of rRNA, tRNA join aminoacids through peptide bond and form polypeptides by the process called translation.

The polypeptides formed by translation process undergo posttranslational modification and are converted to active polypetides. These undergo proper folding with the help of chaperans and functional proteins are produced.

The proteins that can be formed theoretically are many. But the proteins that are selected for in evolution are those that can reliably and repeatedly folded up into stable forms which can be subjected to conformational changes by controlling mechanisms.

Protein effect is a balance between protein synthesis and protein breakdown. Degradation of abnormal proteins is mediated by intracellular proteosome-ubiquitin mechanism. However, in case of degradation of normal protein with specific halflife is dependent on aminoacid sequence on the N-terminal side of the peptide. Certain amino acids offer protection while other such as Arg, Asp and Glu promote proteolytic attack. Thus the aminoacid sequence of this region can be selected over evolutionary time to give protein with appropriate half life.

The protein synthesis at genomic level is regulated by regulating proteins at different levels—

(i) Transcrptional control; (ii) RNA processing control; (iii) RNA transport control; (iv) translational control, mRNA degradation control; Protein activity control.

The control at the transcriptional level involves the actions of repressor protein (inducible enzyme); corepresser - aporepressor complex (repressible enzyme); transcription activators and transcription repressor; transcription factors operating at the enhancer site and repressor site causing enhancement and suppresion of transcription respectively.

Adaptation is achieved by way of formation new structure and development of new functions for better survival in the existing environment. Environmental factors act as stimulus. They in turn modulate intracellular mechanisms by way of genetic alteration and gene expression that ultimately leads to phenotypic changes.

Permanent changes in protein structure can be brought about by mutations and recombinational changes in the DNA sequence of a cell. This can produce subtle changes in enzyme activity, signalling activity and subcellular morpthology, as well as expression of other proteins. All these in turn can lead to permanent heritable

change in the development of organisms. These are seen at a series of levels. At biochemical levels, there may be increased thermal tolerance of an enzyme or expression of a more pH stable allozyme; morphologically there may be developmental changes of gene expression leading to altered positioning of muscle cells, nerves, blood vessels, even completely modified appendages; Physiologically there may be increase in heart rate under the influence of Ca^{++} and cAMP signalling regime on Na^{+} pumping; changes in the rate of Na^{+} pumping as more or different channel proteins are synthesized; at the behavioral level there may be an increase in the speed of attack, greater sensitivity to a sexhormone or a concious change in response. All of these proteins induced changes the raw material for adaptation. Environmental stimuli modulate intracellular system through neurotransmitter and hormones. In relation to adaptive changes the hormones are the major controllers of genetic expression and thus of protein synthesis.

5.4 Significance of body size

An unconditional generalisation cannot be made about adaptive responses like avoidance, conforming and regulating across animal kingdom. It is dependent on body design and habitat.

1. Smaller and soft bodied animals are more likely to be avoider and conformers. They can use microhabitat more effectively. They enter into protected crevices, burrows or rest on and in other organisms. Since these animals have a high surface area to volume ratio (small animal) rapid fluxes will occur (water, ion, air, thermal energy) and so restoration to normalcy is expensive. They also have little inbuilt protection from swelling and shrinkage and lack complex outer layer which can provide same insulation or impermeability. In these animals in estuarine habitats, a switch over to cyclic avoidance is common an conforming is also seen. Those terrestrial habitats where there is both continuous high environmental stress and high fluctuation, conforming may not be an option, rather exceptional strategies for avoidance are common (torpor, estivation, encystment etc.).

2. Animals with hard outer layers (exosceleton) may have better options for some regulations are a greater independance of their environments Some animals like arthropods likely develop partial regulation of osmotic concentration. Because of exoskeleton, the outer surface can have reduced permeability and may be partly thermally insulated by cuticular hairs. So fluxes are slowed. But in these animals, behavioral avoidance, aided by efficient limbs (and sometimes wings) that can be built from an exoskeleton becomes a major part to cope with environmental change especially in the more rapidly changing terrestrial habitats.

3. Large animals are much more likely to be regulating in all environments with

important exception of the relatively equable and unchanging open ocean. Larger animals operate in a larger scale environment where rapid changes are relatively unimportant. Because of larger surface (lower surface area–volume ratio) more time is available for regulatory mechanism to operate. They may have better opportunities for energy storage. They may have room for complex neurohormonal control mechanism. In terrestrial habitats, where environmental changes are faster, all these factors work together and the only option left to large animals is regulation.

Thus the smaller animals adapt to avoidence and conforming specially if they have soft body . In case of small and medium sized animals with hard exoskeleton, some regulation and behavioral avoidance means are adapted to cope with the environment. The large animals cope with the environment by adapting regulation.

Anatomical and functional changes occur in a predictable way with increasing body size. The study of the size-related effect is known as scaling.

4. Heart rate, O_2 consumption, metabolic rate, all are higher in small animals compared to large animals. This is partly related to the disproportional increase of tissue of low metabolic rate like skeleton, fat and connective tissues in large animals. The activity of oxidizing enzymes is higher in corresponding tissues of small animals than larger ones. Cytochrome oxidase and malic dehydrogenase have been shown to be more in small than large mammals. The muscles of small animals consume much more energy during steady running than do the muscles of large ones which is due to higher metabolic rate in small animals. (It is known that large divers remain submerged for longer period of time than small divers. This is because the large animals have less O_2 consumption and less metabolic rate.) Small mammals have relatively larger surface area and so heat dissipates more readily from them. In case of fish, the energy cost for swimming per unit mass of body unit declines with increased size.

For an evolutionary and ecological reason, there is an optimum body size for a taxon. Scaling and size dependency are crucial factors in all comparative physiology and patterns in animal adaptation. The size of a particular species of animal is determined by several factors like —

(i) **Phylogenetic inheritance.** e.g. insects are small and vertebrates are large.

(ii) **Basic physiological design.** e.g. animals with open circulation are larger than those with no circulatory system, and those with close circulation may be larger again.

(iii) **Basic mechanical design.** e.g. animals with hydrostatic skeletons are usually relatively small as are those with exoskleton; while those with tubular endoskeletons are relatively large.

(iv) **Habitat.** e.g. any given design may be larger in aquatic habitat giving internal support than on land where self-weight is a problem.

Scaling effect have been observed in migration. In case of avoidence as a measure to cope with stress birds usually migrate; while terrestrial animals, specially smaller ones, are inclined to escape by hibernation and torpor. The smallest migrating mammals have a body mass of about 20 Kg (some African antilope) ; where as very small birds and even monarch butterfly having one gram body mass are regular migrators. Thus different sizes and locomotory modes affect the ability to migrate effectively. For any given migration time, flying animals will be able to complete a much longer distance than walkers or swimmes of similar size or alternatively, to achieve the same distance in the same time, walkers or swimmers would have to be much larger. Global, terestrial or marine migrations are therefore slow and prolonged often taking many months (as in salmonid fish). Whereas aerial migration by birds can be completed in a matter of days and weeks. Adaptation of the capability to store food is an important factor in migration. Few animals undertake nonstop migration relying on food stores; many animals accumulate fat stores of 25 – 50% of body weight before migrations. Since fishes are generally ectothermic with 10 fold lower metabolic rate than similar sized birds and mammals, they can migrate much further without feeding and have the greatest migratory capability.

Several experimental and other observations indicate that adaptation is related to body size in respect of cellular adaptation as well as in organism level.

Because of small size, the desert rodents evade the heat by retreating to their underground burrows during the day. However ground squirrel, although of small size is similarly handicapped in hot weather but they adapt in a different way. When they are outside their burrows on a hot day they get heated very rapidly. They cannot tolerate very high temperature and die if heated to 43^0 C. However a temperature of $42{\cdot}4^0$ C is tolerable by them without apparent ill effects. When it is heated it goes to a relatively cold burrow as get cooled rapidly due to large surface area and the exposure to hot climate is tolerated.

5.5 Suggested questions

1. Explain acclimatization, acclimatian and adaptation. Write briefly on the levels of adaptation.
2. Give brief accounts of mechanism of adaptation.
3. Name the adaptive responses seen during environmental change. How are they related to body size ?
4. Write briefly on scaling with examples. Comment on or name the different factors that determines the size of an animal.

Unit 6 ❑ Physiological adaptation to different environments

Structure

6.1 Marine adaptations

There are thousands of species of marine life, from tiny zooplankton to enormous whales. Each is adapted to the specific habitat it occupies. Throughout the oceans, marine organisms must deal with several things that terrestrial life do not

- regulating salt intake
- obtaining oxygen
- adapting to water pressure
- dealing with wind, waves and changing temperatures
- getting enough light

Salt regulation : Fish can drink salt water, and eliminate the salt through their gills. Seabirds also drink salt water, and the excess salt is eliminated via the nasal, or "salt glands" into the nasal cavity, and then is shaken, or sneezed out by the bird. Whales don't drink salt water, instead get the water they need from the organisms they eat.

Oxygen : Kish and other organisms that live underwater can take their oxygen from the water, either through their gills or their skin.

Marine mammals need to come to the water surface to breathe, which is why the deep-diving whales have blowholes on top of their heads, so they can surface to breathe while keeping most of their body underwater.

Whales can stay underwater without breathing for an hour or more because they make very efficient use of their lungs, exchanging up to 90% of their lung volume with each breath, and also store unusually high amounts of oxygen in their blood and muscles when diving.

Temperatures : Many ocean animals arc cold-blooded (ectothermic) and their internal body temperature is the same as their surrounding environment.

Marine mammals, however, have special considerations because they arc warm-

blooded (endothermic), meaning they need to keep their internal body temperature constant no matter the water temperature. Marine mammals have an insulating layer of blubber (made up of fat and connective tissue) under their skin. This blubber layer allows them to keep their internal body temperature about the same as ours, even in the cold ocean. The bowhead whale, an arctic species, has a blubber layer that is 2 feet thick.

Water pressure : In the oceans, water pressure increases 15 pounds per square inch for every 33 feet of water. While some ocean animals do not change water depths very often, far-ranging animals such as whales sometimes travel from shallow waters to great depths several times in a single day.

Whales can dive deeply. The sperm whale is thought to be able to dive over $1^{1}/_{2}$ miles below the ocean surface, and they can do that successfully because their lungs and rib cages collapse when diving to deep depths.

Wind and waves : Animals in the intertidal zone do not have to deal with high water pressure, but need to withstand the high pressure of wind and waves. Many marine invertebrates and plants in this habitat have the ability to cling on to rocks or other substrates so they are not washed away, and have hard shells for protection.

Light : Organisms that need light, such as tropical coral reefs and their associated algae, are found in shallow, clear waters that can be easily penetrated by sunlight. Since underwater visibility and light levels can change, whales do not rely on sight to find their food. Instead, they locate prey using echolocation and their hearing.

In the depths of the ocean abyss, some fish have lost their eyes or pigmentation because they are just not necessary. Other organisms are bioluminescent, using light-giving bacteria or their own light-producing organs to attract prey or mates.

6.2 Adaptations for life in an estuary

The term estuary comes from Latin *aestus,* meaning tide. The adjective *aestuarium* means tidal. Thus estuary is defined as 'the tidal mouth of a great river, where the tide meets the current'.

There are a number of vertical and horizontal attributes to estuarine ecosystems. The intertidal zone is alternatively flooded and exposed. There may be salt marsh or mangrove wetlands, algal beds, sand or mud flats, reefs of oysters, mussels or calms in this region. Organisms that live in this region have developed special adaptations.

Like other aquatic ecosystems the vertical gradient of light is a limiting condition for photosynthetic activity. At euphotic zone, where light reaches the bottom plants can live attached to the bottom. Estuarine water clarity tends to be much greater near the ocean, so both rooted and planktonic plants generally photosynthesize in greater

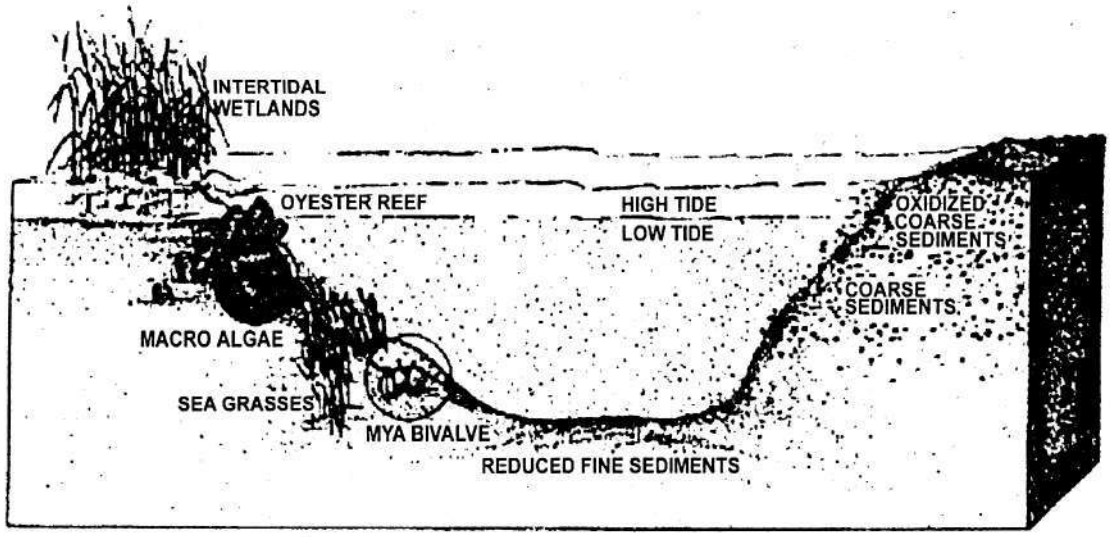


Fig. 5.1 : Vertical zonation of estuarine habitat

depths than in low salinity regions. Animals of the aphotic zone are dependent on transported food from somewhere else. Estuaries are productive environments for aquatic life. They are rich in nutrients compared with rivers and oceans, and they are also good animal refuges, offering protection from storms, competitors and parasites. However, estuaries do experience sudden and often widespread changes in salinity, temperature and dissolved oxygen levels, so aquatic organisms must find ways to cope with these changes.

Dissolved oxygen content is another important parameter for estuarine life. There exists a gradient from oxidizing (aerobic) to reducing conditions (anaerobic or anoxic) in estuaries for biological and chemical processes.

Salinity

Salinity is perhaps the most important factor affecting aquatic species. Most aquatic animals are adapted to life in sea water. These animals vary in terms of the degree to which they are able to tolerate the lower salinities of the estuary. A smaller number of animals are adapted to life in fresh water, and few of these species tolerate salinities in excess of 2 ppt (parts per thousand).

Ocean salinity is 35 ppt. The salinity level of the estuary varies along its length, with depth and with the seasons. There are extreme changes in salinity from almost freshwater conditions in winter to saline (or almost hypersaline) conditions in summer. This change in salinity causes a change in the concentration of dissolved gases (fresh water contains more oxygen than sea water at the same temperature) and in the density and viscosity of the water.

Adaptations to changing salinities

Estuarine organisms possess adaptations or behaviours that enable them to cope with changing salinities. Few organisms remain in an estuary for the whole of their life cycle. Some fish are truly estuarine, spending their whole lives in the estuary. Some use it as a nursery habitat only and others are marine visitors, coming in when the salinity suits them. Mobile animals like fish and crabs can swim away from unfavourable conditions. Prawns and crabs move out of the estuary in winter when waters are less saline. However, less mobile (sedentary) animals such as barnacles and worms have to either seal themselves inside their shells or adapt to the conditions. Many sedentary animals die when conditions are unfavourable and must recolonise when conditions change. Many algae and seagrasses die off during winter periods when salinity levels become too low. Some organisms are able to tolerate extreme conditions for a short time only. A sudden change, such as an unusual heavy summer fall of rain which produces a freshwater flow into the estuary, or extreme conditions of salinity, will produce a variety of responses. For example, worms, molluscs and fish produce slime or mucus to cover and protect their sensitive body surfaces. Some polychaete worms and crabs retreat into holes or burrows, plugging them. Other animals withdraw their sensitive body parts, or close their shells. If an organism cannot escape or reduce contact with the water during times of abnormal salinity it must use a physiological response. The marine species generally reside near the mouth of the estuary, the freshwater species in the low salinity areas and the estuarine species somewhere in the middle. Many estuarine species are osmoregulators, meaning they can maintain a constant salt balance, no matter what the salinity of the water is. Alternatively, an animal may modify its metabolic rate or change its patterns of activity. The physiological response of an animal to salinity changes takes time to complete, so it is often supplemented by a behavioral response that enables it to either delay or moderate exposure to unfavourable conditions, completely avoid them, or slowly adapt their body to the new saline environment. For example, some bivalves close their shell valves when sea water suddenly becomes diluted. After a while they become used to these conditions. Organisms that are capable of dealing with varying salinities are euryhaline, and organisms that can only deal with small changes in salinity are stenohaline. Stenohaline animals rely on coping methods such as moving out of the area, burrows in the sand, excreting excess salts or closing their shells, and worms, molluscs and fish can produce mucus or slime to cover sensitive body parts. The mussel, *Xenostrobus securis,* deals with salinity of 2g/L by closing up their valves for up to many months. When an organism is unable to move or reduce contact with the unfavorable salinity, it then relies on physiological responses such as osmosis (become iso-osmotic with estuarine water) or changes its activities or its metabolic rate. Generally mature organisms are better able to handle the stress of

salinity changes than reproducing adults or newly hatched eggs. Many sessile organisms die if conditions become too harsh and have to re-colonize when conditions are once again suitable.

Euryhaline crustaceans respond to changes in salinity by molecular processes. Vertebrates respond by regulated blood osmotic concentrations and controlling ion fluxes and organic osmolytes. Invertebrates can regulate blood osmotic concentration, cell volume or both.

Sodium chloride is very important in regulating blood osmolarity in euryhaline crabs, and regulating fluxes and permeability's of these ions is how the crabs deal with the salinity. Green shore crabs respond quickly to salinity variations and within six hours, their blood Na^+ level reaches a steady state. It is thought that the crabs possess a sensory organ on their legs known as "hair peg" which is what responds to the salinity variations and sets the wheels in motion for response. Gill structure is also helpful in dealing with changes in salinity as the gills contain two different sets of epithelial cells, which are believed to be crucial to ion and gas exchange.

Adaptations to changing temperatures

Temperatures are more variable in the estuary than they are in the ocean. In winter the estuary is colder than the ocean, and in summer it is warmer. Daily temperature fluctuations can also be extreme, especially in the shallows. Also, the solubility of oxygen depends on temperature. (More oxygen dissolves in cold water than in warm water.) Adverse temperatures may cause responses of avoidance and escape. One of the most common mechanisms that certain organisms use to cope with conditions in the cold winter months is to transform into a resting stage. Another habit is to burrow into the mud or sand on the bottom of the estuary. Fortunately in the Swan-Canning system temperatures are rarely extreme, never reaching freezing in the winter or going above 40°C in the summer. Organisms that are able to withstand varying temperatures are eurythermal. Those that can't must use other responses such as moving, burying themselves, or to transform to their resting stage. Temperature can fluctuate over seasonal cycles in temperate estuaries. Bacteria have adapted to this by reorganizing their biochemical pathways and adjusting protein and DNA synthesis rates.

Adaptations to changing oxygen levels

In water containing low amounts of oxygen, organisms have physiological and behavioral mechanisms to survive. In sessile animals that cannot move to a more oxygen rich environment, they must lower their energy demand, maintain their metabolism or use a method of creating energy without oxygen. Many animals do this by increasing their heart rate and the flow of water past respiratory surfaces.

Some organisms have respiratory pigments, which allow them to maximize the oxygen they do get. Others are very good regulators of oxygen uptake and maintain constant rates of uptake until the oxygen falls to a 'critical pressure'. They then switch to anaerobic metabolism. This can only be maintained for a short time. Some organisms, such as oyster larvae, have been shown to lower there aerobic metabolism in response to hypoxia.

Nektonic adaptations to estuarine environment

General adaptation for true nektons include the existence and position of swimming organs (e.g. fins), a smooth streamlined body, a specific gravity close to water, the ability to extract oxygen at a relatively rapid rate from the surrounding water in order to support the large energy requirements of swimming. Additionally, well developed sensory organs are necessary for most organism with active nektonic existence. Fishes mostly achieved these requirements. They have streamline body and are able to control their depth through varying their specific gravity by changing the amount of gas in the swim bladder. Fishes have gills well adapted for rapid oxygen exchange, thus meet an active and sustained movement. Their specialized sensory organ, the lateral line, is very sensitive to sound waves and to changes in water density. The chin barbells in some fishes help to locate food.

There are specific adaptations of nekton that allow them to flourish in estuaries. This is interesting since the estuarine environment is apparently hostile or stressful. The rapidly changing physical and chemical environment imposes great energy demand of fish so that most species cannot survive there. Estuarine species may be divided according to tolerance depending on temperature and salinity. Some species require high salinities and are not found in the riverine reaches of estuaries. Some other species possess a wide range of salinity tolerance. There are only a few species which are adapted to withstand extreme variation in both temperature and salinity.

6.3 Fresh water adaptations

Aquatic animals develop various types of adaptations according to changes in aquatic environment. This is essential for the maintenance of life in aquatic medium.

Majority of herbivores are algal grazers. Animals show variation as they live in ripples and send bottom areas. But most animals are seen in these two areas.

First fautra of the hill stream of India are very rich and varied. *Botia dayi, Botia dario, Nemacheilus aureus, Nemacheilus beavani, Nemacheilus montanus, N. rupicola* are bottom living fishes usually found under stones and boulders. Because living at the bottom, they have broad head, dorso-ventral flattening of the anterior part of the body, presence of thorasiç adhesive apparatus, horizontal alignment of paired fins. Many fishes at the bottom stay under stones and boulders and thus avoid water

currents. The paired fins are fanlike and horizontally placed. The ventral surface is smooth, without scales, and flattened.

Those, which are inhibitant of fast flowing water, need to move without the favour of water current, by means of mascular effect. *M. armatus* has an elongated cylindrical and cel like body and the posterior part is highly compressed laterally. The snout is produced into a long beak-like process to reduce the pressure exerted by the water current. Some fishes like *N. botia* are inhibitants of slow running water and usually abundant in ports. These fishes assemble just below the surface of the water and move through the weed-beds for feeding and returns to the pools. In fishes which feed on insects, mainly larva and nymphs, the mouth is ventral and suited to the bottom feeding habit. They make adjustments to withstand the current. Many nymphs have efficient adaptations, which enable themto tolerate current strength upto 300 cm/sec. There animals cling to the surface of the rock keeping their heads towards the current and when pressed firmly against the substratum. The water current exerts a downward pressure which helps to held the animal in position. Body flattering is thus one of the adaptations to enable organisms to take shelter beneath stones and avoid the force of the current. The inches, which usually live under the stones have flattened streamlined bodies.

Streamlining of the body helps to offer the least resistance to water and is also found among many stream-invertebrates. The tapering bodies of the nymphs of the mayfly illustrate this very well.

Development of suckers helps the animal to attach if itself to the substrate. The larva of the blackfly, *Simulum* are found in large numbers, attached to stones in stream-bed. They are one of the mostly adapted animals to the life in fast-flowing water. Simulum possess large salivary glands which secrete a pad of sticky silk on the substratum to which it attaches itself by means of a pair of modified prolegs at the rear end of the abdomen. At the posterior end of the semicrect body is a circle of rows of outwardly directed hooks, which when the muscles of the disc are relaxed, move outward and cling on to a silk web previously placed there by the larvae. The anterior end of the body then swings freely in the current. There is a fan like food gathering organ on each side of the mouth which traps food particles from the water current.

The development of hooks is also an adaptation to current. Caddies fly larvae occur most abundantly in streams with moderate to swift currents. They usually construct cases made of leaves, twigs, sand grains and stones cemented together with silk secreted by the animal itself. In swift water, the cases are stout, cylindrical, tappered posteriourly and are usually swollen and more solidly so constructed of sand, pebbles or rock fragments.

Hill stream fishes have also become highly adpated to this dynamic environment.

The chief factors affecting the life cycle of these organisms are : (A) Strength of the current, (B) Hill streams are shallow and clear, so they have to bear intense light during day-time, (C) Sufficient food is available but in the form of algae covering stones and rocks.

Fishes have to develop special adaptations to live in the environment. The head and the body of most hill stream fishes are greatly flattened and in the highly specialized form the body is leaf like. The ventral profile is straight while the dorsal profile is only slightly arched. The head is usually small and semicircular. The size is generally small so that they can easily take shelter any rock and beneath them, and can conveniently live in shallow water protected from direct sunlight. The scales of these fishes undergo reduction or are very minute or embedded in the skin as in *Nemacheilus.* This is of special significance to adhesions to rocks and stones so that although scales are present in the dorsal in the dorsal and lateral aspects, they are absent in the region of the chest. The paired fins specially the pectorals are modified for adhesion and the number of inner rays is increased. The fins are shifted that it can act as hydroplanes and adhesion to rocks and stones is facilitated. In most species, the lower lobe of caudal fin is longer than the upper one. Some species possess (eg. *Nemacheilus,* and *Glyptosternum*) develop a band shaped caudal peduncle which appears to be an adaptationto life in fast flowing waters.

The position of the mouth is shifted from the anterior and of the snout to the ventral side towards the tip of the snout. Instead of a transverse the mouth is generally horny covering which help to scrape the algal material from the stone-surface, for feeding.

The mouth is surrounded by sensitive barbules beneath the head. These help in testing the substratum. In some species (eg. *Nemacheilus*) the lips are divided in the middle and are swollen so that they form a saucer, when pulled outwards. In *Glyptosternum,* the lips are reflected and spread round the mouth to form a bread sucker for attachment. The eyes are generally small in size and pushed towards the upper surface where they come to lie close together.

Besides formation of additive sucker for the lips, the skin is thrown into grooves and ridges on the ventral side of the body, specially between the pectroal fins of the body. Such straited structures serve as friction plates for attachment to stones.

The gill openings are restricted to sides only and do not extend beyond the pectoral fin. Any effect on respiration is copresneted by the well acrated water of the stream. The bottom lining forms required negative buoyancy ; the air bladder is considerably reduced or degenerate and becomes completely a useless hydrostatic organ. This helps the bottom living organisms to aquire negative buoyaney. Most fishes expend much energy in fighting the current, hence they spend long periods in resting.

Amplibians found in torrent waters generally develops small lungs, presumbly to reduce buoyamoy. Size of the appendages and body are reduced or a streamline shape is achieved to reduce the surface areas exposed to the full impact of current.

Most animal species, though have developed devices for clinging to the bottom, they are more abundant at the underside of the rocks in riffles, than they are on the upperside. Some species, however, such as rotifers, water mites, protozoans find shelter within the mass of algae that may cover the top of the rock. Even swift water fishes take all the advantages of whatever protection available.

Inspite of various mechanisms for maintaining position against the current, a continuous drift animals occurs downstream. The driff is more prominent at night than daytime. A continuous drift of downstream is compensated by the adult insects to lay their eggs in shallow water. The adult insects emerging downstream commonly fly upstream for reproduction. Some fishes move upstream for spawning.

Lake is a stretch of water surronded on all sides by land. Different types of lakes are there like glacial lakes of high altitude, tectonic lakes formed due to movement of deeper parts of earth's crust, lakes formed due to volcanic activity, lakes formed by wind action or land slide.

In the lakes, there is a gradual rise of temperature as one moves from the depth towards the surface.

Amount of O_2 present in a lake or pond depends on extent of contact between water and air, oi. the circulation of water and on amount produced and consumed by the lake community. Animals and microgamisms present in water, use O_2 and produce CO_2.

According to habitat characteristics, the pond and lakes exhibit 3 zones— littoral, limnotic and profundal zones. The shallow water which has light for the surface to the bottom is called litteral zone. The bottom part that does not receive day light is called profundal zone.

Light serves as the initiator or driving force of the ecosystem. It controls the development of pigmentation in animals, and stimulates the development of various adaptations. The development of vision depends on availability of light in the media. Fossorial forms have generally reduced vision and so do deep sea forms. In some organisms locomotory activity depends on light intensties. This is called photokenesis.

In cave dwelling and fossorial or borrowing animals, the size of the eyes are generally reduced due to absence of light or is found embedded in the integument. Eyes are absent in cave dwelling fishes and cave amphibians. In burrowing amphibious such as the Apoda, the eyes are covered by a fold of the integument. In abyssal depth, where there is absence of light, the organisms either have reduced vision or the eyes are highly developed to preceive low light, which spread the occurrance of biofuminiscente at those depths.

Light controls locomotor activity of many animals by a direct action on locomotion. This is called photokinesis. The rate of swimming in mussel crab has been shown to be upregulated where the intensity of light is increased in experimental condition.

Temperature

Heat is an ecological factor which has an wide ranging influence directly on the biota as well as indirectly in combination with other factors. It is a direct regulator of the climatic conditions of a place as well as modulator of other factors such as density and salinity. Organisms have adjusted themselves, and have developed various adaptations to meet the temperature ranges. Solar radiation falling on earth supplies energy to the living organisms in the form of heat, the activity of the organisms depend on transfer of heat from the media in which they live. In their adaptations, organisms have specialized themselves to live whether narrow limits in small scale environment (microhabitat) that are significantly different from the larger environments (habitats) of which they form a part.

In case of poikilothermic animals, the physiological activities are adapted to function in spite of the temperature strees. Poikilotherms are ectotherm and the body temperature is dependent on the amount of heat they take up from water (environment).

In case of homeothermic animals, the heat generated in the body is retained. Hence, it has high body temperature. They have a high metabolic rate compared to poikilothermic animals. They are endothermic animals.

Homeothermy is necessary for organisms to inhibit the terrestrial environment, because air and land experience violent fluctuations of temperature and climatic conditions.

1. Animals produce eggs that can survive thermal extremes.

2. In arctic regions animals have their freezing point of plasma lowered for –80°C to –1.47°C, in response to seasonal changes in temperature. They produce anti-freeze, which reduces their & risk from freezing.

3. Amphibians in land lose water very fast by evaporation through skin. Their risk of dessication is partially compensated by a reuction in urine flow, update of water for urinary bladder and absorption through skin, all promoted by vasotocin. Their survival is further aided by a retread to underground burrows, where the temperature is much lower than on the surface.

Reptiles have impermeable skin, but they also lose temperature byinceased evaporation, although the rate is lower than amphibia. However, some reptiles can control the body temperature to some extent. In *Varanus*, as the temperature of environment rises, powerful pumping movement of the mouth and neck occur. This is the functional equivalent of ponting in mammals. It can also increase its body temperature by increasing its metabolic rate.

In general most of the poikilotherms undergo hibernation and estivation to meet temperature extremes, while homeotherms are able to regulate their body temperature through special adjustments. Hibernation occurs to tide over extremes of cold. The term is used to describe conditions or instances in which metabolism is reduced, animal puts itself in a state of rest with reduced metabolic rate. There occur a marked drop in body temperature. It is of wide occurrance in poikilotherms specially at higher latitudes.

Terrestrial reptiles avoid overwintering by retreating into rocks, crevices, or borrows as do the amphibians. Many aquatic turtles spend the winter buried in mud beneath ponds and stream borders. Among mammals hibernation is reported for the order Monotremata, Marsupialia, Insectivora, Chiroptera, and Reptilia. At least one bird is known to hibernate. During awakening from hibernation a sequence of events occur— (i) rapid rise of body temperature, (ii) warming of heart promotes circulation of blood which carries heat in a restricted way such that the heat passes largely to the respiratory system and the brain during the early stages of awakening.

Aestivation

Dormancy during summer months, when temperature is high, or excessive dryness and shortage of food products is called onset of dry season, burry in the mud of a swamp or river bed and forms mucous lined cell in which it aestivates. This may last for as long as seven months, with the return of rains and subsequent rise of water level, the fish energes from its aestivation chamber.

In many insects, dormancy take the form of diapause (during this time growth and development are suspended or retarded). In homeothermic animals, body temperature is kept constant in spite of changes of environmental temperature. These animals are curythermal and are capable of regulating their body temperature by thermostatic regulation through physiological adjustments.

Amphibia move from direct sun to sheddy place. Insects move from the sun to the stones and beneath leaves. Some insects freeze at night and thaw out in the day in cold climate. O_2 content of water is less compared to air. Thus, to ensure O_2 supply, respiratory adaptations occur. These include (i) increase in respiratory surface, (ii) use of higher affinity oxygen storing pigments, (iii) modulation of ventillatory or circulatory rates.

In fresh water vertebrates, gills are the most common adaptations to ensure O_2 supply. Diferent forms of gills develop eg. serial by repeated filamentous gills, tentucular crown gills, enclosed lamellate gills are seen in molluses and annelids, tracheal gills, spiracle gills, rectal gills in insects. Freshwater vertebrates use cutaneous or gill based respiration. Skin breathing is quite common in fresh water fish larvae and many eels. Cat fish rely on it most extensively to sustain their metabolic rate. Many species depend largely on skin respiration. eg. Salamander, frogs. Although

most fishes, and juvenile or neotenic amphibions use gills and nearly all adult amphibians, reptiles, birds and mammals (being secondarily aquatic) rely on air breathing through lungs.

The haemoglobin saturation curve shows a 'left shift', when the O_2 demand is increased than occur increased rate of ventillation in fresh water animals which may arise from hypoxia of the environment, or increased metabolic activity. In small animals like sponges and rotifiers, ventillation is mainly done by cilia and flagella, their activity increasesas needed.

In cray fish, hyperventillation is seen initially, but as hypoxia sets in, the animal shows brady cardia and the circulatory flow pattern alter to give increased flow anteriorly over the brain.

Fresh-water vertebrates show sophisticated ventilatory response to hypoxia. In fish, hypoxia is sensed by oxygen receptors located in the brain and aorta. Both rate and stroke volume are increased in hypoxia, by changing the buccal and opercular pumping pattern.

In freshwater vertebrates, when O_2 demand is increased due to hypoxia in environment or increased metabolic activity, ventillation is increased, by changing the buccal and opercular pumping pattern. In snakes and turtles, ventilation is intermilttent at rest, but becomes continuous during steady swimming. Green turtle shows a seven fold increase in mean ventillation frequency and increase in both pulmonary and aortic blood flow.

In bimodal breathers, besides aquatic breathing, air breathing devices are present. For example, in arthropods, air bubbles and plastrons are used as air supply systems. Special diving adpatations are found in fresh water reptiles, birds and mammals.

Many essentially freshwater animals are able to use both aerial and aquatic oxygen and may switch over to air breathing when their aquatic habitat begins to dry up or overheat and becomes hypoxic, or it becomes too rich in H_2S, due to decomposition.

Custaccans, fishes and amphibians are most notable as bimodal breathers. Many bimodal crabs have reduced gill-area, typical of air breathers.

Fresh water habitat give rise to problems related to reproductive strategies. There is a general tendency for the fresh water invertebrates to have very short life cycles. Commonly there is a very high reproductive output in freshwater invertebrates. In *Drussena* (Zebra mussel) one female may release 1 million oocytes per year. Reproduction in fresh water vertebrates also show a similar pattern. However, secondary freshwater vertebrates that are essentially land animals, usually resort to a terrestrial site for reproduction. Turtles lay their eggs in the upper shore.

Fresh water is rarely very deep, the largest lakes being only 1000-1500 meter, so that freshwater animals do not normally have to cope with great pressure. However,

buoyancy is more difficult in fresh water due to reduced specific gravity of the medium giving very little lift. Swim-bladder is present in fish and it is 6-9% of body volume. Swim bladder gives the buoyant effet.

Suckers and hooks for adhesion develop specially in case of bottom dwellers. Certain fish have ventral sucker and blackfly have hooks to hold on to rocks.

The four eyes fish Anableps can focus simultaneously on terrestrial and aquatic image having practically two eyes with a pear shaped lens and two sets of pupils above and below the water miniscus. Fresh water beatles, bugs, and fishs such as pike have eyes with large pupil. This is also seen in surface dwelling betles and bugs. Chemoreception is also an important device in many freshwater animals for prey location, predator avoidance and the location of hosts and mates in river. Many aquatic insects have water pressure receptors ; flow receptors are also found. In crustaceans and fishes they are located in cephatic appendoges like antennae or vibrissae. Electro reception occurs in a range of fresh water fishes with a very weak and strong field producers. Echolocation is seen in freshwater dolphins, but it is largely lost in freshwater animals.

Most of the freshwater zooplankton are filter feeders. Hydra takes small prey that brush against and trigger the shinging cells in its tentacles. Aecolerated cutrophication is an well recognised process in both rivers and lakes. The reduced O_2 level leads to submerged plants disappear as light is cut off and animal casualty follow. The submerged animals change eg. *Tubifex* worms replace crustaccans and then surface dwellers decline.

Fresh water can be polluted in different ways. Contaminated natural freshwater is the single biggest source of human disease of the world today. Some of the diseases is due to parasites and pathogens naturally present in freshwater. But a huge additional hazard is added by pathogens derived from sewage and animal wastes.

Global warming leading to acid rain makes the water acidic. This causes— (i) increased mortality of fresh water animals, (ii) It solubilizes minerals. Alluminium sloubility causes aluminium poisoning of many fishes, (iii) Cuticle formation in insects and crustaceams is affected, thus leads to defective osmoregulation. (iv) Hatching of some salmonid fishes is decreased.

Thus, abudance of many species decline, particularly snails, amphibians, small crustaccans, leading to decrease in zooplankton. Other fishes such as eel, insects and rotifers are also destroyed it the pH stays below 5 for long time.

6.4 Extreme aquatic environments

The deepsea physical environment has special characteristics. To cope with this environment, the animals living there undergo suitable adaptation.

Four remarkable characteristics prevail in this area. These are :

(1) ***Absence of sunlight.*** Beyond 200 fathoms their is no sunlight.

(2) ***Quiescence :*** Because of depth the movement of water is almost absent.

(3) ***Cold environment :*** In the deepsea the temperature is nearing the freezing point and the temperature remains constant.

(4) ***Pressure :*** The pressure is high as it increases with depth.

(5) ***Lack of green vegetation :*** Total absence of light is the reason for total absence of green vegetation in the deep sea environment.

For survival in the deepsea environment, the animals have to develop adaptive changes to cope with adverse physical condition present there. The general characterstics of deepsea animals are as follows :

(i) The deepsea animals are weak and delicate.

(ii) The body is generally simplified.

(iii) They are either totally blind or they possess powerful telescopic eyes to catch maximum amount of light.

(iv) They develop long feelers to act as tactile organs.

(v) Almost all the deepsea animals are luminescent.

(vi) Most of the deepsea fishes live on the exudes of decaying matters and so the animals lose the masticatory power. There are other animals which possess powerful jaws.

(vii) Most of these develop wonderful devices for caring the youngs. other produces large number of youngs to overcome the hostile environment.

(viii) Small size is an important charasteristic of deep sea living.

Structural modifications in deep sea animals.

Almost all the phyla have representatives who lead deep sea life. The modification of the invertebrates are diverse compared to vertebrates.

Modification of the vertebrates of the deep sea have been found to be as follows. Amongst the elasmobranchs the true sharks do not exhibit deep-sea charasteristics excepting the luminous sharks. The silver shark, however, show deep-sea characteristics in having huge eyes and long alternated body and tail.

Amongst the teleosts, the typical deep sea form is *Cetomimus*. It has a long mouth, small teeth, very small eyes and scaleless body. However, in *Ipnops* there are no eyes and only two large luminous organs are found on the head. Scaleless body and well developed luminescent organs are the characteristic features in Stomiatidae. In *Gastrostomus*, the body is long, slender with rows of luminous organs on the lateral sides of the body and the mouth is bounded by very large jaws. The Gadiformes

(cod like forms) have reduced mouth and dentition; the eyes are extremely large, the trunk is reduced and has a filamentous tapering tail.

The anglers show typical deep sea characteristics. The paired fins are adapted for crawling on the bottom of the sea and the anterior finrays of the dorsal fin function as a lure. In *Linophryne*, the finrays are provided with luminous organs to attract the prey. The *Oneirodes*, however, is blind but has luminous organs to compensate for the loss of eyes. In another deep sea fish, *Protostomias*, specialised light producing organs are present in rows on the lateral side of this body.

Another important feature of deep sea fishes is flatness of the body to adjust with high pressure. The body in these fishes becomes flattened and the mouth is shifted to the lateral side of the body.

These structural modifications in deep sea forms are due to peculiar physical condition of the deep sea environment. The deep sea forms are geologically very recent in origin. These forms were originally the inhabitants of the pelagic or littoral regions which migrated to the deep sea and become adapted.

6.5 Parasitic habitats

Parasites depend fully on their host for their living. Hence, they adapt themselves accordingly. Adaptation to specific environments is a dynamic feature of all living organisms. Pararitism starts with an accidental meeting of few animals. Gradually one, the guest, starts to lead a more dependent life upon its host for food and shelter. This change from a free living life to a life in which food and shelter become available without any effort bring about profound modifications in the make up of parasites.

The helminths are modified morphologically and physiologically to live in their particular environment. The modification depend on the degree of parsitism.

Morphological adaptations

Every part of the body of a helminth parasite exhibits twist due to parasitic mode of life. The strutural modifications involves two aspects — degeneration and attainment of new organs.

Degeneration : There occur loss or simplification of unused organs or parts. In helminthes the loss or degeneration involves particularly the digestive and locomotory organs.

(a) ***Digestive / Alimentation.*** As the parasites begin to live on digested or semidigested food of the host, there occur reduction in their alimentation and digestive glands.

In the adult trematodes, the digestive tube is a blind gut. In the larvae of

trematodes, the gut is either very simple or completely eliminated. In adult tape worm the alimentation is completely absent. As the animal lives amidst digested and semidigested food, the food is absorbed directly through the general surface of the body. *Trichinella* and cystacercus larvae in vertebrate muscles occur in such location that they are constantly surrounded by rich nutritious food. The food is absorbed through the outer layer of the body.

(b) ***Locomotor organs :*** Thes locomotory organs are not necessary for the parasites because they live their entire life in the body of the host. Hence, locomotor organs are totally reduced. However, in the free living larval stage of the parasites, such as miracidian and hexacanth the ectoderm is ciliated.

(c) ***Sense organs :*** These sense organs in helminthes are also simple structures. This happens because they lead a sedentary life in a sheltered habitat. This correlation is especially seen in the endoparasites.

2. Attainment of new structures/organs

Parasitic existance is made possible due to modification of old structures and formation of new structures. These are necessary and helpful in food absorption, protection, attachment and vast reproduction.

(a) ***Integumet :*** The outer integument or cuticle of helminth parasite becomes highly modified and is so adapted as to resist against the digestive juices, passage of food and for adhesion.

Cuticle become thin to absorb food in those parsites which live in rich nutritious environments e.g. adult liver fluke in bile, blood flukes in blood, tapeworms in intestine. *Trichinella* and cysticereus in vertebrate muscle, several larval forms developing in lymph spaces and in blood stream.

In case of some gut parasites e.g. tapeworms, nemotodes etc. which remain attached to the wall of the gut the cuticle is suitably modified— it becomes thick impregnated with chitin like substance and enzyme resistant, so that it is not digested by digestive enzymes but remain permeable to water.

In most trematodes the outer integument contains spines, spinnules or scales of various kinds. These cuticular modifications protect the outer surface of the worms against the abrasive action of food and roughage flowing around them. In the chinese liver fluke *Cloriorchis sinesis*, the larval stage has a spinous cuticle, this suggests that possibly it was a gut parasite and subsequently converted to a parasite of the bile passage.

(b) ***Musculature :*** The muscles are well developed in tapeworm (e.g. *Taenia*). This enables *Taenia* to spread and elongate their bodies along the length of the intestine of their host. Similarly, power of locomotion enables the round worm (e.g. *Ascaris*) to counteract gut peristalsis and thus maintain their position in the intestine. In this way the parasites becomes capable to obtain predigested nutrients of the host.

(c) ***Organs for attachment :*** All parasites develop suitable devices for attachent to their hosts, either to the exterior or to the interior of host cavities. Helminthes are variously modified for adhesion to the body of their hosts. In all adult parasitic flatworm acetabula or suckers develop. The liver fluke (e.g. *Fasciola*) has two suckers on the ventral side of the body one anteriorly and another posteriorly. In the tape worms the scolex bears either four sucking cups (e.g. *Taenia solium*) or accessory suckers (e.g. *Myzoplyllo bothrium*) or lateral sucking grooves or bothria (e.g. Diplyllo bothrium).

Some cestodes and nematodes also develop hook-like structures near the cephalic ends that further help in attachment. In some, a basal circlet of hooks (*T. solium*) or rows of hooks develop (*Dipylidium carninum*). A buccal armature bearing toothlike structure has been found in *Macracanth orhynchus*.

In some helminths e.g. miracidium and cerceria, unicallular secretary glands develop which help during penetration into the host. The secretion contain a lytic enzyme that digests the host's time make passage through which the worms move. In larval trematodes cystogenous glands develop whose secretion help in cyst formation. These glads degenerate after their functions are over. Hook worm contain buccal glands which pour secretion that are anticoagulant and also has histolytic action.

(d) ***Vast reproduction :*** Parasitic adaptation involves a significant development of reproductive organs with much increased capability of reproduction. In both flat worms and round worms, the interior of the body is mostly occupied by the genital organs. The chances for survival is increased by astronomical production of eggs. Self fertilization is more common than cross fertilization. The life history usually includes several larval stages for multiplication and for easy and sure transfer from one host to anohter. The nervous system in all parasitic helminthes and excretory system, paticularly in trematodes show little deviation or adaptation to perticular mode of life.

Physiological adaptation

Apart from structural modification physiological or functional adaptation are also seen in parasitic adaptation.

1. ***Intracellular digestion :*** The parasites develop intracellular digestion because they feed on tissue elements and inflammatory excudates. This has been observed in flukes.

2. ***Osmoregulation :*** The osmotic pressure of the interior of parasitic worms remains less than or same as that of their hosts and thus there occur no difficulty in the exchange of water. It has been found that cestodes have well developed osmoregulatory system and their pH tolerance is also high.

3. ***Anaerobic respiration :*** The parasites adapt to decrease their oxygen demand

because they live in less or no oxygen environment. The intestinal parasites live in an environment completely devoid of free oxygen. They develop a very low metabolic rate which requires a minimum amont of oxygen. In the absence of free oxygen, energy is obtained by the fermentation of glycogen which is broken by glycolysis carbon di-oxide and fatty acid. The glycogen and lipid contents in than body tissues have been found to be high whereas the protein content is less.

4. ***Antienzyme :*** One important parasitic adaptation is to develop antienzymes so that they are not attacked on destroyed by the host enzymes. Most of the helminth parasites, particularly intestinal parasites, secrete antienzymes in order to protect themselves from the gastric juices and digestive enzymes of the host. A dead worm cannot secrete these enzymes and so they are digested by the host enzymes. Medicines are used to destroy this anti-enzyme action and thus the parasites are subsequently destroyed by the host.

Group A(II)
Biochemistry

Unit 1 ❐ Glucose Metabolism *via* EMP and HMP Pathway, TCA Cycle

Structure

1.1 Introduction

The major function of carbohydrate in metabolism is as a fuel to be oxidized and provide energy for other metabolic processes. In this role, carbohydrate is utilized by cells mainly in the form of glucose. The three principle monosaccharides resulting from the digestive process are glucose, fructose and galactose. Fructose may assume considerable quantitative importance if there is a large intake of sucrose. Galactose is of major quantitative significance only when lactose is the principal carbohydrate of the diet. Both fructose and galactose are readily converter to glucose by the liver. Pentose sugars as xylose, arabinose, and ribose may be present in the diet, but their fate after absorption is obscure. D-Ribose is synthesized in the tissues for incorporation into nucleotides.

1.2 Glucose metabolism *via* EMP and HMP pathway, TCA cycle

1.2.1 Glycolysis through EMP pathway

Glucose enters into the glycloytic pathway by phosphorylation to glucose-6-phosphate. It is completed by the enzyme hexokinase (here glucokinase) in liver parenchymal cells. ATP is required as phosphate donor and Mg^{++} as cofactor. Glucose-6-phosphate is the junction of several metabolic pathways (glycolysis, gluconeogenesis, hexose monophosphate shunt pathways, glyconeogenesis, glycogenolisis). Glycogen is hydrolyzed to glucose-6-phosphate *via* glucose-1-phosphate by the enzyme phosphorylase. In glycolysis it is converted to fructose-6-phosphate by phosphohexose isomerase. This fructose-6-phosphate in presence of the enzyme phosphofructokinase coverted to fructose-1, 6-bisphosphate. It requires ATP and Mg^{++}. Fructose-1, 6-

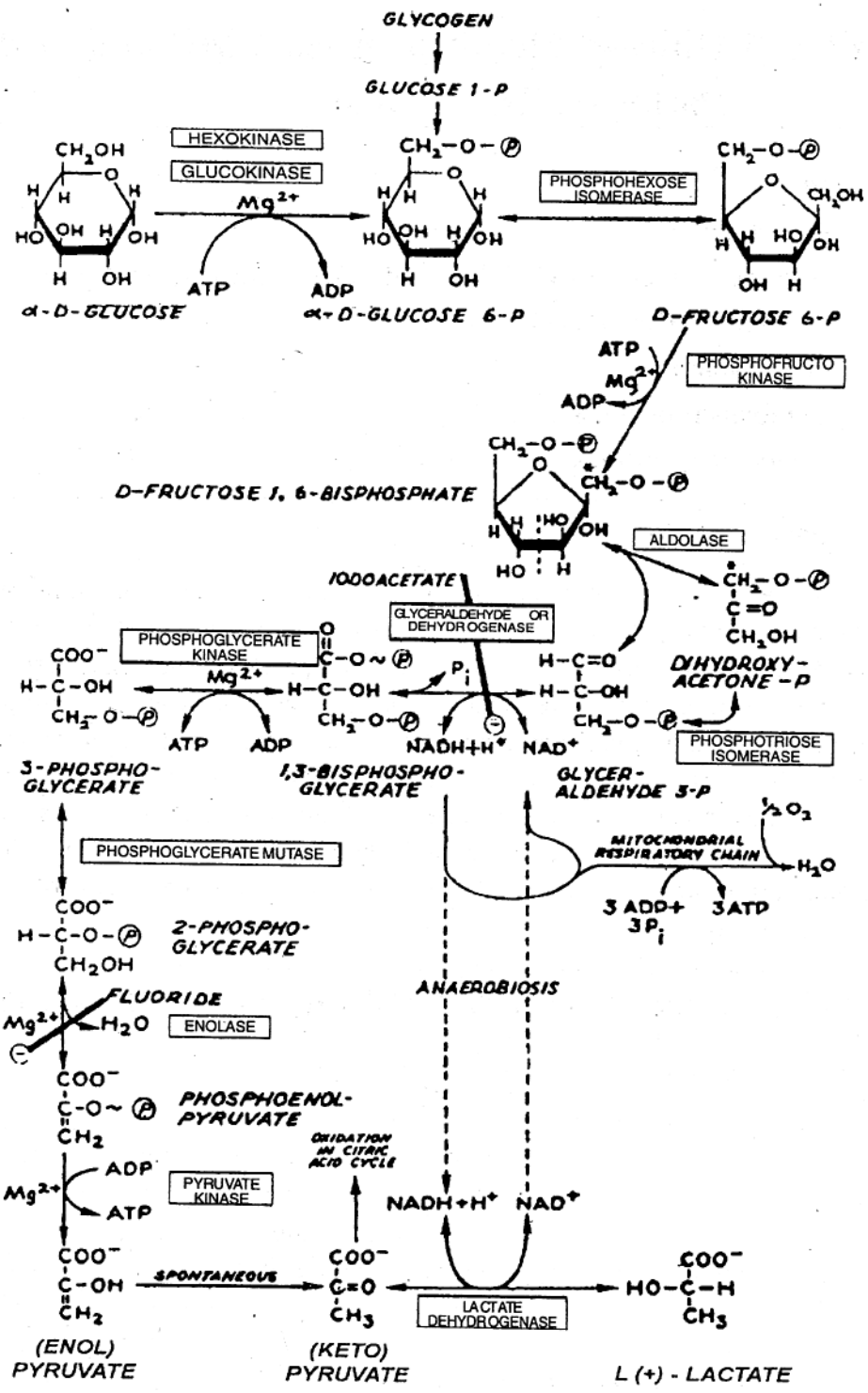


Figure 1.1 : Embden-Meyerhof pathway of lycolysis. (Ⓟ– PO_3^{3-}; Pi, $HOPO_3^{3-}$; ⊖, inhibition)

bisphosphate splits by aldolase into two triosephosphate (glyceraldehyde-3-phosphate and dihydroxyacetone phosphate). Glyceraldehyde-3-phosphate and dihydroxyacetone phosphate are inter converted by the enzyme phosphotriose isomerase. Glyceraldehyde-3-phosphate by oxidation in presence of NAD is converted to 1, 3-bisphosphoglycerate. 1, 3-bisphosphoglycerate is converted to 3-phosphoglycerate in presence of the enzyme phosphoglycerate kinase. It is subsequently converted phosphoenol pyruvate by the enzyme enolase. Phosphoenol pyruvate is converted to pyruvate (Keto) and the spontaneously to Keto pyruvate for the activation of the enzyme ATP and Mg^{++} is very much essentia. Keto form of pyruvate converted to lactate by lactate dehydrogenase. NADH + H^{++} generated for the oxidation of glyceradehyde-3-phosphate is utilized for the activation of the enzyme lactate dehydrogenase.

1.2.2 Hexose monophosphate shunt/pentosephosphate pathway

Major functions of the hexose monophosphate shunt pathway are to provide NADPH for reductive syntheses outside the mitochondria and to provide ribose for nucleotide and nucleic acid, synthesis. This pathway for the oxidation of glucose occurs in certain tissues. It is active notably in liver, lactating mammary gland, adernal cortex, thyroid gland, erythrocytes, testis in addition to Embden Meyerthof. pathway of glycolysis. It is clear that this pathway is markedly different from the EMP of glycolysis. Oxidation occurs in the first reactions and CO_2, which is not produced at all in the EMP. Three molecules of glucose-6-phosphate give rise to three molecules of CO_2 and three 5-carbon residues. The latter are rearranged to generate two molecules of glucose-6-phosphate and one molecule of glyceraldehye-3-phosphate. In shunt pathway NADP and not NAD^+ used as a hydrogen acceptor. The enzymes of this shunt pathwasys are found in the extramitochondrial soluble portion of the cells.

In the first phase glucose-6-phosphate undergoes dehydrogenation and decarboxylation to give pentose, ribose-5-phosphate, *via* ribulose-5-phosphate. In the second phase ribulose-5-phosphate derived from second molecules of glucose-6-phosphate is converted to xylulose-5-phosphate and then by series of reaction forms fructose-6-phosphate and erythrose-4-phosphate. Third molecules of glucose-6-phosphate converted to xylulose-5-phosphate. This xylulose-5-phosphate and erythrose-4-phosphate form fructose-6-phosphate & glyceraldehyde-3-phosphate.

Three molecules of glucose-6-phosphate finally form two molecules of fructose-6-phosphate and one molecule of glyceraldehyde-3-phosphate. So it can be considered that three molecules of glucose-6-phosphate finally forms 2½ molecules of glucose-6-phosphate.

A summary of the reaction is given below :

3-glucose-6-phosphate + $6NADP^+ \rightarrow 3CO_2$ + 2-glucose-6-phosphate+ glyceraldehyde-3-p + 6NADPH + $6H^+$

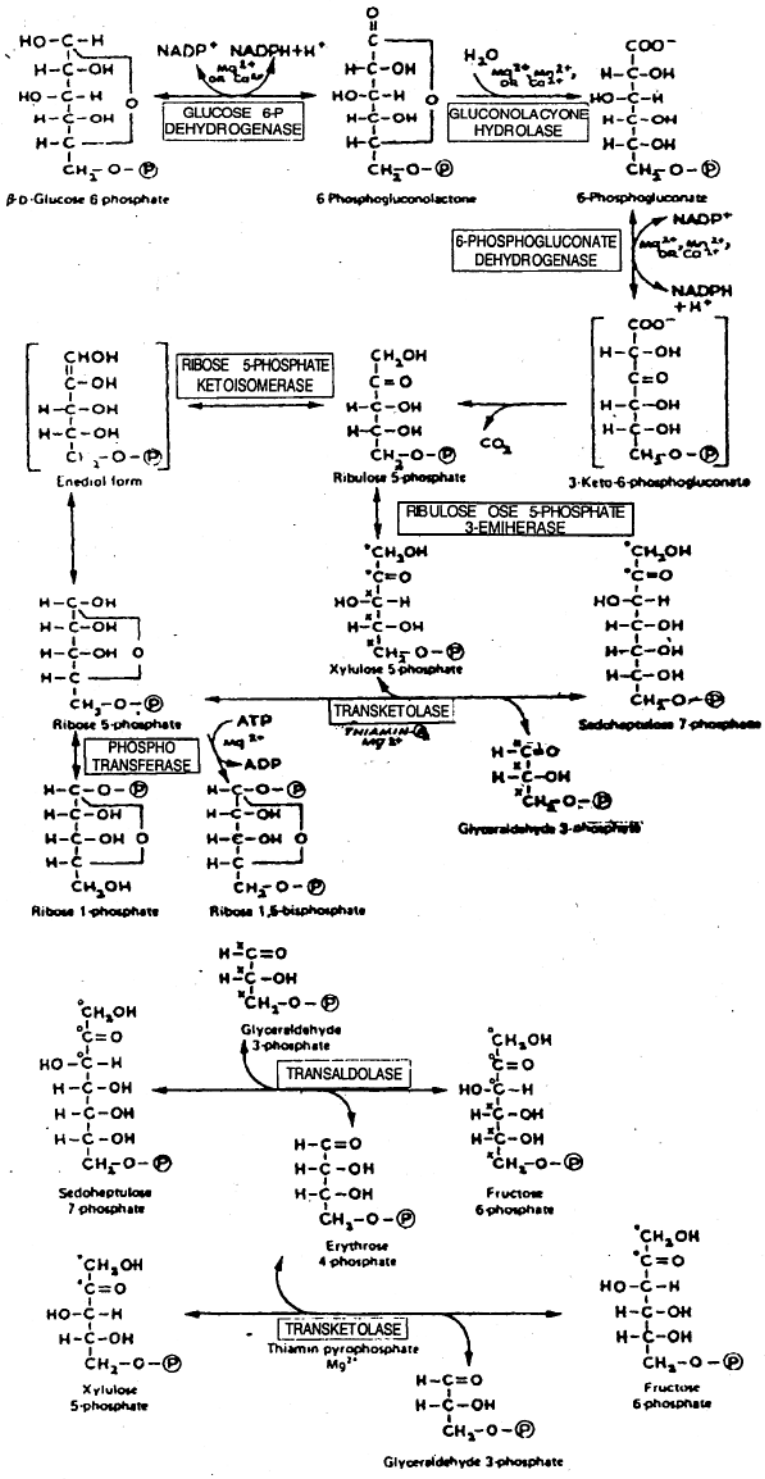


Figure 1.2 : The hexose monophosphate shunt (pentose phosphate pathway) (Ⓟ – PO_3^{3-})

Metabolic significance :

It is active in liver dipose, adrenal cortex, thyroid peritrocytes, testis, lactating mammary and, NADPH of the shunt pathway help in the synthesis of fatty acids and steroid. This NADPH also responsible for the synthesis of the amino acid via glutamet dehydrogenase. In erythrocytes oxidised glthatione is reduced to reduced gluthatione by this NADPH, catalised by gluthatione reductase. Reduced gluthatione remove H_2O_2 from the erythrocyte whih is catalyzed by glutathione peroxidese. This reaction is important. The accumulation of H_2O_2 may decrease the life span of erythoracyte by increasing the oxidation of haemoglobin to methemoglobin. The hexose monophosphate shunt pathway provides pentoses for nucleotide and nucleic acid synthesis.

1.2.3 Oxidation of pyruvate of acetyl-CoA

Before pyruvate can enter the citric acid cycle, it must be transported into the mitochondria via a special pyruvate transporter that aids it passage across the inner mitochondiral membrane. Within the mitochondria, pyruvate is oxidatively decarboxylated to accetyl-CoA. This reaction is catalyzed a mltienzyme complex and designated as pyruvate dehydrogenase complex.

$$\text{Pyruvate} + NAD^+ + \text{CoA} \rightarrow \text{Acetyl-CoA} + NADP + H^+ + CO$$

1.2.4 The Citric Acid cycle (Krebs cycle, tricarboxylic acid cycle)

At the beginning of the cycle Acetyl-CoA combines with 4-carbon carboxylic acid oxaloacetate, resulting in the formation of a 6-carbon tricarboxylic acid citrate. A small quantity of oxalocacetate is needed to facilitate the conversion of a large quantity of acetyl units to CO_2, oxalocetate may be considered to play a catalytic role.

The major fnction of the cycle is to act as the first common pathway for the oxidation of carbohydrae. protein and lipid or glucose, fatty acid & amino acid. Further the citric acid cycle liberated a good amount of energy during the oxidation of carbohydrate, lipid & protein. During the course of oxidation of Acetyl-CoA in the cycle reducing equivalents in the form of hydrogen or of electrons are formed. These reducing equivalent then enter the respiratory chain, where large amont of ATP generated in the process of oxidative phosphorylation. the enzymes of the citric acid cycle are located in the mitochondrial matrix, either free or attached to the inner surface of the mitochondrial membrane. It facilitates the transfer reducing equivalent to the adjacent enzymes of the repiratory chain. It is of further significance that citric acid cycle has a dual or amphibolic role. Through this cycle both synthesis and catabolism of glucose, fatty acid and amino acid take place.

In citric acid cycle at the beginning the oxaloacetate combines with acetyl CoA in presence of citrate synthatase to form citrate. Citrate is then converted to isocitrate

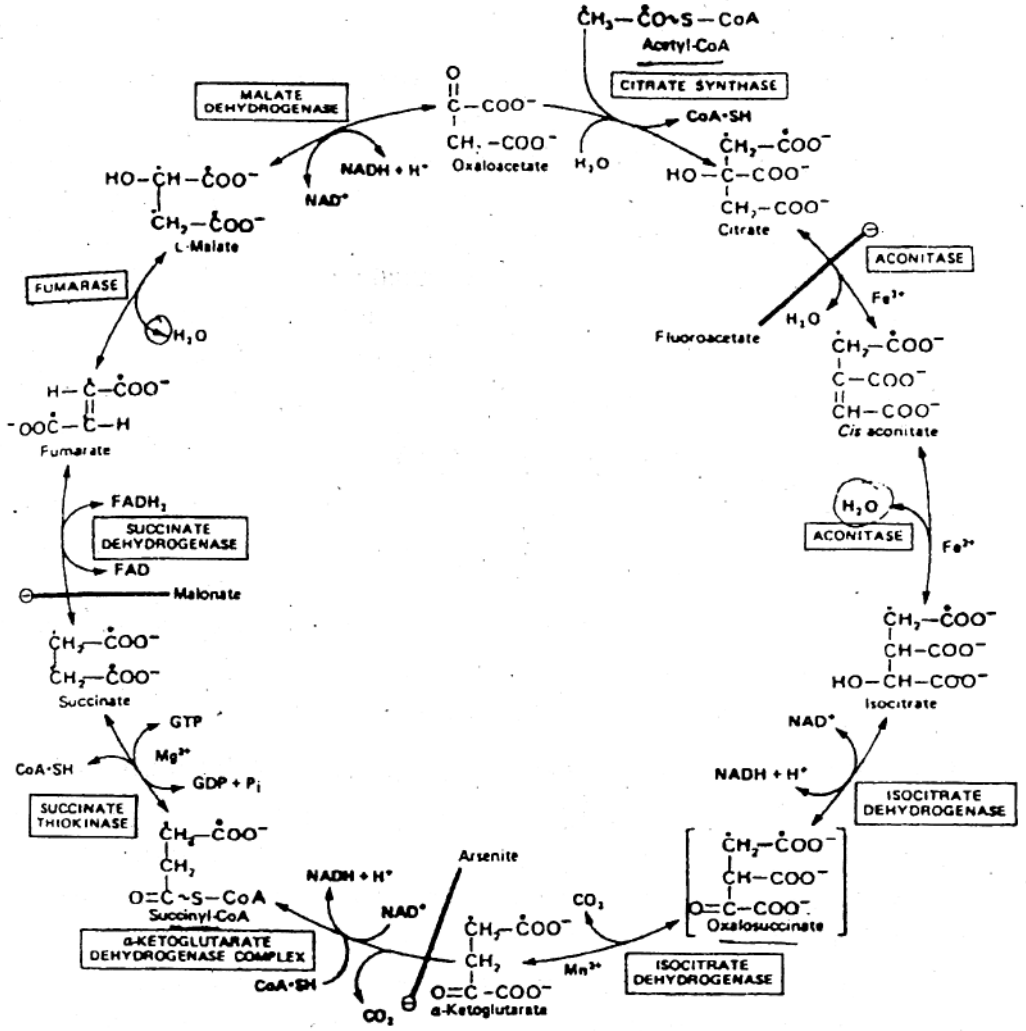


Figure : 1.3 The citric acid (Krebs) cycle. Oxidation of NADH and $FADH_2$ in the respiratory chain leads to the generation of ATP via oxidative phosphorylation. In order to follow the passage of acetyl-CoA through the cycle, the 2 carbon atoms of the acetyl radical are shown labeled on the carboxyl carbon (using the designation [•]) and on the methyl carbon (using the designation [•]). Although 2 carbon atoms are lost as CO_2 in one revolution of the cycle, these atoms are not derived from the acetyl-CoA that has immediately entered the cycle but from that portion of the citrate molecule which derived from oxaloacetate. However, on completion of a single turn of the cycle, the oxaloacetate that is regenerated is now labeled. Which leads to labeled CO_2 being evolved during the second turn of the cycle. Because succinate is a symmetric compound and because succinate dehydrogenase does not differentiate between its 2 carboxyl groups, "randomization" of label occurs at this step such that all 4 carbon atoms of oxaloacetate appear to be labeled after one turn of the cycle. During gluconeogenesis, some of the label in oxaloacetate is incorporated into glucose and glycogen. In this process, oxaloacetate is decarboxylated by release of the carboxyl group adjacent to the CH_2 group. As a result of recombination of the resulting 3-carbon residues in a process that is essentially reversal of glycolysis, the eventual location of label from acetate in glucose (or glycogen) is distributed in a characteristic manner. Thus, if oxaloacetate leaves the citric acid cycle after only one turn from the entry of labeled acetyl-CoA (acetate), label from the carboxyl carbon of acetate is found in carbon atoms 3 and 4 of glucose, whereas lebel from the methyl carbon of acetate is found in carbon atoms 1, 2, 5 and 6. For a discussion of the stereochemical aspects of the citric acid cycle, see Greville (1968). The sites of inhibition (⊖) by fluoroacetate, malonate, and arsenite are indicated.

via cisaconitate in presence of Aconitase. Isocitrate inpresence of NAD^+ and isocitrate dehydrogenase converted oxalosuccinate which is decarbonylated to α-ketaglutarate by the enzyme, Mg^{++} -dependent isocitrate dehydrogenase. α-ketaglutarate by the utilization of NAD^+ in presence of α-ketaglutarate dehtdrogenase complex oxidized to form succinyl CoA which is oxidized to succinate by sccinate thiokinase with the help of GDP. Sccinate in presence of succinate dehydrogenase & FAD forms fumarate. The enzyme fmarase form malate from fumarate which is ultimately oxidized to oxaloacetate by enzyme malate dehydrogenase and NAD^+. Again the reaction sarts and it is repeated in cycle order.

Table 1. Generation of high-energy bounds in the catabolism of glucose

Pathway	Reaction Catalyzed By	Method of ~Ⓟ Production	Number of ~Ⓟ Formed per Mole of Glucose
Glycilysis	Glyceraldehyde 3 phosphate dehydrogenase	Respiratory chain oxidation of 2 NADH	6*
	Phosphoglycerate kinase	Osidation at substrate level	2
	Pyruvate kinase	Osidation at substrate level	2
			10
Allow for consumption of ATP by reactions catalyzed by hexokinase and phosphofructokinase			–2
			Net 8
	Pyruvate dehydrogenase	Respiration chain oxidation of 2 NADH	6
	Isocitrate dehydrogenase	Respiration chain oxidation of 2 NADH	6
	α-ketaglutarate dehydrogenase	Respiration chain oxidation of 2 NADH	6
Citric acid	Succinate thiokinase	Oxiadation at substrate level	2
cycle	Succinate dehydrogenase	Respiratory chain oxidation of 2 $NADH_2$	4
	Malate dehydrogenase	Respiratory chain oxidation of 2 NADH	6
			Net 30
Total per mole glucose under aerobic conditions			38
Total per mole glucose under anaerobic conditions			2

*It is assumed that NADH formed in glycolysis transported into mitochondria via the malete shuttle. If the glycerophosphate shuttle is used. only 2 ~ □ would be formed per mole of NADH, the total net production being 36 instead of 38. The calculation ignores the small loss of ATP due to a transport of H^+ into the mitichondrion with Pyruvate and a similar transport of H^+ inthe operation of the malate shuffle, totaling about 1 mol of ATP.

1.3 Oxidation of fatty acids

β-oxidation of fatty acids

Several enzymes, known cllectively as "fatty acid oxidase" are found in ,otechondral matrix adjacent to the respiratory chain (in inner membrane). These catalyze the oxidation of acyl CoA to acetyl-CoA, the system being coupled with the phosphorylation of ADP to ATP.

After formation of acyl CoA and its penetration through the mitochondrial membreane via the carnitine transporter system, there follows the removal of 2 hydrogen atom from the 2 (α) and 3 (β) carbon atoms, catalyzed by acyl-CoA

dehydrogenase. The results in the formation of α, β unsaturated or Δ^3 unsaturated acyl CoA. The coenzyme for the dehydrogenase is a falvoportein, contaning FAD as prosthetic group, whose reoxidation by the repiratory chain requires the mediation of another flavoprotein termed electron transferring falvoprotein.

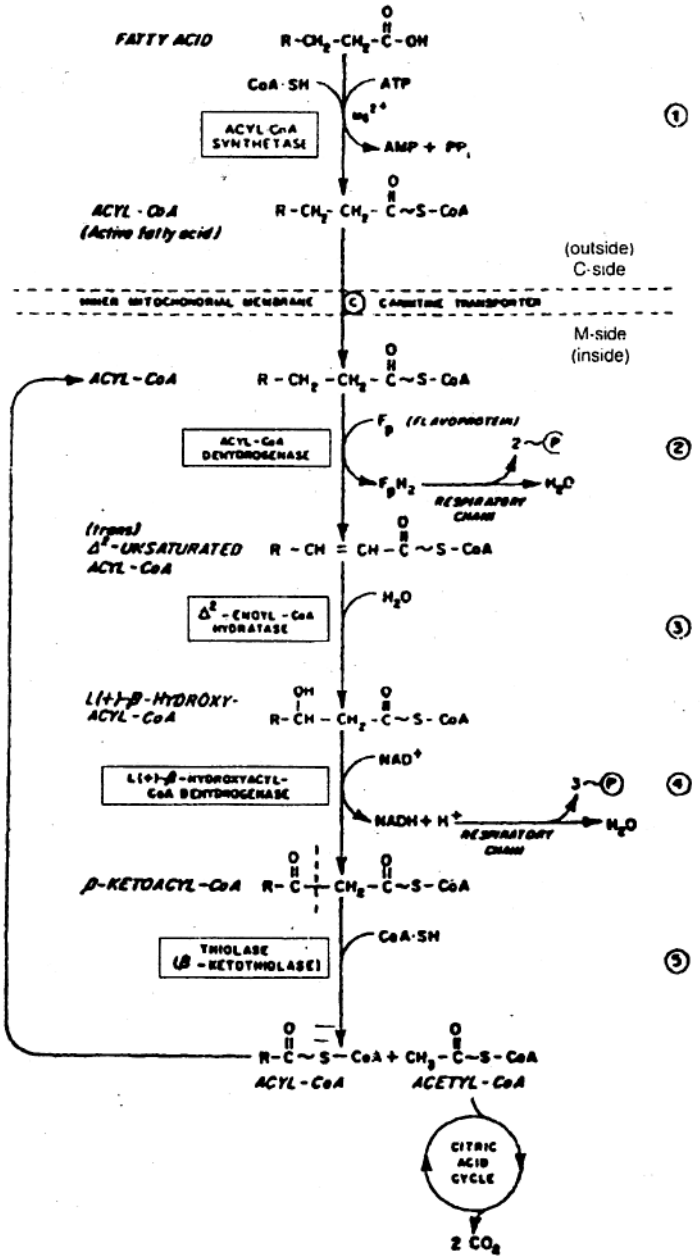


Figure 1.4 : β-oxidation of fatty acids. Long-chain acyl-CoA is cycled through reactions ② – ⑤ acetyl-CoA being split off each cycle by thiolase (reaction ⑤). When the acyl radical is only 4 carbon atoms in length, 2 acetyl-CoA molecules are formed in reaction ⑤

Water is added to saturate the double bond and form β-hydrxyacyl-CoA catalyze by the enzyme crotonase. By further dehydrogenation β-hydroxy derivative [β-carbon (β-hydroxylacyl CoA dehydrogenase)] form the corresponding β-ketol CoA compound NAD acts as coenzyme. β-keto acyl CoA splits at β-position by thiolase involving another molecle of Coa and forms acetyl CoA and acyl CoA contaning 2 carbon less that original acyl CoA. In this way a long chain fatty acid may completely degraded to acetyl CoA and this can be completely oxidized to CO_2 and water *via* citric acid cycle.

Fatty acids with an odd number of carbon atom are oxidized by this pathway.

Peroxisomal fatty acid oxidation :

This is the modified form of β-oxidation of long chain fatty acid by the peroxidomal enzyme shortens the CoA thioester of very long fatty acid to octanoyl CoA.

α and ω oxidation of fatty acid :

α-oxidation is the removal of one carbon at a time from the carboxyl end of molecule. It doesn't require CoA intermediate and does not generate high energy phosphate.

ω-oxidation is brought about by hdroxylase enzymes involving cytochorme P-450 in microsomes. The CH_3 group is converted to a CH_2OH group that sbsequently is oxidized to COOH, thus forming a dicatboxylic acid.

Oxidation of unsaturated fatty acid :

The CoA ester of unsaturated fatty acid are degraded normally by the enzyme responsible for β-oxidation until either a Δ^3-cis-acyl-CoA compound or a Δ^2-cis-acyl-CoA compound is formed, depending upon the position of the double bond. The former compound is isomerized to the corresponding Δ^2-trans-CoA state, which in turn is hydrated by Δ^2-enoyl-CoA hydratase to L (+) β-hydrozyacyl-CoA. The Δ^2-cis-acyl-CoA compound is first hydrated by Δ^2-enoyl-CoA hydratase to the D(–) β-hydroxyacyl-CoA derivative. This ndergoes epimerization [D(–)-β-hydroxyacyl-CoA epimerase] to give the normal L (+)-β-hydroxyacyl-CoA state is β-oxidation.

Microsomal Peroxidation of Polyunsatured fatty acids :

Polyunsaturated fatty acids in membrane initially destroyed by lipid peroxidation a hydrogen atom is removed, leaving a lipid free radical. After rearragement of the double bond, a lipid hydroperoxide or endoperoxide is formed by the addition of moleclar oxygen. NADPH dependent peroxidation of unsaturated fatty acids is catalyzed by microsomal enzymes.

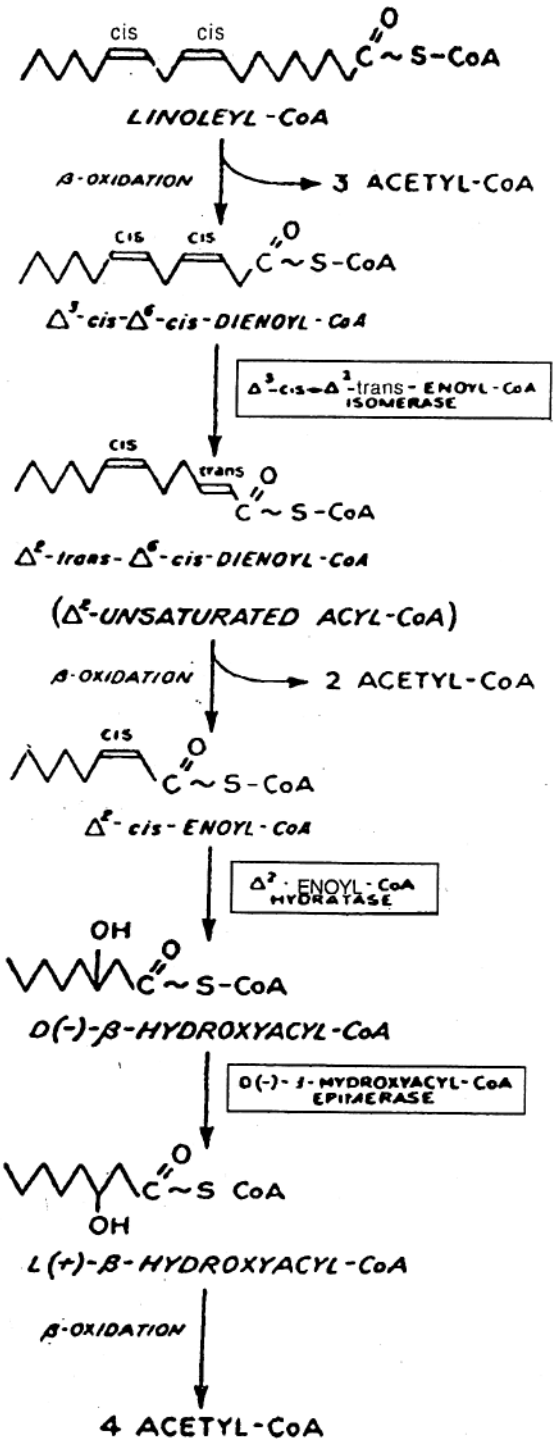


Figure 1.5 : Sequence or reactions in the oxidation of unsaturated fatty acids, e.g. Inoleic acid.

1.4 Phenylalanine

Phenylalanine is a glycogenic and ketogenic amino acid, through tyrosine & it forms (a) catecholamines, (b) melamins, (c) thyroid hormones, (d) fumerate and acetoacetate.

(a) The metabolism of catecholamines

The pathway for the formation of catecholamines are give below :

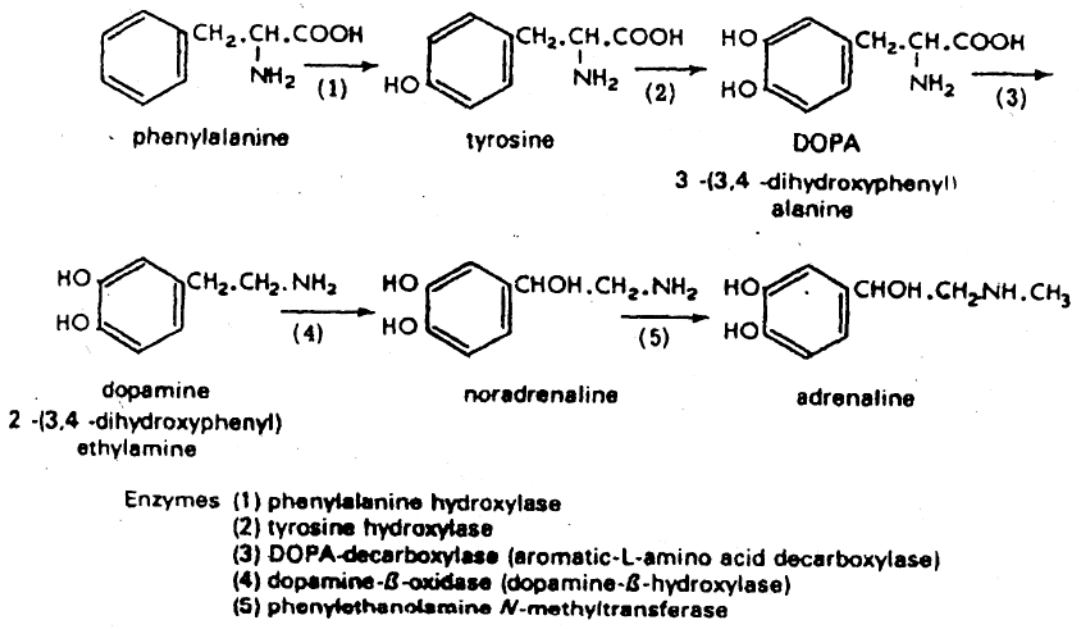


Figure 1.6 : The principal biosynthetic pathway for dopamine, noradrenaline and adrenaline

Catecholamines are speedily inactivated by methylation at eh orthohydroxyl group by catecholamine-O-methyltransaminase (COMT) & adenosylmethionin and oxidative deaminetion by monoamine oxidase (MAO) and aldehyde dehydrogenase. Conjugated and unconjugated end products such as metnoradrenatine, metadrenatine, vanilic acid are excreted through urine.

(b) Melanin :

This is protien-bound polymeric pigment present in skin melanocytes, hair, iris, choroids, retinal pigmented cell and substantia nigra.

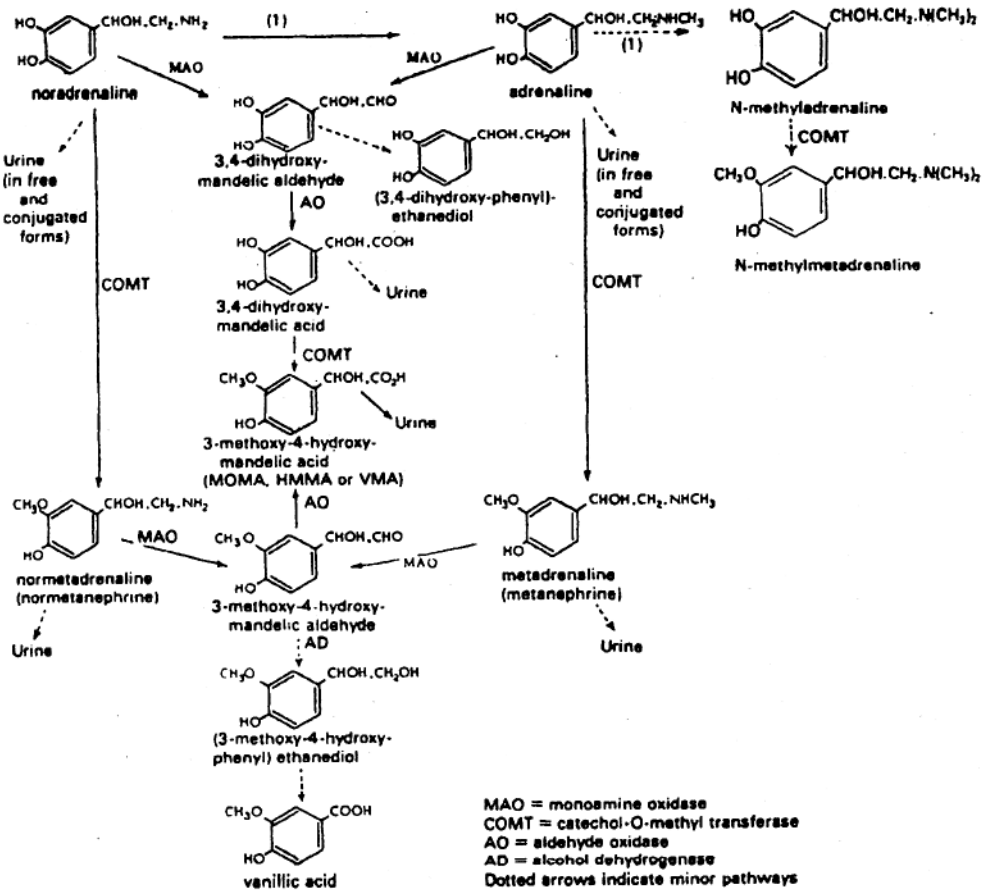


Figure 1.7 : The breakdown of adrenaline and noradrenaline after Axelrod (1963, 1965), Nodified from Crossland, 1967, by kind permission of Messrs Butterworths.

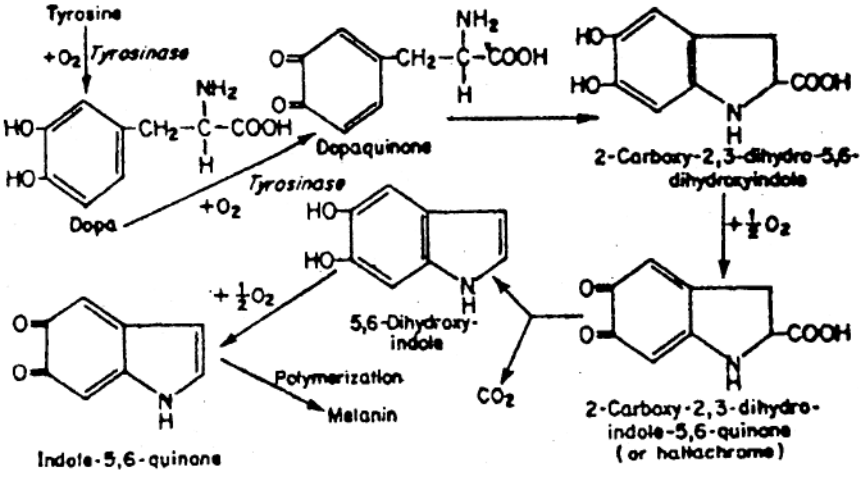


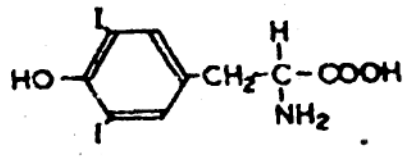
Figure 1.8 : Possible pathway for melanin synthesis

(c) Thyroid hormones :

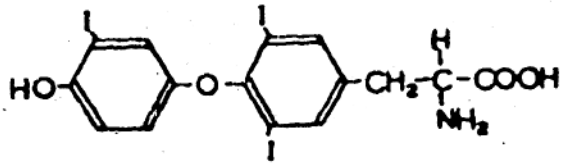
Phenylalanine by phenylalanine hydroxylase are converted to tyrosine. By iodination at 3 position forms monoiodotyrosine (MIT) and then by further iodination at the 5th position it forms diiodotirosine (DIT). Then by coupling of one mono-and one-di-or by coupling of 2 diiodotyrosine formation of triiodtyrosine (T_3) or tetraiodo thyronime (T_4) formation take place. The enzyme responsible for the T_3 and T_4 formation are peroxidase.

The thyroid hormones themselves are deiodinated in the liver, salivary glands and kiney. The iodine can be used again. the remainder of the molecules are disposed of as conjugated in the urine.

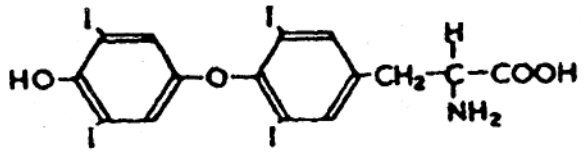
(d) Phenylalanine and tyrosine form acetoacetate and fumerae. Then through TCA cycle they are completely degrated or through this cycle they synthesize, or lipid or protein.



3,5-Diiodotyrosine



3,5,3'-triiodothyronine (T_3)



3,5,3',5'-tetraiodothyronine (thyroxine, T_4)

Figure 1.9 : Iodinated derivatives of tyrosine

1.5 Catabolism of purine

Primates, dalmation dogs and uricotelic animals, catabolize purine to uric acid. Human eliminates 600-800mg. of uric acid daily; 80 to 90% through urine and rest through bile. Catabolism of purinase to uric are given in figure 1.12.

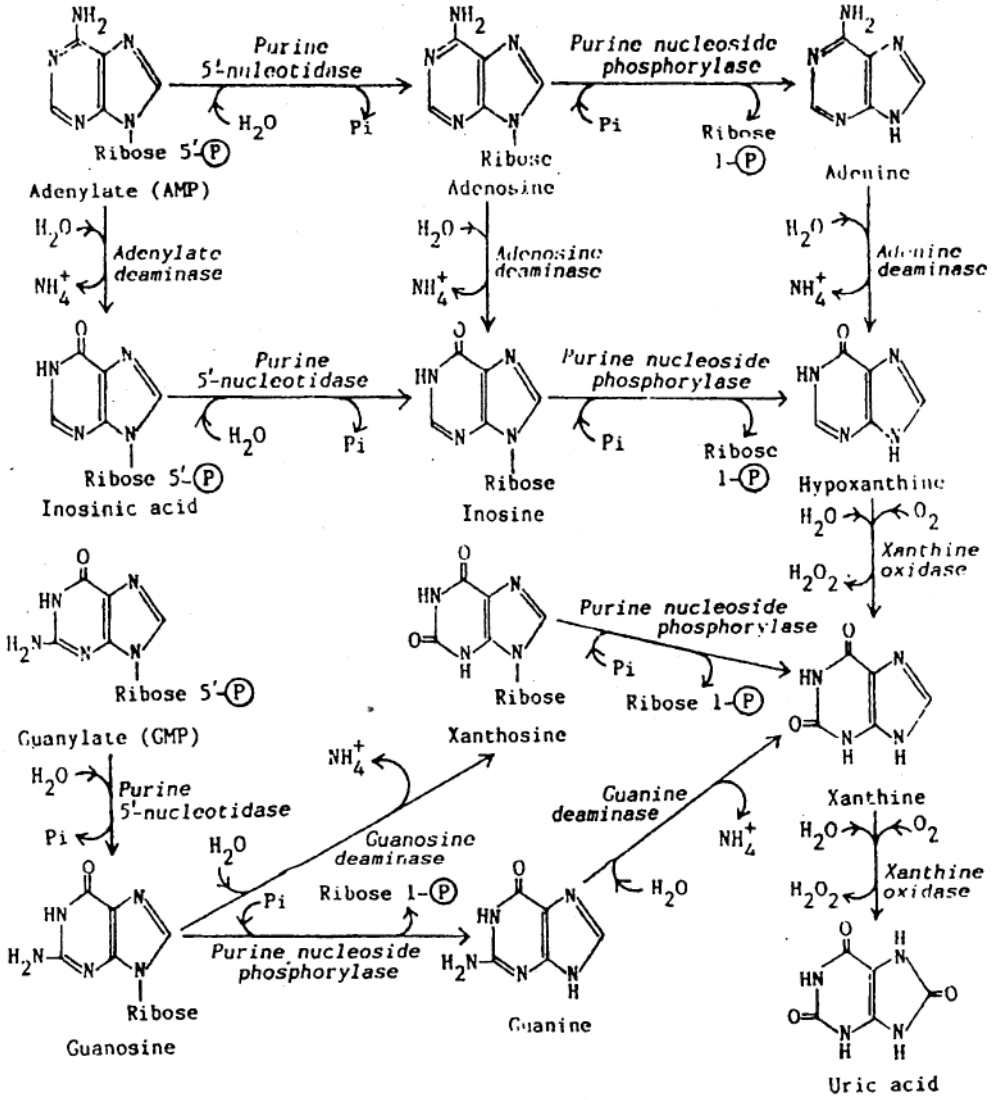


Figure 1.10 : Catabolism of purines to uric acid

1. Transamination

Transamination transfer the amino group of one amino acid to a ketoacid, changing the later into a new amino acid and original amino acid into a new keto acid.

Alamine + α-ketogutarate = Pyruvate + glutamate

Except threonine, lysine, proline and hydroxyprolinem all α-amino acid can participate in transaminations and change into respective keto acid by donating α-amino groups. Transaminations are double displacement type of bisubstrate reations. The two substance (i.e. amino acid and keto acid) bind separately and successively

with the prosthetic group of the enzyme. L-glutamic acid first combines with the enzyme bound pyridoxal phosphate and forms enzyme bound Schiff base with the liberation of water. It is hydrolysed to release glutamic acid as a product leaving enzyme bound pyridoxamine phosphate (PMP). The pyruvic acid (second substrate) binds with the PMP to form a new schiff base. After intermolecular rearrangement, dissociates with the help of water forms alamine and enzyme bound pyridoxal phosphate.

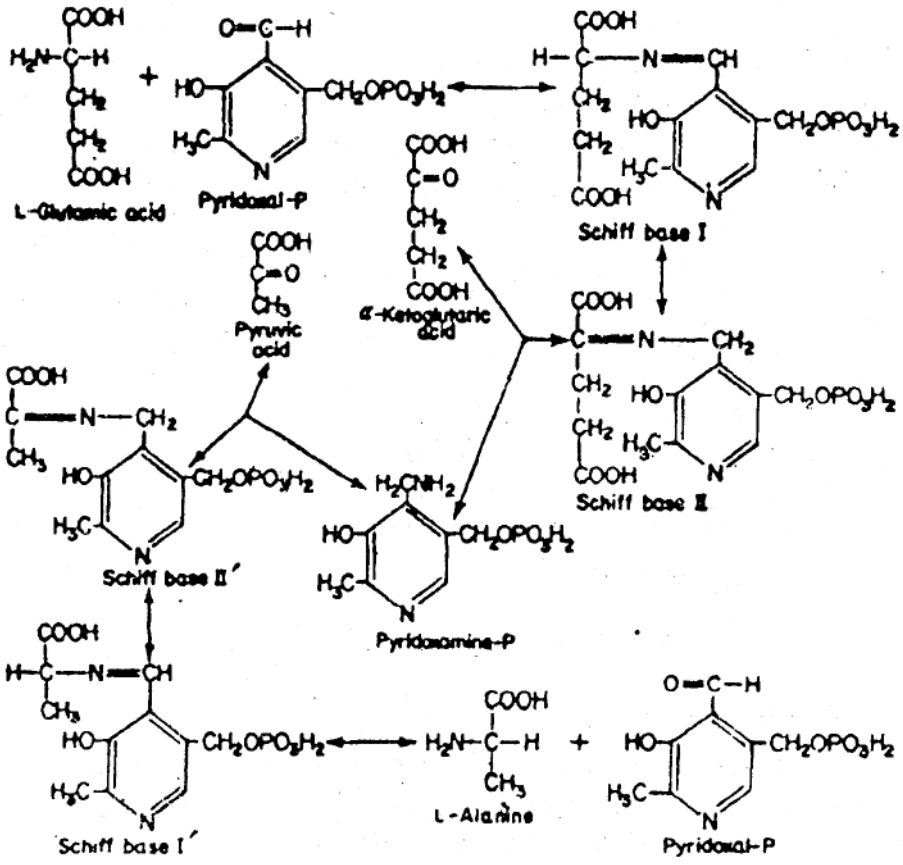


Figure 1.11 : Mechanism of transamination by glutamate pyruvate aminotransferase

The reaction may be summarized as follows :

α-Amino acid, + enzyme PLP = a-keto acid, + Enzyme PMP.

α-keto acid 2 + enzyme PMP = α-amino acid 2 + enzyme PLP.

Each ransaminase specific for each amino acid as substrate for the first stage of transmination.

Cliunically two imporant transaminases are–

Serum-glutamate oxaloaretate ransaminase (SGOT)

Serum-gltamate pyruvate transminase (SGPT).

2. Transdeamination

In mammulian tissue particlarly the liver amino acids are deaminated by transamination with the help of aminotransferases (transaminase) followed by oxidative deamination of the resulting L-glutamate in the mitochondria by a polymeric and allosteric enzyme. L-glutamate dehydrogenase, present in liver mitochondria requires NAD^+ or $NADP^+$ by the action of L-glytamate dehydrogenase.

Dehydrogenase is reversible. It may also synthesize glutamate by the reductive animation of α-ketoglutarae. The better is then spontaneously hedrolyzed into α-ketoglutamate and ammonia by cytoplasmic glucomate dehydrogenase with has low Km ammonia. It mainly catalyzed the biosynthesis of glutamate sing NADPH, NADH, ATP and GTP. It is allosterically activated by NAD^+, $NADP^+$, ADP, GDP and AMP.

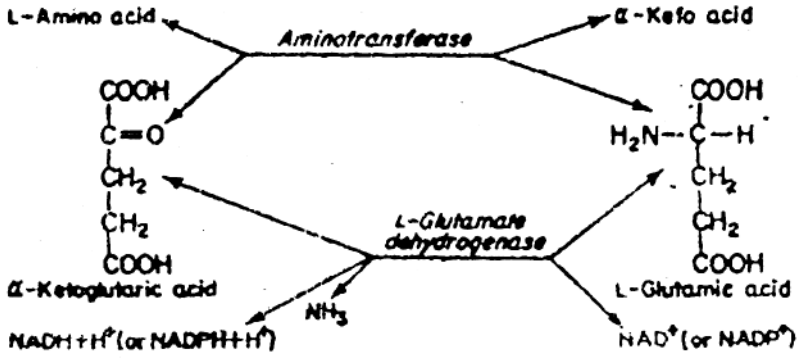


Figure 1.12 : Transdeamination of L-amino acid

3. Oxidative deamination

Oxidative deamination is the direct removal of the amino group as ammonia by flavoprotein containing enzyme amino acid oxidase. It is found in peroxisomes of mammalion liver and kidney cells. It ocurs in mamalian renal mitochondria and microsome and too poor inmammalian livercell.

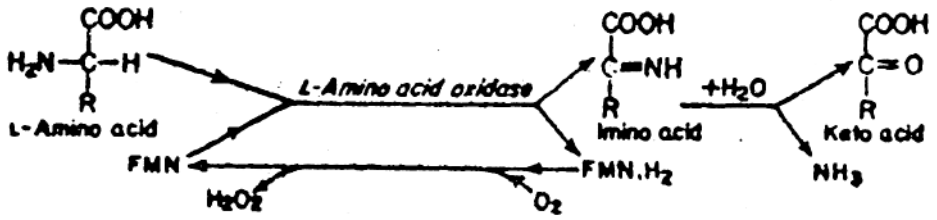


Figure 1.13 : Oxidative deamination by L-amino acid oxidase

It oxidize corresponding imino acid by transferring reducing equivalent from the amino acid to the flavin mucleotide in its prosthetic group. The amino acid reacts spontaneously with water to give an α-keto acid and ammonia. The reduced nucleotide of the prosthetic group is reoxidised directly by molecular O_3, producing H_3O_3.

1.6 Terminal questions

1. Write the significance of the citric acid cycle.
2. State schematically the generation of high energy bonds in the catabolism of glucose.
3. Discuss about the glycolysis.
4. Write about the glycogenesis.
5. Discuss critically about the hexose monophosphate shunt pathway.
6. Write about the metabolic significance of hexose monophosphate shunt pathway.
7. Discuss about the role of phenylalanine for the generation of biologically active molecules.
8. Discuss about the sybthesis & function of neutrotransmitters / T_3 & T_4 or thyroid hormones/melanin from phenylalanine.
9. Write about the catabolism of purine.
10. Discuss about the β-oxidation of fatty acid.

Unit 2 ❐ Biological Oxidations

Structure

2.1 Introduction

You will read about biological oxidations in this unit. Oxidation and reductions in tissues consist of transfers of *reducing equivalence* (viz., electrons and H^+) from one substrate to another; the electron-donor substrate is thereby oxidized while the electron-acceptor substrate gets simultaneously reduced. In this way, electrons may flow between successive substrates of one or more pairs, depending mainly on the relative electron-affinities of the electron-donor and the electron-acceptor substrates each time. In *aerobic metabolic pathways,* molecular O_2 is required to serve as the final electron-acceptor during such electron-flow from intermediate to intermediate of the pathway. But in *anaerobic metabolism,* no O_2 is required to act an electron acceptor. Glycolysis and TCA cycle are examples of respectively anaerobic and aerobic oxidative pathways. Both anaerobic and aerobic oxidations can generate high amounts of free energy which may be trapped by forming high-energy bonds such as those of ATP and GTP. Thus, electrons removed from substrates of aerobic oxidations are transported by successive electron-carriers of the *mitochondrial respiratory chain,* ultimately to be accepted by molecular O_2; the energy released during such electron transports to O_2 is harnessed to phosphorylate ADP to ATP by *oxidative phopshorylation.*

Objectives

Reading of this unit should enable you to :

- Understand the nature and actions of different classes of oxidoreductases,

- Know the natures of redox potentials and electron-transfer potentials and their roles in electron transfers between substrates.
- Describe different components of the mitochondrial electron transport chain and their respective roles in electron transport.
- Explain the free energy changes during electron flow along different respiratory chain complexs of the electron transport chain.
- Understand the difference between substrate-level and oxidative phosphorylations,
- Describe the structural organization of ATP synthase of inner mitochondrial membrane particles and its mode of action in oxidative phosphorylation of ATP.
- Explain the coupling of the mitochondrial oxidation and ATP synthesis in terms of the chemiosmotic theory.
- Describe the redox loop mechanism and the proton pump mechanism to explain the formation of a transmembrane proton gradient with the energy released by mitochondrial electron transport.
- Discuss the regulation of oxidative phosphorylation.

2.2 Oxidoreductases

Oxidation of a substrate consists of either the removal of reducing equivalent (viz, electrons and H^+) or the addition of O_2 to it. Reduction is the reverse process of addition of reducing equivalents to a substrate. Oxidation and reductions are mostly catalyzed in biological systems by specific enzymes called *oxidoredctases*. You will know the natures and modes of action of different classes of oxidoredctases from their brief classification given below.

2.2.1 Oxygenases

They oxidize their respective substrates by incorporating oxygen into the molecules of the latter. They belong to two main subclasses.

(a) Dioxygenases :

These incorporate both oxygen atoms of O_2 into each substrate molecule, often breaking the C-C bond between the oxygen-accepting Carbons in the latter; e.g., tryptophan 2, 3-dioxygenase and homogentisate, 1, 2-dioxygenase of liver oxidizing respectively tryptophan and homogentisate.

(b) Monooxygenases or hyodroxylases :

These catalyze the hydroxylation of their substrate by incorporating it in one of the oxygen atoms from O_2 and similtaneously reduce the other oxygen atom to H_2O

by reducing equivalents from specific coenzymes/cofactors such as cytochrome b_5 and tetrahydrobiopterin; e.g. phenylalanine hydroxylase of liver and dopamine β-hydroxylase of adrenal medulla, hydroxylating their substrates respectively using tetrahydrobiopterin and L-ascorbate.

2.2.2 Oxidases

The catalyze the oxidation of their respective substance by transferring reducing equivalents (electrons and H^+) from the substrate to molecular O_2 alone, producing H_2O in most cases.

(a) Copper-containing aerobic oxidases :

These are copper-protein complexes. The Cu^{2+} ion of the enzyme receives the electron from the substrate (AH_2) to from Cu^+ which then transfers the electron to molecular O_2 directly; the corresponding proton (H^+) from the substrate follows the electron to join the O_2 to form H_2O.

$$AH_2 + 2C^{2+} \rightarrow A + CU^+ + 2H^+$$

$$2Cu^+ + 2H^+ + ½O_2 \rightarrow 2Cu^2 + H_2O$$

E.g., cytochrome oxidase, a Cu^{2+} –heme-protein complex of inner mitochondrial membrane, oxidized reduced cytochrome C.

(b) Flavoprotein oxidases :

These are flavoproteins (EP) with FMN or FAD as prosthetic group. They oxidize their substrates (AH_2) by transferring reducing equivalents (e and H^+) from the latter to their own prosthetic groups, reducing them to $FMNH_2$ or $FADH_2$. $FMNH_2$ of $FADH_2$ is reoxidized by the transfer of reducing equivalents directly to molecular O_2, producing H_2O_2; eg. L- and D-amino acid oxidases of respectively kidney and liver, bearing FMN and FAD respectively; xanthine oxidase of liver, containing Fe^{3+}, Mo^{6+} and FAD. Flavoprotein oxidases can also transfer the reducing equivalents to methylene blue, reducing the latter.

$$AH_2 + FP \rightarrow A + FPH_2$$

$$FPH_2 + O_2 \rightarrow FP + H_2O_2$$

2.2.3 Hydroperoxidases

These hemoproteins of hepatic and renal peroxisomes, thyroid cells, granulocytes and erythrocytes serve to reduce H_2O_2 with reducing equivalents from specific substrates such as glutathione, halides (Cl^- or I^-) and cytochrome C, which are thereby oxidized; eg., glutathione peroxidase (RBC) reduces H_2O_2 while oxidizing

glutathione (GSH) to GS-SG, myeloperoxidase (granulocytes) oxidized Cl^- to hypochlorite (Cl^-) for reducing H_2O_2, and iodine peroxidase (thyroid) oxidized I^- to hypoiodite (IO^-) for reducing H_2O_2.

$$2GSH_2 + H_2O_2 \rightarrow GS - SG + 2H_2O$$

$$Cl^- + H_2O_2 \rightarrow ClO^- + H_2O_2$$

$$I^- + H_2O_2 \rightarrow IO^- + H_2O_2$$

Catalases of peroxisomes, in addition, can reduce H_2O_2 to H_2O with electrons from another H_2O_2 molecule :

$$H_2O_2 + H_2O_2 \rightarrow 2H_2O_2 + O_2$$

2.2.4 Dehydrogenases

They transfer reducing equivalents from their substrates to specific electron-aceptors other than molecular O_2 |——| never to the latter. The electron-acceptor thus reduced gets reoxidized by transferring the reducing equivalents in turn to some other electro-acceptors. A chain of such electron-transfers may be carried out by the serial actions of several dehydrogenases (an *electron transport chain*) until an oxidase transfers the electrons from the preceding electro-acceptor to molecular O_2 directly.

(a) Pyridine-linked dehydrogenases :

While oxidizing their substrates, these enzyme transfer two electrons and a proton |——| as a hydride ion (H^-) ion |——| from the substrate to a pyridine nucleotide (either NAD^+ or $NADP^+$) serving as the electron-acceptor coenzyme; the other proton from the substrate (AH_2) is released in the medium while the coenzyme is reduced to BADH or NADPH.

$$AH_2 + NAD^+ \rightarrow A + NADH + H^+$$

$$AH_2 + NADP^+ \rightarrow A + NADPH + H^+$$

The reduced coenzyme is next released from the dehydrogenase, joins some other dehydrogenase as the electron-donor coenzyme of the latter, and gets reoxidized by donating its hydride ion to substrate (B) of that enzyme and reducing it.

$$B + NADH/NADPH + H^+ \rightarrow BH_2 + NAD^+ / NADP^+$$

In most cases, NADH gets reoxidized by donating its H^- ion to an electron-acceptor of the *mitochondrial respiratory chain* (see 2.4). But NADPH is mainly reoxidized by donating its H^- ion to specific substrates/intermediates of the *reductive synthesis* pathways such as those for palmitic acid and cholesterol synthesis.

NAD^+ *–dependent dehydrogenases* include cytoplasmic L-lactate dehydrogenase and glyceraldehyde 3-phosphate dehydrogenase for glycolysis, mitochondrial pyruvate dehydrogenase, α-ketoglutarate dehydeogenase and malate dehydrogenase for aerobic oxidative metabolism of pyruvate, and mitochondrial 3-hydroxycyl-CoA dehydrogenas for beta-oxidation, *$NADP^+$ – dependent dehydrogenases* include cytoplasmic glucose 6-phosphate dehydrogenase and 6-phosphogluconate dehydrogenase of HMP shunt.

(b) Flavin-linked dehydrogenases :

Their prosthetic groups consists of riboflavin-derivatives such as FMN and FAD. While oxidizing their substrates (AH_2), these enzymes transfer two elections and two protons from the substrate to their flavin prosthetic grops, reducing the latter to $FMNH_2$ and $FADH_2$ respectively.

$$AH_2 + FM/FAD \rightarrow A + FMNH_2/FADN_2$$

FMN-dependent dehydrogenases include NADH dehydrogenase of mitochondrial respiratory chain. *FAD-dependent dehydrogenases* include mitochondrial succinate dehydrogenase of the TCA cycle and acyl –CoA dehydrogenase of beta-oxidation. Dihydrovatate dehydrogenase, a mitochondrial enzyme for pyrimidine synthesis bears both FMN and FAD as its prosthetic groups, along with Fe^{3+}.

Most flavin-dependent dehydrogenases get their reduced prosthetic groups ($FMNH_2$) oxidized be transferring their reducing equivalents through mitochondrial iron-sulfur proteins to coenzyme Q of the respiratory chain.

$$FMNH_2 / FADH_2 + 2Fe_2^{3+}S_2 \rightarrow FMN/FAD + 2Fe_2^{3+}S_2 + 2H^+$$

$$CoQ + 2Fe_2^{+2}S_2 + 2H^+ \rightarrow CoQH_2 + 2Fe_2^{3+}S_2$$

(c) Iron-sulfur proteins :

These are electron-transporting components of the mitochondrial respiratory chain (see 2.4). They occur in the inner mitochondrial membrane mainly as protein-bound $2Fe_2S_2$ and Fe_4S_4 clusters, in which respectively 2 and 4 nonheme iron atoms are linked respectively to as many acid-labile inorganic sulfides and are also coordinately bound to the side chain SH groups of specific cystaine residues of apoproteins. Iron-sulfur proteins transport electrons from the reduced flavin prosthetic groups, such as $FMNH_2$ of NADH dehydrogenase and $FADH_2$ of succinate dehydrogenase, to coenzyme Q of the mitochondrial chain. During such process, the Fe^{3+} ions-sulfur proteins get first reduced to Fe^{2+} ions on accepting electrons from the reduced flavins and next oxidized back to Fe^{3+} ions by donating those electrons to CoQ.

(d) Cytochromes :

These are hemoproteins containing Fe^{3+} in their oxidized ferricytochrome ($CyFe^{3+}$) forms and Fe^{2+} in their reduced ferrocytochrome ($CyFe^{2+}$) forms. In aerobic cells such as hepatocytes, myocytes, adipocytes and cardiac myocytes, cytochromes b_H, b_L, b_{560}, C_1, C, a and a_3 occur as components of the mitochondrial respiratory chain; they differ from each other in their apoproteins (apocytochromes) and/or in their heme prosthetic groups (eg. *heme a, heme b* and *heme c*). Ferricytochrome ($CyFe^{3+}$) gets reduced to ferrocytochrome ($CyFe^{2+}$) receiving electron from an iron-sulfur cluster or another ferrocytochrome, and gets reoxidized $CyFe^{3+}$ by donating that electron to CoQ or iron-sulfur cluster or another ferricytochrome or molecular O_2. Of the cytochromes of the inner mitochondrial membrane, cytochromes a and a_3 together constitute an oxidase, *cytochrome oxidase* (Cy a-a_3), that transfers the electron directly to molecular O_2; but the others function as dehydrogenases, not donating electrons directly to molecular O_2 and instead, carrying electrons in a specific serial order to reach the electrons to Cy a-a_3 for the final flow of those electrons to molecular O_2 : $CoQ \rightarrow Cyb_L \rightarrow Cyb_H \rightarrow Cyc_1 \rightarrow Cyc \rightarrow Cya \rightarrow Cya_3 \rightarrow O_2$.

2.3 Redox potential

Each substance occurs in any system in two forms : its oxidized or oxident form can accept electrons from another substance, thereby oxidizing the latter and itself getting reduced while its reduced or *reductant* form can donate electrons to some other substance, thereby reducing the latter and itself getting oxidized to change into the oxidant form. The oxidant and reductant forms of any substance remain in equilibrium with one another and constitute a *conjugate redoxpair or redox couple;* e.g., pyruvate (oxidant)/lactate (reductant), $FAD/FADH_2$, $FMN/FMNH_2$, $NAD^+/NADH$, $NADP^+/NADPH$, $Fe_4^{3+}S_4$, oxaloacetate/malate, fumarate/succinate, $CoQ/CoQH_2$, ferricytochrome a/ferrocytochrome a.

Redox potential or oxidation-reduction (OR) potential of any substance is a measure of its electron-affinity. As the electron-affinity is determinant for the donation or acceptance of electrons by a substance, its redox potential gives an estimate of its tendency to accept or donate electrons.

Redox potential of substance is estimated by placing an electrode in an equimolar solution of its oxidant and reductant forms and measuring its potential against a standard hydrogen electrode. For measuring the *standard redox potential* (*Eo*) of a substance at pH o.o, a *standard reference half-cell* is constituted by dipping a hydrogen electrode in a 1 molar solution of H^+ at 25°C and bubbling H_2 gas at 1 atmosphere pressure through hat solution (Fig. 2.1) A *sample half-cell* is also constituted by

immersing another electrode in an equimolar solution of the oxidant (Ox) and reductant (Red) forms of the substance, each of a concentration of 1M. The electrodes in the two half-cells are connected to a voltmeter while the electrical continuity is maintained by connecting the solutions of the half-cells by an agar or salt bridge. The voltmeter reading gives the E_0 of the substance under in vestigation. The *standard redox potential of biological system* (E'_0) is, however, determined using a standard reference half-cell containing 10^{-7} M solution (pH 7.0) of H^+ ions.

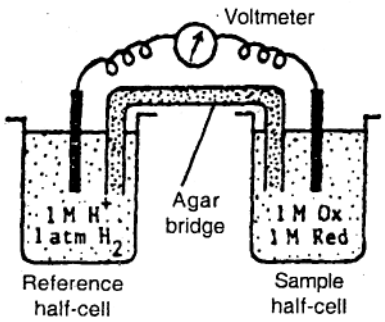


Figure 2.1 : Set-up for measuring the Eo of a redox couple
(From D. Das, Biochemistry, Academic Publishers, 2002)

For a substance with a *higher electrons-affinity* than H_2, electrons donated by H_2 of the reference half-cell will flow from there to the sample half-cell to join the oxidant member (O_x) of that substance; this will make the sample half-cell electrode positive to the reference electrode and the E'_0 and from the voltmeter will be positive. Thus, a higher, i.e., more positive or less negative, redox potential indicates a *stronger electron—affinity,* and consequently a *stronger oxidizing capacity* of the oxidant (O_x) member of the relevant redox couple. For a substance with a *lower electron-affinity* than H_2, electrons donated by its reductant form will flow from the sample half-cell to the reference half-cell to join H_2 ions there; this will make the sample half-cell electrode electronegative to the reference electrode and the E'_0 given will become negative. Thus, a lower, i.e., more negative, redox potential indicates a *lower electron-affinity* of the reductant. It follows, therefore, that :

(*i*) the higher or more positive the E'_0 of a redox pair the stronger is its oxidant member (e.g., ferricytochrome c) as an oxidizing agent and the weaker is its reductant member (e.g., ferrocytochrome c) as a reducing agent;

(*ii*) the lower or more negative the E'_0 of a redox pair, the weaker is its oxidant member (e.g. $NADP^+$) as an oxidizing agent and the stronger is its reductant (e.g. NADPH) as a reducing agent;

(*iii*) in any oxidation-reduction reaction, electrons will flow form reductant member of a redox couple with a lower E'_0 to oxidant member of another redox couple with

a lower E'_0 to the oxidant member of another redox couple with a higher E'_0 |—| this would oxidize the reductant of the first redox couple to its oxidant form and would reduce the oxidant of the second redox couple to its reductant form.

To cite an example, isocitrate gets oxidized to α-Ketoglutarate by donating electrons to NAD^+ and reducing the latter to NADH, because the E'_0 (–0.38 volt) of E'_0 the α-Ketoglutarate / isocitrate redox couple for lower than the (–0.35 volt) of the NAD^+/NADH redox couple.

However, the redox potentials of two redox couples would be the sole determinant of electron-flows between them, only when the oxidant and reductant members of each couple occur in-equimolar concentrations, i.e., in the so-called standard conditions. Whenever the oxidant and reductant members of either couple are not present in the system in equimolar concentrations, electron-flow between the two redox couples will depend on the resultant effect of their respective E'_0 values and the ratio of oxidant and reductant concentrations of each couple, and is consequently determined by the determined by the difference between the *electron transfer potentials (E)* of the two couples. For each couple. E is given by the following *Nernst equation.* Where *R* is the molar constant (8.314 joules/degree/mol). *T* is one absolute temperature (assumed here as 298 K or 25°C), *F* is the Faraday constant (96487 coulombs/mol) and *n* is the number of electrons transferred.

$$E = E'_0 + \frac{RT}{nF} \ln \frac{[Ox]}{[Red]}; \text{of}, E = E'_0 + \frac{0.059}{n} \log \frac{[Ox]}{[Red]}$$

Thus, a change in the relative concentrations of the oxidant and reductant members of one or both the redox couples may after the difference (ΔE) between their electron transfer potentials so as to reverse the direction of electron-flow between those couples. For example, during anaerobic glycolysis, pyruvate is reduced to lactate by the reducing equivalents (*e* and H^+) from NADH, because (*i*) the pyruvate / lactate redox couple has a higher E'_0 of –0.185 volt than the E'_0 (–0.315 volt) of the NAD^+ / NADH couple and (*ii*) the [OX]/[Red] ration of the first couple is increased while that of the second one is lowered by the rise in cytoplasmic concentrations of respectively pyruvate and NADH of glycolysis, thus lowering the electron transfer potential of the second couple much below that of the first one. In the aerobic condition after glycolysis, on the contrary, (*i*) the difference ($\Delta E'o$) between the *E'o* values of the two redox couples remains as before, but (*ii*) the [Ox]/[Red] ratio of the pyruvate/lactate couple is lowered by the accumulated lactate while that of the NAD^+/NADH couple is raised by the reoxidation of the accumulated NADH; these result in a rise in the electron transfer potential (*E*) of the latter couple for above that of the former, bringing about electron flow from lactate to NAD^+ and redoxidation of lactate to pyruvate.

2.4 Mitochondrial respiratory chain

Mitochondrial *respiratory or electron transport chain* consists of a specifically arranged set of electron transporters in the inner mitochondrial membrane for carrying out aerobic oxidations of NADH and reduced flavins ($FMNH_2$ and $FADH_2$) by transporting reducing equivalents (*e* and H^+) from them to molecular O_2. Each electron transporter of this chain acts as a *redox couple* |——| you have read about redox couples in the preceding section 2.3 on redox potential. Such electron carriers are located in proper and steric relation to each other in order of progressively rising redox potentials. This enables electrons to flow from the reduction member of each such redox couple of the chain to the oxidant member of the next one having a higher E_0 unit they are finally donated to molecular O_2. The mitochondrial ET chain they constitutes the *final electron-flow* path from tissue substrates to molecular O_2. Two components of the ET chain, viz., cytochrome c and co-enzyme Q, are not intergral proteins of the membrane |——| cytochrome c is a peripheral protein on the cytoplasmic (outer) surface of the inner membrane while coenzyme Q (ubiquinone) is nonpolar, lipid-soluble benzoquinone derivative, not bound to any protein and diffusing freely in that membrane. Other components of the ET chain, viz, NADH dehydrogenase, succinate dehydrogenase, iron-sulfur clusters and cytochromes b_{560}, b_{11}, b_1, c_1, a and a_3, remain bound to specific integral proteins of the inner membrane; each such set of intergral protein-bound components |——| except cytochrome C and CoQ |——| is organized into one or other of four *respiratory chain complexes* which together constitute a *respiratory chain assembly.* Thousands of such respiratory chain assemblies may occur per cubic micrometer of the inner membranae of an aerobic cell.

2.4.1 Respiratory chain complexes

The integral protein-bound components of each ET chain are organised into the following four respiratory chain complexes in the inner membrane. Components of each such complex are oriented asymmetrically in the membrane and arranged in specific steric relations to each other and in an ascending order of their redox potentials.

a. Complex I or NADH-Q reductase :

This large transmembrane oligomeric integral protein (MW 850 Kdal) has FMN and several FE_2S_2 and FE_4S_4 iron-sulfur clusters as its prosthetic groups. It functions as NADH dehydrogenase transferring reducing equivalents (*e* and H^+) from NADH through FMN and iron-sulfur c lusters of its prosthetic groups to coenzyme Q, thereby oxidizing NADH to NAD^+ and reducing CoQ (ubiquinone) to $CoQH_2$ (ubiquinone).

$$NADH + H^+ + FMN \rightarrow NAD + FMNH_2;$$

$$FMNH_2 + 2Fe_2^{+3}S_2 \rightarrow FMN + 2Fe_2^{2+}S_2 + 2H^+;$$
$$COQ + 2Fe_2^{2+}S_2 + 2H^+ \rightarrow COQH_2 + 2Fe_2^{3+}S_2$$

You have already read that *COQ* is a lipid-soluble benzo-quinone derivative, not bound to any protein and not belonging to any of the respiratory chain complexes. Diffusing freely in the lipid bilayer of the inner membrane, CoQ can make alternate contacts with other electron transporters such as iron-sulfur clusters and cytochromes of different respiratory chain complexes and transports electrons between them. In this way, it functions as a mobile electron transporter between either complexes I and III or complexes II and III. Mammalian CoQ is called Q_{10} because of then C_5 isoprenoid units forming its long nonpolar sidechain.

b. Complex II or succinate-Q reductase :

This oligomeric nonspanning integral protein (MW 127 Kdal) consists of a dimeric flavoprotein called *succinate dehydrogenase* with FAD prosthetic group, a few Fe_4S_4 ion-sulfur clusters, and cytochrome b_{560} with a heme to prosthetic group. Succinate dehydrogenase oxidized succinate to fumarate using FAD as the electron acceptor. The $FADH_2$ is reoxidized by the transfer of its reducing equivalents through cytochrome b_{560} and iron-sulfur clusters to CoQ, reducing the latter to QH_2. You should bear in mind that electrons cannot pass from succinate to COQ through complex I because succinate has a higher *E'o* (+ 0.30 V) than *E'o* (– 0.315) of the NAD^+ of that complex.

c. Complex III for QH_2-cytochrome c reductase :

This large oligomerie integral protein (MW 280 Kdal) consists of a single transmembrane *apocytochrome b* bearing two *heme b* prosthetic groups at two different sites, one *cytochrome* c_1 having a *heme c* prosthetic group, and nonspanning and nonheme *Rieske iron-sulfur protein* with a Fe_2S_2 cluster. Because of their locations at different sites on apocytochrome b molecule, the two heme b moleculés, differ in their electron-affinities and consequently function as cytochromes b_L (*E'o* = –0.03 V) and b_H (*E'o* = + 0.03V) respectively.

Complex III reoxidized QH_2, formed by the electron-transport by either complex I or complex II, to COQ by transferring its electrons through a postulated *Q cycle* (see 2.5.5) to ferricytochrome c_1 (Cye_1 Fe^{3+}) thereby reducing the latter ferrocytochrome c_1 (Cyc_1 Fe^{3+}).

Ferrocytochrome c_1 next gets reoxidized by donating its electron to ferricytochrome c (Cye Fe^{3+}) and reducing the latter to ferrocytochrome c (Cye Fe^{2+}). You should realize that *cytochrome c,* though not a component of any respiratory chain complex, serves as an electron-transferring link between complexes III and IV by receiving electrons from cytochrome c_1 or complex III and donating them to cytochrome a of complex IV.

d. Complex IV or cytochrome oxidase :

This large oligomeric transmembrane protein (MW 200 Kdal) carries two *heme* a molecules (heme a and a_3) bound to two different sites on its peptic subunit I, and two copper ions (Cu_a and Cu_{a3}) bound respectively to its subunits II and I. Because of the bindings of heme groups and copper ions to separate sites on the apoprotein, the heme a-Cu_a combination has a lower electron affinity (lower *E'o*) than that of the heme a_3-Cu_{a3} combination. The meme a-Cu_a combination forms the *cytochrome c-binding site* and acts as the *anaerobic oxidizing unit* (cytochrome a) of complex IV|——| it accepts electrons from ferrocytochrome c to reoxidize it to ferricytochrome c and subsequently donates them to the heme a_3——Cu_{a3} combination; the latter, having a higher *E'o* forms the *O_2-binding site* and acts as the *aerobic reduction unit* (cytochrome a) donating the electrons, received from heme a-Cu_a unit, to molecular O_2. Both the heme iron and the copper ion of each of the two units alternate between respectively ferric $\rightleftarrows$ ferrous and cupric $\rightleftarrows$ cuprous forms while transporting electrons. Complex IV thus functions as a *copper-containing aerobic oxidase or cytochrome a–a_3* complex (see 2.2.2).

2.4.2 Electron-flow paths

You will read below now the electrons from different metabolites, being oxidized, follow different routes along the mitochondrial ET chain to reach molecular O_2.

(*a*) Some metabolites such as pyruvate from glycolysis, α-ketoglutarate and isocitrate of the TCA cycle, and 3-hydroxy cacyl-CoA of beta-oxidation posses electron transfer potentials lower than that of NAD^+ of their respective dehydrogenases at their respective normal [Ox] / [Red] ratios in mitochondria. So, these substrates donate their electrons to NAD^+ producing NADH during their oxidations. Again, because of a higher *E'o* of the FMN of NADH-Q reductase than that of NADH, the latter can transfer the electrons so that FMN, thus reaching the electrons to complex I. Electrons from the $FMNH_2$ of complex I then flow through the successive components of

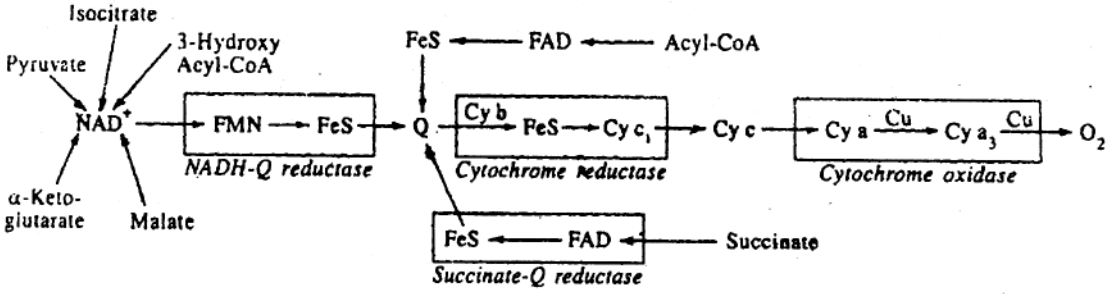


Figure 2.2 : Electron-flow paths through the respiratory chain.
(From D. Das, Biochemistry, Academic Publishers, 2002)

complex I to CoQ along their progressively rising *E'o* and *E* values, reducing CoQ to QH_2 (Fig. 2.2)

(*b*) Succinate from the TCA cycle is oxidized by succinate dehygenase which transfers its reducing equivalents to FAD. Similarly, acyl-CoA dehydrogenase oxidizes acyl-CoA in betaoxidation, using FAD as the electron-acceptor. From the $FADH_2$ thus produced, electrons can pass through iron-sulfur relatively higher *E'o* and *E* values of $FADH_2$ (see 2.4.1).

(*c*) Electrons from QH_2, whether coming from complex I or complex II, pass successively through the components of complex III, cytochrome C and complex IV. Because of their progressively rising *E'o* and *Eo* values, and finally reach molecular O_2.

2.4.3 Free energy changes during electron flow

For any oxidation-reaction, the *negative free-energy change* ($\Delta G^o'$) due to electron-flow from the reductand of one redox couple to the oxidant of another redox couple with a higher *E'o* is given by the following equation depending on the difference in redox potential ($\Delta E'o$) between the two redox couples, the calorie equivalent (*F*) of the Faraday constant (*F* = 23.063 Kcal), and the number (*n*) of electrons passing from one redox couple to the other.

$$\Delta G^{o'} = - nF\Delta E'o.$$

The $\Delta G^{o'}$ of the elections transfer between two successive redox couples at any step of mitochondrial electron transport must-exceed the $\Delta G^{o'}$ (– 8 Kdal) of hydrolysis of a high-energy ATP bond it any such bond has to be formed with the energy released at that step of mitochondrial electron transport. You will understand from the following computation that when two electrons (*n* = 2) flow over a $\Delta E'o$ + 0.22, $\Delta G^{o'}$ amounts to –10.15 Kcal and may be almost the minimum for generating an ATP bond ($\Delta G^o'$ –8 Kcal) :

$$\Delta G^{o'} = - nF\ \Delta E'o = - 2 \times 23.063 \times 0.22$$
$$= - 10.15 \text{ Kcal}$$

Negative free energy changes (– $\Delta G^o'$) along the rising $\Delta E'o$ at different sites of mitochondrial electron transport are brought about by the *exergonic transfer of electrons* between successive components of the respiratory chain complexes. If the negative $\Delta G^{o'}$ resulting at a particular site of such electron transfer is sufficiently higher in magnitude than the positive $\Delta G^{o'}$ of the endergonic phosphorylation of ADP to ATP the exergonic electron transport at that site may by coupled with the endergonic synthesis of ATP to drive the latter reaction.

(*a*) *Complex I* (NADH-Q reductase) transports two electrons over a $\Delta E'o$ of more than +0.36V from NADH ($E'o$ = 0.315 V) to CoQ ($E'o$ = + 0.045V). So, using the aformentioned equation, the $\Delta G^{o'}$ per electron-pair transported by complex I amounts to about – 16.61 Kcal/mol which suffices for the formation of one ATP bon (Table 2.1).

$$\Delta G^{o'} = nF\Delta E'o = -2 \times 23.063 \times 0.36 = -16{,}61 \text{ Kcal.}$$

(*b*) *Complex II* (succinate-Q reductase) transports two electrons over the $\Delta E'o$ of +0.015V from succinate ($E'o$ = +0.030V) to CoQ ($E'o$ = +0.045), thus generating the $\Delta G^{o'}$ of – 0.69 Kcal/mol. This poor $\Delta G^{o'}o$ does not suffice for forming any hing energy bond of ATP.

$$\Delta G^{o'} = nF\Delta E'o = -2 \times 23.063 \times 0.015 = -0.69 \text{ Kcal.}$$

(*c*) *Complex III* (QH_2 –cytochrome C reductase) transports two electrons over $\Delta E'o$ of + 0.19 V from QH_2 ($E'o$ = + 0.045V) to cytochrome C ($E'o$ = + 0.235 V), thus generating the $\Delta G^{o'}$ of – 8.76Kcal/mol. |——| this suffices for forming a high-energy ATP bond.

$$\Delta G^{o'} = nF\Delta E'o = -2 \times 23.063 \times 0.19 = -8.76 \text{ Kcal.}$$

(*d*) *Complex IV* (cytochrome oxidase) transports two electrons over the $\Delta E'o$ of + 0.58 IV from ferrocytochrome C ($E'o$ = + 0.235V) to molecular O_2 ($E'o$ = + 0.816 V), thus generating the $\Delta G^{o'}$ of – 26.80 Kcal |——| this suffer for forming another ATP bond.

$$\Delta G^{o'} = nF\Delta E'o = -2 \times 23.063 \times 0.581 = -26.80 \text{ Kcal.}$$

Table 2.1 ET Chain sites for forming high-energy bonds.

Complex	Electron transfer	$\Delta E'o$	$\Delta G^{o'}$	~P bonds formed
I	NADH → CoQ	+ 0.360 V	–16.61 Kcal	I
II	Succinat → CoQ	+ 0.015 V	–0.69 Kcal	0
III	QH_2 → Cyc Fe^{3+}	+ 0.190 V	–8.76 Kcal	I
IV	Cyc Fe^{2+} → O_2	+ 0.51 V	–26.80 Kcal	I

Thus, the $\Delta G^{o'}$ resulting from the transport of an electron-pair from pyruvate, α-ketoglutarate, malate, isocitrate or 3-hydroxyacyl-CoA to molecular O_2 by complexes I, III and IV successively, can be utilized to from one high-energy bond of ATP at each of those sites, making a total of *three high-energy bonds* during the aerobic oxidation of those substrates. But the transport of an electron-pair from succinate of acyl-CoA bypasses complex I and instead passes through complex II, III and IV,

resulting in a total yield of *two high energy bonds.*, because the $\Delta G^{o'}$ of electron-transport but complex II does not suffice in forming any ATP.

2.5 Oxidative phosphorylation

Phosphorylation of ADP to ATP involves the endergonic formation of a high-energy phosphorisc anhydride bond, having a high positive $\Delta G^{o'}$. This endergonic reaction has to be driven by being coupled with and exergonic reaction with a negative $\Delta G^{o'}$ of higher magnitude. This can be accomplished by either substrate-level or oxidative phosphorylations.

2.5.1 Substrate-level and and oxidative phosphorylations

These are two important processes of ATP formation by the phosphorylation of ADP.

(*a*) *Substrate-level phosphorylation :*

These can take place in both anaerobic and aerobic metabolisms. A high-energy metabolic intermediate is first formed as an intermediate in a metabolic pathway; e.g., 1, 3-bisphosphoglycerate and phosphoenolpyruvate, each carrying a high-energy phosphate bond, are formed in glycolysis while succinyl-CoA having a high-energy thioester bond is formed in the TCA cycle. Next, the exergonic cleavage of the high-energy bond in the metabolic intermediate, with a high negative $\Delta G^{o'}$, is coupled with the endergonic phosphorylation of ADP or GDP respectively to ATP and GTP directly utilizing the bond energy of the metabolic intermediate. Such phosphorylations of ADP or GDP do not involve aerobic respiration, use of molecular O_2, and mitochondrial electron transport.

(*b*) *Oxidative phosphorylation :*

These can take place during aerobic metabolisms only and involve the transport of electrons from substrates, being oxidized, to molecular O_2 by the components of the mitochondrial ET chain. Yo may recall that such mitochondrial electron transport is *exergonic* because it takes place along the progressively rising redox potentials of the successive components of respiratory chain complexes; at specific sites of the chain, the exergonic electron transport has negative $\Delta G^{o'}$ values of higher magnitudes than the positive $\Delta G^{o'}$ of the *endergonic* phophosylation of ATP to ADP. So, the exergonic mitochondrial electron transport can be used for driving the endergonic phosphorylation of ADP by the ATP *synthesis* of inner membrane particles. This is known as oxidative phosphorylation.

2.5.2 P : O ratio

It is ratio between the number of phosphate groups incorporated into ATP by esterification to form high-energy phosphate bonds, and the number of oxygen atoms reduced by electron-pairs transferred to them by mitochondrial electron transport to accomplish that esterification.

$$P : O = \frac{\text{number of phsohate groups esterified}}{\text{number of oxygen atoms reduced}}.$$

The P : O ratio serves as an index of oxidative phosphorylation, indicating the number of high-energy phosphate bonds formed by the transfer of electron-pairs from specific substrates to one oxygen atom. For example, oxidations of substrates such as pyruvate and malate by NAD^+_dependent dehydrogenases produce a P : O ratio of 3, because each electron-pair from NADH is transported of one oxygen atom successively by repiratory chain complexs I, III and IV, generating three ATP bonds |——| one at each of those three sites of the respiratory chain. On the contrary. FAD-dependent oxidations of substrates such as acyl-CoA and succinate produce a P : O ratio of 2, because each electron-pair from $FADH_2$ would reduce an oxygen atom by being transported only by complexes II, III and IV, by-passing complex I and thus generating only two ATP bonds |——| one at each of the sites in complexes III and IV. However, these prevalent P : O ratios are now thought actually to amount to 2.5 and 1.5 respectively because of factors such as mitochondrial proton leakages and presence of AMP in mitochondria in addition to ADP and ATP.

2.5.3 ATP synthesis

An inner membrane protein, called the *proton-translocating ATP synthesis,* catalyzes the endergonic phosphorylation of ADP to ATP in the mitochondrial matrix the negative $\Delta G^{o'}$ of the exergonic electron transport along the mitochondrial ET chain. This enzyme occurs in the headpiece of each of the numerous dumpbell-shaped *inner membrane particles* projecting into the mitochondrial matrix from the

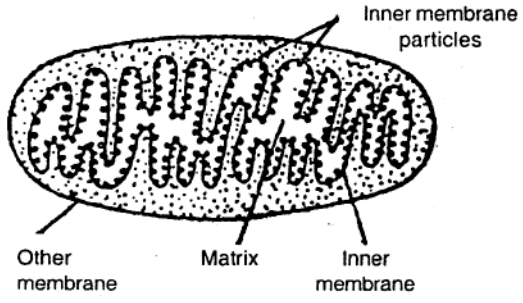


Figure 2.3 : Sectional view of a mitochondrion
(From D. Das Biochemistry, Academic Publishers, 2000)

matrix surface of the membrane and its cristase (Fig. 2.3). Each such particle (MW 450 Kdal) consists of the following three main parts.

(*i*) A cylindrical *basepiece,* 22 nm in diameter, is lodged in the inner membrane as nonpolar transmembrane oligomeric integral protein (*Fo unit*). A circular ring of a *DCCD-binding proteolipid* (MW 8 Kdal) eneireles a polar *proton channel* running across the membrane from its cytoplasmic surface to its matrix surface (Fix. 2.4).

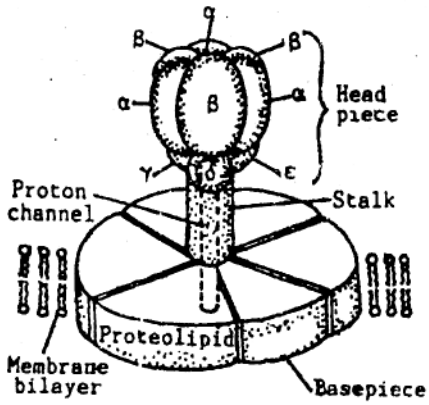


Figure 2.4 : Schematic diagram of an inner manbrane Particle
(From D. Das, Biochemistry, Academic Publishers, 2000)

(*ii*) *An oligomycin-sensitivity conferring protein* (MW 23 Kdal) and a *coupling* F_6 *protein* (8 Kdal) form a narrow 5-mm long *stalk* joining the F_O unit of basepiece to the head-piece of the particle.

(*iii*) A 370-KD polar, oligomeric, peripheral membrane protein (F_1 *unit*) constitutes the spheroidal *head piece* of the particle, and held by its stalk, projects into the mitochondrial matrix. The F_1 unit 10nm × 8nm in size, is made of three α, three β, and one each of γ, δ and ε peptide subunits. The α and β subunits are arranged alternately in a circle in the F_1 unit, in the form of three αβ pairs; the α and β subunits of each such pair partly differ from those of the other two pairs in their conformations as also in their special relations with the single γ subunit. The β subunit of each αβ pair bears a catalytic site of the *ATP synthase.* The δ subunit of F_1 is coupled to the matrix and of the proton channel running through the F_o unit and thus constitutes functionally a gate of that channel leading to F_1 unit.

So long as the transmembrane proton channel remains open in the inner membrane particle and an inward proton gradient is maintained across the inner membrane, the energy released by the exergonic down-gradient inward translocation of protons through the proton channel drives the endergonic phosphorylation of ADP to ATP by the ATP synthase of the F_1–F_o particles. But if (*i*) the proton channels are closed, eg., by the action of DCCD (*ii*) the trans-membrane proton gradient is abolished by the

action of proton-translocating ionophores such as dinitrophenol or (*iii*) the F_1 unit is isolated from the F_o unit by the action of urea, then the F_1 unit nolonger acts as the ATP synthesis and functions instead as *ATP* to hydrolyze ATP to ADP and P_1.

Boyer's binding-change model, in its present modified from, proposers that in every F_1 unit, the catalytic site on each β subunit acts cyclically due to conformation changes of the F_1 unit brought about by a transmembrane inward translocation of protons down a proton gradient resulting from electron transport. This model is briefly described below.

The β subunits of three αβ pairs of each F_1 unit have different conformations at any time : (*i*) the catalytic site on the β subunit of one αβ pairs exists as an *inactive O site* with an open conformation and possesses no affinity to bind to the substrates. ADP and Pi; (*ii*) the catalytic site on the β subunit of the second αβ a pair occurs as catalytically inactive *L site* binding only loosely to the substrates ; (*iii*) the catalytic site of the β subunit of the third αβ pair exists as the tightly binding and active *Tsite* (Fig. 2.5). The down-gradient proton inflow through the proton channel in the inner membrane particle produces the following changes cylically in the β subunited of each αβ pair of the F_1 unit.

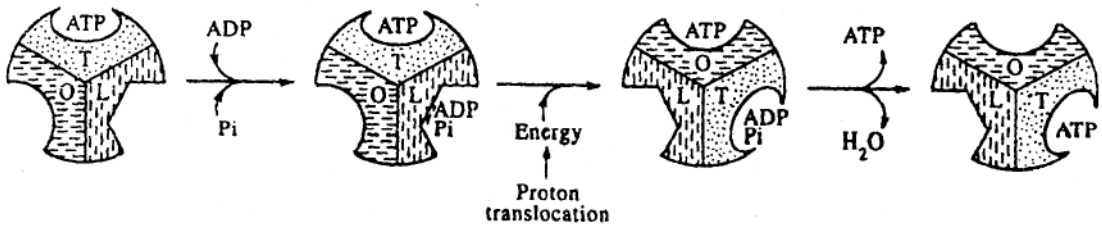


Figure 2.5 : Boyer's binding change model for ATP synthese action
(Atter R. L. Coss, Annu Rev Biochemistry, 50 : 687, 1981)

(*a*) The L site holds ADP and Pi initially loosely, the T site holds and ATP tightly, and the O site holds no substrate.

(*b*) The down-gradient proton-inflow through the proton channel of the F_1 unit releases energy which causes the rotation of its γ subunit; this changes (*i*) the conformation of the *T site to an O site,* releasing the ATP from it, (*ii*) that of the *L site to a T site* which now holds on the loosely-bound ADP and Pi moretightly, catalyzes the formation of ATP from them and then continues to hold that ATP tightly, and (iii) the conformation of the *pre-existing O site to a new L site* which binds loosely to ADP and Pi now.

These changes are cyclically repeated to make the β subunit of each αβ pair pass successively through O → L → T conformational and associated functional changes, bringing about the successive steps of ATP synthase activity.

2.5.4 Chemiosmotic theory

There is no direct physical association between the respiratory chain complexes of the inner membrane and the ATP synthase in the headpiece of the inner membrane particle; so, the energy released during mitochondrial electron transport has to be conserved by its conversion to a form that can be subsequently utilized by ATP synthease in ATP formation. This process is known as the *energy transduction or coupling between mitochondrial oxidation and ATP synthesis* Peter Mitchel proposed his *chemiosmotic theory* to explain this coupling.

According to this theory, the free energy released by electron transport along the mitochondrial ET chain is immediately harnessed in actively translocating protons from the mitochondrial matrix across the inner membrane to the outer compartment of mitochondrion. This active outward trans-membrane proton-translocation results in a higher proton (H^+) concentration outside that membrane. This in turn leads to (*i*) a *pH gradient* (ΔpH) with the outside pH 1.4 units below that in the mitochondrial matrix, and (*ii*) a *transmembrane potential* ($\Delta\Psi$) with the inner side of the membrane electron-negative to the extent of 0.14 volt relative to its outer side. The resultant of these two is an *electrochemical potential* (ΔP) with a consequent inward transmembrane proton (H^+) gradient. Where R and F are respectively the molar gas constant (8.314J/K/mol) and the Faraday constant (96-187C/mol. and T is the temperature of 298K.)

$$\Delta P = \Delta\Psi - 2.303\frac{RT}{F}\Delta pH = 0.14 - 2.303 \times \frac{8.314 \times 298}{96487} \times (-1.4)$$

$$= 0.14 - 0.059\ (-\ 1.4) = 0.224\ \text{V}.$$

As the protons, thus pumped out, flow inward down the consequent proton gradient through the proton channel of the inner membrane particle, the energy released by their down-gradient exergonic inflow is used by ATP synthase of F_1 unit in catalyzing the endergonic phosphorylation of ADP to ATP.

At least two protons must be actively translocated outward across the inner membrane to suffice for the formation of own ATP bond; Mitchell proposed in his *redox loop model* (See 2.5.5) that two protons can be so translocated during the transport of each electron-paid by each of the repiratory chain complexes I, III and IV. So, three ATP bonds should be expected to be formed in consequence of the transport of an electron-paid from NADH to molecular O_2 through complexes I, III and IV. However, *in vivo* nearly three protons may have to be translocated to provide for each ATP bond, because some translocated protons may diffuse back across the inner membrane, lowering the electrochemical gradient ΔP below what is theoretically anticipated.

2.5.5 Redox loop mechanism

According to this model proposed by Mitchell, the mitochondrial ET chain is organized into *three redox loops* (Fig. 2.6) Each loop includes a respiratory chain complex and transports electrons first from the inner (matrix) surface to the outer (cytoplasmic) surface of inner membrane and then, back again to the inner surface. During such trnasport of each electron-pair, each loop simultaneously translocatres two protons outward and releases them into the cytoplasm from the outer surface of the membrane. For functioning in this manner, the respiratory chain complex are oriented a symmetrically in the inner membrane.

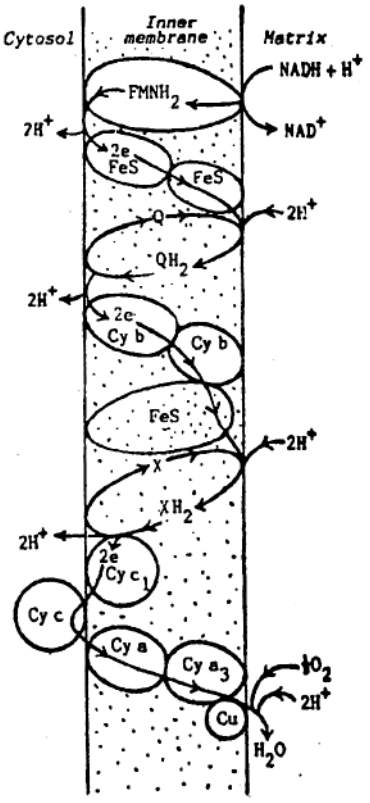


Figure 2.6 : Mitechell's redox loop model
(From D. Das, Biochemistry, Academic Publishers, 2000)

(*a*) *First redox loop :*

This is constituted by an integral membrane protein which spans the entire thickness of the membrane and functions as *NADH-Q reductase* or respiratory chain *complex 1.* Its. FMN prosthetic group receives a pair of electrons along with protons from NADH of the matrix and gets reduced to $FMNH_2$ The latter donates the lecrons

to the iron-sulfur clusters of complex I near the outer surface of the membrane and simultaneously release two protons from that surface to the cytoplasm. The iron-sulfur clusters then transport the electrons back to the inner surface of the membrane and transfers them to CoQ there; the latter is reduced to QH_2 (ubiquinol) by taking up two protons simultaneously from the matrix.

(*b*) *Second redox loop :*

This consists of another transmembrane complex which functions as *complex II* or *cytochrome c reductase.* QH_2 (see above) diffuses through the membrane lipid bilayer to its outer surface where it gets reoxidized to CoQ by donating its electrons to cytochrome b_L (b_{560}) of complex III and releasing its two protons into the cytoplam. These electrons are then transferred to cytochrome b_H(b_{562}) and iron-sulfur cluster of complex III. It was initially supposed that a hypothetical mobile electron-carrier (X) would receive the electrons from the iron-sulfur cluster of complex III near the inner surface of the membrane and two protons from the adjacent matrix to become reduced to XH_2 which would then carry the electrons to the next redox loop by diffusing through the membrane lipid bilayer.

(*c*) *Third redox loop :*

This loop, Mitchell proposed, consists of *cytochrome C* and *complex IV* or *cytochrome oxidase.* The reduced hypothetical carrier (XH_2) was supposed to diffuse across the inner membrane to its outer surface where it would get reoxidized by donating its electrons to cytochrome c_1 and simultaneously releasing its protons into the adjacent cytoplasm. Electrons received by cytochrome c_1 would then be carried successively by cytochromes c, a and a_3 to the inner surface of the membrane where the electrons would finally be transferred to molecular O_2 in the matrix.

Mitchell tried to explain in this way how three pairs of protons were translocated from the matrix to the cytoplasm during the transport of one electron-pair successively by complexes I. III and IV. But the hypothical electron-carrier (X) supposes to link the second and third loops could never be found.

Q cycle :

This cycle of events has been proposed to avoid assuming a hypothetical electron-carrier (X) for linking complexes III and IV, and to suggest as alternative way of proton translocation other than the direct involvement of complex IV as a third redox loop. The Q cycle proposes the translocation of *four protons* by the second redox loop itself during its role in electron-transfer, evading the need to assumed the translocation of two protons by each of second and third loops. Its proposed steps are summarized below (Fig. 2.7).

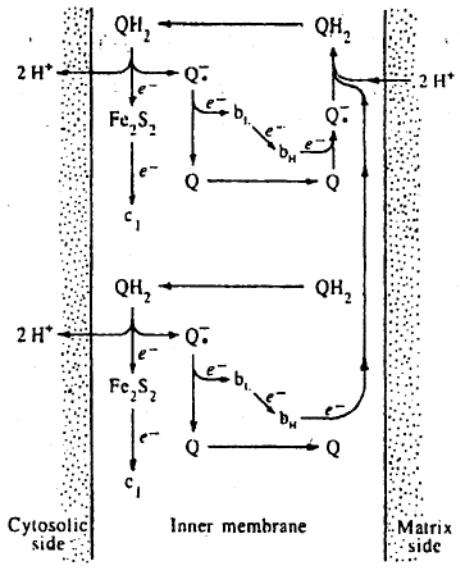


Figure 2.7 : The Q cycle (From D. Das, Biochemistry, Academic Publishers, 2005)

(*i*) QH_2, formed by the transfer of electrons from complex I of the first redox loop, diffuses from the inner (matrix) side to the outer (cytosolic) side of the inner membrane where it donates one electron to the Rieske iron-sulfur protein of complex III, releases two protons in the cytosol and forms a semiquinone anion (Q^-). The latter is reoxidized to CoQ by donating its other electron to ferricytochrome b_L, reducing it to ferrocytochrome b_L.

(*ii*) The Rieske iron-sulfur protein gets reoxidized by giving the received electron to ferricytochrome c_1 of complex III, to reduce it to ferrocytochrome c_1.

(*iii*) Two ferrocytochrome b_L molecules are reoxidized by donating their respective electrons to two cytochrome b_H (b_{562}) molecles reducing them to ferrocytochrome b_H.

(*iv*) Two ferrocytochrome b_H molecules get reoxidized by donating their electrons to a CoQ molecule near the inner (matrix) surface of the membrane; the CoQ molecule simltaneously accepts two protons from the matrix to be reduced to QH_2. The latter diffuses back to the outer (cytosolic) surface of the membrane to restart the cycle from step (i).

Thus each Q cycle translocates a total of four protons outward across the inner membrane during the transport of each electron-pair by the complex III itself.

2.5.6 Proton pump mechanism

This is an alternative model proposed to explain the outward proton translocation across the inner membrane during mitochondrial electron transport. According to this

model, some proton-translocating proteins occur in the inner membrane with two separate conformations in their oxidized and reduced forms, repectively. The oxidized form of such a proton-translocating protein has specific amino acid side-chains of its proton-binding site exposed on the inner (matrix) surface of the inner membrane; in this conformation, those amino acid side chains behave as weak acidic groups with high acid ionization exponent (pk) values and consequently possess high proton-affinities. So, protons form the matrix bind easily to such weak by acidic groups of the proton-binding site (Fig. 2.8). *Cytochrome oxidase* of the ET chain may act as such a proton-translocating protein, but there may also be other such proteins that are not direct components of the ET chain.

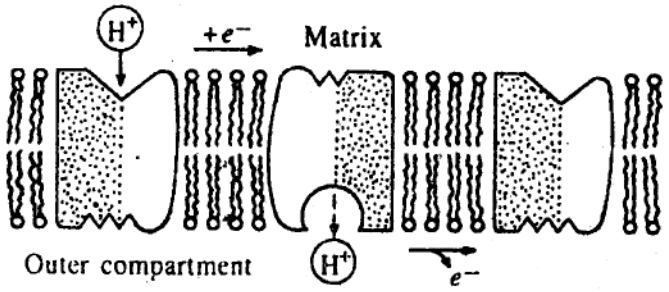


Figure 2.8 : Changes in a portion - translocator during a proton pump cycle
(From D. Das, Biochemistry, Academic Publishers, 2005)

As the protonated proton-translocator gets reduced by accepting electrons from a mitochondrial respiratory chain complex, it undergoes changes in conformation and proton-affinity. As a result, the amino acid side chains at the proton-binding site of the reduced and protonated translocator get exposed now on the outer (cytoplasmic) surface of the membrane, behave as stronger acid groups with low pK values and weaker proton-affinites, and consequently release their protons into the cytosol. The subsequent reoxidation of the translocator by the transfer of its electrons to the next-electron-acceptor changes its conformation to its original deprotonated form so that its proton-acceptor site gets exposed again on the inner surface of the membrane, regains its original weak acid groups and higher proton-affinity (higher pK) and awaits fresh protonation to star the next cycle.

2.5.7 Regulation of oxidative phosphosrylation

The rate of oxidative phosphorylation depends on the mitochondrial ATP concentration. In other words, the ATP *mass action ratio* in the mitochondrical matrix determines the rate of oxidative phosphorylation—this ratio is given by [ATP] / [ADP][Pi] which is the ratio of the molar concentration of ATP and the product of molar concentrations of ADP and Pi in the matrix. A fall in this ratio or a rise in the

ratio of mitochondrial concentrations of the reductant and oxidant forms of the NAD^+/ NADH redox couple, viz., [NADH]/ [NAD^+], enhances the mitochondrial concentration of ferrocytochrome c ($CycFe^{2+}$); the latter being the substrate of *cytochrome oxidase,* the activity of the enzyme is consequently enhanced, leading to a rise in the rate of mitochondrial oxidation as also of the oxidative phosphorylation coupled with it. Stated in a different way, mitochondrial oxidation and oxidative phosphorylation are controlled largely by the mitochondrial ADP concentration (*respiratory or acceptor control*).

To maintain the rate of oxidative phosphorylation in an active tissue, the ATP mass action ratio is normally kept adequately low in the mitochondrion by the action of a homodimeric intergral protein. *ATP-ADP translocator,* of the inner mitochondrial membrane. This 60-KD protein carries out and *ATP-ADP antiport* across the inner membrane. Thus, the ATP-ADP translocator, driven by the membrane potential, transports ATP molecules outward and ADP molecules inward across the inner membrane. This normally keeps the matrix [ADP]/[ATP] ratio at about tenfold of the cytoplasmic [ADP]/[ATP] ratio. In this way, the ATP-ADP translocator functions as a *rate-limiting factor* in maintaining an adequately low ATP mass action ratio in the mitochondrion. Its inhibition by the mould anitibiotic bongkrekic acid or the glycoside atractyloside decreases the transmembrane ATP-ADP transport, there raises the mitochondrial ATP mass action ratio, and consequently brings about the inhibition of mitochondrial oxidation as also of oxidative phosphorylation. You may read further details of this translocator action in 3.4.5.

In a resting tissue, relatively less ATP needs to be to be hydrolyzed for energy production and the ATP mass action ratio remains high. This keeps the ferrocytochrome c concentrations low in the mitochondrion with a consequent decline in cytochrome oxidase activity. So rates of mitochondrial electron transport and ATP formation remain low. On the contrary, in a highly active tissue, ATP is hydrolyzed in high amounts to ADP and Pi for meeting the enhanced energy requirement, and ATP-ADP antiport is heigthened. The mitochondrial ATP mass action ratio consequently falls and ferrocytochrome c concentration rises. This leads to a rise in cytochrome oxidase activity, with resultant rises in mitochondrial oxidation and oxidative phosphorylation.

Uncoupling of oxidation and phosphorylation :

Some lipid-soluble substance such as penta-chlorophenol and dinitrophenol possess strong proton-affinities because of their weakly acid nature. So, they can bind to protons on the outer side of the inner membrane to change into neutral protonated forms which diffuse across the membrane and release the protons into the mitochondral

matrix. Thus, acting as *monile, protontranslocating ionophores,* they lower or abolish the transmembrane proton gradient, resulting in (i) a decline in proton inflow through proton channels of F_1-F_0-particles and a consequent lowering of the ATP synthase action of F_1 unit, and (*ii*) an enhanced ATP ase action of F_1 unit because of the rise of proton concentration in the matrix. The resultants of these two effects consist of an enhanced mitochondrial oxidation due to the rise in matrix ADP concentration, and decline in oxidative phosphorylation. So, the energy generated by enhanced mitochondrial oxidation is given of as heat. In this way, these lipid-soluble proton-translocators bring about an *uncoupling* of oxidation and phosphorylation. The *thermogenic* (temperatue-raising) *affect* of thyroid hormones may result partly from the uncoupling of mitochondrial oxidation and oxidative phosphorylation.

In hibernating mammals and new-born furless mammals, adrenaline secreted on cold exposure brings about an uncoupling of mitochondrial oxidation and oxidative phosphorylation. The secreted adrenaline activatives the lipolysis of brown fat in adipocytes into fatty acids which open up inner membrane proton channels made of an integral protein called *thermogenin.* This causes inward proton translocation across that membrane through the thermogenin channels and nullifies the transmembrane proton gradient, uncoupling oxidation and phosphorylation. Thus the energy from mitochondrial oxidation of fatty acids from brown fat is dissipated as heat instead of being utilized for ATP synthesis.

2.6 Summary

Biological oxidations and reductions are mostly catalyzed by oxidoreductases of four main classes, viz, oxygenases which incorporate oxygen into their substance, oxidases which oxidize their substrates by transferring electrons from the latter to molecular O_2 directly, hydroperoxidases that transfer electrons from substrates of H_2O_2, and dehydrogenases which transfer electrons from their substrates to electron-acceptors other than molecular O_2. Of the dehydrogenases, pyridine-linked dehydrogenases use either NAD^+ or $NADP^+$ as electron-acceptor coenzymes, flavin-linked dehydrogenases posses either FMN or FAD as the prosthetic group for accepting electrons from the substrates, iron-sulfur proteins have nonheme iron-sulfur clusters which accept electrons from the substrates, while cytochromes are hemoproteins with their heme-iron functioning as electron-acceptors.

Every substance exists in an oxidant form which may accept electrond from other substance, oxidizing the latter and itself being reduced to a reductant form; the latter, in turn, may donate these electrons to other electron-acceptor substances, reducing the latter and itself being reoxidized to the oxidant form. Electrons can flow

in this way from the reductant member of a substance having lower electron-affinity to the oxidant member of another substance having higher electron-affinity. Redox potential is a measure of the electron-affinity of a substance. So, it is an estimate of the tendencies for accepting and donating electrons by a substance. Electrons would flow from the reductant form of a substance having a lower or more negative redox potential, i.e., a lower electron-affinity, to the oxidant form of another substance having a higher or more positive redox potential, i.e., a higher electron-affinity. Besides, the direction of electron flow between two substances also depends on the ratios molar concentrations, of their oxidant and reductant forms in a system.

A number of redox (oxidant-reductant) couples, mostly integral proteins of the inner mitochondrial membrane, constitute a mitochondrial electron-transport or respiratory chain, transporting electrons from various substrate to molecular O_2 ultimately, in accordance to the order of their progressively ascending redox potentials. Each such ET chain consists of four respiratory chain complexes, viz., complex I transporting electrons from NADH to coenzyme Q, a mobile nonprotein nonpolar ubi-quinone of the inner membrane, complex II transporting electrons from succinate of CoQ, comprlex III carrying electrons from CoQ. H_2 to a peripheral protein cytochrome c, and complex IV receiving electrons from reduced cytochrome C and donating them to molecular O_2. Except cytochrome oxidase, an oxidase of complex IV, all other integral proteins of the ET chain function as dehydrogenases. The exergonic transport of each electron-pair along the rising redox potentials of complexes I, III and IV release sufficient free energy for forming three high-energy ATP bonds.

Oxidative phosphorylation is the endergonic phosphorylation of ADP to ATP, harnessing the energy released by the exergonic electronic-transport along the respiratory chain. This phosphorylation is catalyzed ATP synthase activity of the F_1 protein of the headpiece of inner membrane particles. According to Peter Mitchell, for this oxidative phosphorylation the ATP synthase activity is coupled with the mitochondrial electron-transpost by a transmembrane proton gradient created across the inner membrane by the outward transfer of protons through the latter during the election transport. Mitchell proposed a redox loop mechanism, and subsequently the Q cycle in modifying that mechanism, to explain how mitochondrial electron transport may lead to the inward transmembrane proton gradient. An alternative proton pump mechanism has also been proposed to explain the coupling of mitochondrial oxidation and ATP synthase action, assuming changes in the proton-affinities and conformations of some proton-translocator proteins of the inner membrane during mitochondrial electron transport. In his binding-change model of ATP synthase action. Boyer has proposed cyclic changes in that action of the beta-subunits of each of three $\alpha\beta$-subunit-pairs of the F_1 protein during electron transport.

Oxidative phosphorlation is mainly regulated by the ATP mass action in the mitochondrial matrixm, viz., [ATP]/[ADP][Pi], which is normally kept adequately low in the matrix by a transmembrane ATP-ADP anitport, carried out by an inner-membrane ATP-ADP translocator. Uncoupling of mitochondrial oxidation and oxidation and oxidative phosphorylation enhances mitochondrial. Oxidation and decrease oxidative phosphorylation, leading to the release of the energy from aerobic oxidation as heat.

2.7 Terminal questions

1. (a) Discuss the differences between substrate-level phosphorylations and oxidative phosphorylations, citing examples.
   (b) Describe the components of the inner membrane particle and its F_1 unit housing the ATP synthaes, using a suitable diagram.
   (c) Discuss with a suitable diagram Boyer's binding-change model for the mitochondrial ATP synthase action.
2. (a) Describe the chemiosmotic theroy about the coupling of mitochondrial oxidation and ATP synthesis.
   (b) Describe the proton-pump mechanism for proton translocation across the inner membrane during mitochondrial electron transport.
   (c) Discuss the roles of ATP mass action ratio and ATP-ADP translocator in regulating oxidative phosphorylation.
3. (a) Give an outline of electron transports by pyridine-linked dehydrogenases, flavin-linked dehydrogenases and iron-sulfur proteins.
   (b) Give brief accounts of oxygenases and oxidases, with examples of the subclasses of both.
   (c) Discuss how to mitochondrial cytochromes belong to two different classes of oxidoreductases, giving examples.
4. (a) Explain what you understand by redox potential and standard redox potentials. Discuss the significance of redox potential in the electron flow between electron-donors and electron-acceptors.
   (b) Describe the relation between electron transfer potential and redox potential, quoting the relevant Nernst equation.
   (c) Explain how electrons flow from NADH to pyruvate during anaerobic glycolysis, but from lactate to NAD^+ in the post-glycolysis aerobic condition.

5. (a) Describe the components and functions of the repiratory chain comlexes constituing each mitochondrial respiratory chain assembly.
   (b) Describe using a suitable flowcahrt the different paths followed by electrons along the respiratory chain complexes during their flow from different metabolites to molecular O_2.
   (c) Discuss, quoting the relevant equation, the free energy changes during electron flow along different repiratory chain complexes, and identify there from the sites of the repiratory chain where high-energy bonds may be formed.
6. (a) Described the oxidation-phosphorylation coupling using Mitchell's chemiosmotic theory.
   (b) With a suitabl_ .iagram, describe the redox loop mechanism of electron transport to explain proton translocation during mitochondrial repiration.
7. (a) Describe the Q cycle with a sitable flowchart.
   (b) Discuss the regulation of mitochondrial oxidative phosphorylation.
   (c) How can mitochondrial respiration and phosphorylation be uncoupled? Describe such uncoupling in hibernating mammals.
8. Write notes on the followin :
   (a) Redox couples.
   (b) Hydroperoxidases.
   (c) ATP-ADP translocator.
   (d) ATP-ADP mass action ratio.
   (e) Proton-translocating ATP synthase.
   (f) P : O ratio

2.8 Answers

1. (a) See Section 2.5.1.
   (b) See Section 2.5.3.
   (c) See Section 2.5.3.
2. (a) See Section 2.5.4.
   (b) See Section 2.5.6.
   (c) See Section 2.5.7.
3. (a) See paragraphs (a), (b) and (c) Section 2.2.4.

(b) See Sections 2.2.1 and 2.2.2.

(c) See paragraph (d) of Section 2.2.4.

4. (a) See first to fifth paragraphs Section 2.3.

(b) See sixth paragraph of Section 2.3.

(c) See seventh (last) paragraph of Section 2.3.

5. (a) See Section 2.4.1.

(b) See Section 2.4.2.

(c) See Section 2.4.3.

6. (a) See Section 2.5.4.

(b) See Section 2.5.5.

7. (a) See Section 2.5.5.

(b) See Section 2.5.7.

(c) See fourth and fifth paragraph of Section 2.5.7.

8. (a) See first, fourth, fifth and sixth paragraphs Section 2.3.

(b) See Section 2.2.3.

(c) See second and third paragraphs of Section 2.5.7.

(d) See second paragraph of Section 2.5.7.

(e) See Section 2.5.3.

(f) See Section 2.5.2.

Unit 3 ❐ Membrane Transport, Biosyntheses and Muscle Contraction

Structure

3.1 Introduction

You will learn in this unit about the characteristics, kinetics and significances of active transports across membranes and some mechanisms of such transports.

You will then proceed to read about the biosynthesis of several substances of biological importance such as norepinephine, epinephrine, glutathione, serotonin, melatonin, glycogen, saturated fatty acids, prostaglandins and urea.

Finally, major muscle proteins will be briefly descibed, sliding filament model of muscle contraction and the molecular mechanism of muscle contraction will be presented to you.

Objectives

- After reading this unit, you should be able to :
- Differentiate between nonmediated and carrier-mediated transports,
- Distinguish between different stoichiometric categories of active transports,
- Describe the characteristics, significance and kinetics of actine transports,
- Discuss the mecahnism of ATP-driven active transports such as the sodium-potassium pump.
- Understand and describe the mechanisms of ion gradient-driven active transports such as the ATP-ADP antiport,
- Describe the pathway, energetics and significance of urea biosynthesis and explain its significance,
- Give an account of how glucose is converted to glycogen for storage in tissues,
- Discuss how glucose can be formed from non-carbohydrates,
- Describe how amino acids are used in synthesizing specialized products such as glutathione, norepinephrine, epinephrine, serotonin and melatonin,
- Give an account of synthesis of the saturated fatty acid, palmitic acid, and of its elongation into longer saturated fatty acids,
- Explain how polyunsaturated fatty acids can be used in synthesizing prostaglandins by the cyclo-oxygenase pathway.
- Develop a basic idea about the structural and molecular characteristics of major muscle proteins,
- Understand the concept of the sliding filament model of muscle contraction,
- Understand and explain the molecular mechanism of muscle contraction, based on actine-myosin interaction in striated muscles,
- Describe the myosin-based contraction of smooth muscles.

3.2 Membrane Transport

There are two major processes for the transport or flow of solutes across in **intact membrane.**

(a) Nonmediated transport or diffusion :

It is a simple physical process in which the substrate flows down its electrochemicl gradient across the membrane, requiring neither any energy expenditure nor any transporter or carrier protein to ferry it from one side to the other, continuing only

so long as the transmembrane gradient exists, and getting affected by the Gibbs-Donnan effects of nondiffisuble ions.

(b) Carrier-mediated transports :

They are far speedier processes with specific transporters, translocases or carrier proteins transferring specific molecules or ions across the membrane. Such mediated transports may again be of two main types, according as cellular work is or is not involved in the process.

(i) Passive mediated transport or facilitated diffusion :

Here, specific carriers transfer specific substrates for more rapidly than diffusion and down their respective concentration or electrical gradients across the membrane so long as such a gradient exists, but requiring no energy expenditure and not being affected by cold, lack of O_2 or metabolic inhibitors; e.g., transports of fructose and mannose across the intestinal cell membrane.

(iii) Active transport : This type of transport is even more rapid and is carried out by carrier proteins having high specificities for their substrates; but active transport transfers substrate molecules against their electrochemical gradients, consequently involves the expenditure of cellular energy and so, gets depressed by cold, metabolic inhibitors and hypoxia; e.g., transmembrane transports of Na^+ and K^+ and Na^+K^+ pump.

3.3 Active Transport

Active transports are *endergonic, carrier-mediated transports* of specific solute molecules or ions by specific carriers in a specific direction across a membrane and even against the transmembrane electrochemical gradient of the transported substrate. Like facilitated diffusions, active transports also belong to three categories *stoichiometrically;* viz., *(i) uniport* or the transmembrane transport of a single substrage, e.g., extrusion of cytosolic Ca^{2+} to the extracellular fluid; *(ii) symport or* the simultaneous transport of two different substrates in the same direction across a membrane, e.g., simultaneous active transport of boty N^{a+} and glucose from the intestinal lumen to the enterocyte across its membrane; (iii) *antiport* or *counterport* which is the simultaneous transport of two different substrates in opposite directions across a membrane, such as the active Na^+ efflux and K^+ influx by $Na^+ - K^+$ pump across the plasma membrane.

3.3.1 Characteristics of active transports :

(a) Active transprot is *far more rapid* than both nonmediated transport and passive mediated transport.

(b) The carrier or transporter for any active transport possesses very high *substrate-specificity,* transferring only a specific substrate or a specific pair of substrates across the membrane.

(c) Active transport is *unidirectional*–unlike diffusion and facilitated diffusion which can take place in either direction across a membrane depending on the direction of the electrochemical gradient of the substrate, active transport of a substrate always takes place in a single specific direction across the membrane. This can be explained by the *asymmetric disposition of the transporter* for any active transport in the membrane. When not bound to its substrate, this asymmetric disposition keeps its substrate-binding site on a particular surface of the membrane and binding to the substrate only on that surface; this changes the conformation of the carrier either to place the carrier-bound substate on the other surface of the membrane or to constitute a polar channel across the membrane to its other surface. This enables the carrier to release the substrate from that surface. The carrier does not itself diffuse through the membrane during the process and regains its original conformation after the substrate has been released from it.

(d) The darrier always transports the substrate in a *single specific direction* across the membrane, *even against the electrochemical gradient* of the substrate. This enables the active transport to absorb the entire amount of a substrate from the lumen of intestine or renal tubule. It also produces a steep electrochemical gradient of the substrate across a membrane.

(e) Active ion transport is *electrogenic,* producing a potential difference between the two sides of the membrane when it segregates positive and negative ions on its opposite sides. But if it is either a symport of two counterions or an antiport of two like-chanrged ions, active ion transport is *electroneutrall* with no trans-membrane potential difference resulting from it.

(f) As active transport is carried out against the electrochemical gradient of the substrate, it is an *endergonic process* with positive free-energy change ($\Delta G^{o'}$). Where R is the molar gas constant (8.314J/K/mol) and T is the absolute temperature (298K ≡ 25°C), the ΔG^{01} for the active transport of n moles of solute from its lower molar concentration (C_1) to its higher molar concentration (C_2) across a membrane, and the work done (Wcal) by the cell in that transport are given by :

$$\Delta G^{o'} = RT\ l\ n\ \frac{C_2}{C_1} = 2.303\ RT\ \log\ \frac{C_2}{C_1}\ ;$$

$$W = 2.303\ nRT\ \log\ \frac{C_2}{C_1}.$$

(g) Because active transports are engergonic, they have to be driven by being *coupled with exergonic reactions* with high negative $\Delta G^{0'}$ such as the hydrolysis of high-energy bonds of ATP (*ATP-driven active transports*) and the electrochemical potential gradients created by transmembrane ion gradients (*ion gradient-driven active transports).* [See 3.3.4 and 3.3.5 below]

(h) Because transportable substrate molecules for outnumber the limited number of specific carrier molecules, progressive rise in substrate concentration [S] leady to a gradual decline in the rate of increase in the transmembrane flow or *flux (J)* of the substrate and ultimately reaches a maximum flux (*J*max) with a saturating substrate concentration—the *J*max cannot be exceeded by raising the [S] further. This is reflected in the *rectangular hyperbolic graph* obtained by plotting *J* values against the respective [S] values (Fig. 3.1). Thus the flux of active transport is a *rectangular hyperbolic*

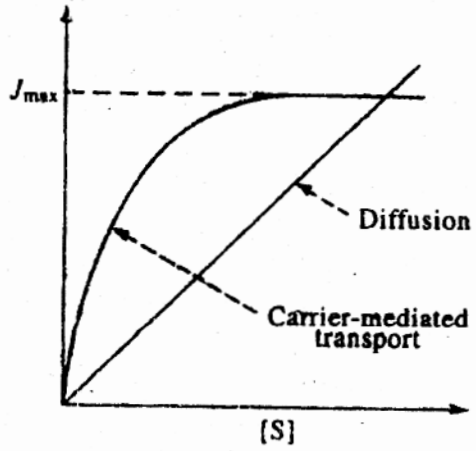


Figure 3.1 : Hyperbolic Kinetics of active transports, in contrast to the kinetics of diffusion. [From D. Das, *Biochemistry*, Academic Publishers, 2005]

function of the molar concentration of substrate, and follows the equation for the *Michaelis-Menten hyperbolic Kinetics;* where K_m is the molar concentration of sybstrate for attaining $^1/_2$ *J*max, $J = \frac{J\,\max[S]}{K_m + [S]}$

(i) Some cases of simultaneous active transports of more than one substrate molecule, however, follow the *sigmoid saturation Kinetics* owing to a *positive cooperatively* between the substrate molecules.

(j) Some active transports are *competitively inhibited* by suitable substrate analogues competing with the substrate for occupying the substrate-binding sites of the carrier. This indicates the existence of only *a limited number of sites* available in the carrier for such mediated transports.

3.3.2 Significances of active transports

By transporting specific solute molecules or ions in a particular direction across a membrane, active transports help to accomplish the following :

(i) Absorb or reabsorb the specific solutes from intestinal or renal tubular lumens very rapidly, independent of their transmembrane electrochemical gradients and consequently, even totally from those lumens,

(ii) Create and maintain ion gradients in specific directions across a membrane.

(iii) Keep the resting membrane polarized, a repolarize it after its depolarization,

(iv) Create the resting membrane potentials of neuromuscular tissues and also participate in the course of action potentials during their excitation and conduction of impulses, and

(v) Help in the distribution of water in different body compartments due to the obligatory osmotic transfer of water across the membranes along with the actively transported solutes.

3.3.3 Mechanisms of active transports

It may be recalled that being *endergonic* in nature, active transports have to be conducted by being *coupled with exergonic reactions* which in most cases consist of either ATPase catalyzed hydrolysis of high energy phosphate bonds of ATP, or dissipation of ion-gradients across the membrane to be crossed by the transported substrates. These two mechanisms are described below with examples. (See 3.3.4 and 3.3.5.)

3.3.4 ATP-driven active transports

In these cases, endergonic active transports are conducted by being coupled with the exergonic hydrolysis of ATP so that the free energy released by the latter chemical reaction can be harnessed for the endergonic physical process of active transmembrane transport. Four types of ATPases, catalyzing ATP hydrolysis for active ion-transports, occur in animal membranes, viz., *P-type ATPases* of plasma membrane and endoplasmic reticulum membrane, *V-type ATPases* of lysosomal membrane, *F-type ATPases* of inner mitochondrial membrane particles—all these three types are associated with active transports of specific cations such as Na^+, K^+, Ca^{2+}, and H^+, and A-type ATPases for active transports of anions such as Cl^- ions. Two such ATP-driven active transports are described below.

(a) $Na^+ - K^+$ pump :

This integral protein of plasma membrane is also known as *$Na^+ - K^+$ ATPase* and carries out the active antiport of three Na^+ ions outward and two K^+ ions inward across the membrane against their respective concentration gradients, hydrolyzing

one ATP molecule into ADP and Pi : $3Na^+_{in} + 2K^+_{out} + ATP + H_2O \rightarrow 3N_a{}^+_{out} + 2K^+_{in} + ADP + Pi$. This antiport helps to maintain the extracellular and intracellular Na^+ concentrations respectively at about 150 mM and 15 nm, and the respective K^+ concentrations at about 5mM and 100mM. This ATPase, a tetrapeptide consisting of two nonglycosylated α and two glycosylated β peptidesubunits, is a *P-type ATPase* —like other P-type ATPases, this enzyme also occurs in the plasma membrane, is inhibited by vanadate, and gets phosphorylated by ATP during its action. The enzyme exists in two conformation,s viz., E_1 and E_2 (see below).

The $Na^+ - K^+$ ATPase is activated **in vivo** by any rise in intracellular Na^+ concentration or in extracellular K^+ concentration. Then, it carries out the $Na^+ - K^+$ antiport across the plasma membrane in the following way to maintain the normal inward concentration gradient of Na^+ and the normal outward gradient of K^+ (Fig. 3.2).

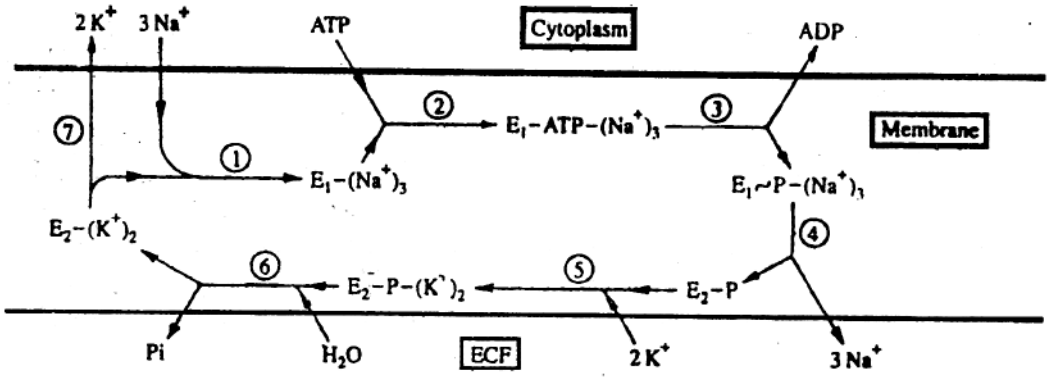


Figure 3.2 : Active antiport of Na^+ and K^+ by Na^+–K^+ [From D. Das, *Biochemistry*, Academic Publishers, 2005]

(i) The cation-binding site of the E_1 form of $Na^+ - K^+$ ATPase has a high Na^+ affinity, is situated on the cytoplasmic surface of the membrane, and binds to three Na^+ ions from the cytosol, forming $E_1 - (Na^+)_3$ complex; the latter then binds to a cytosolic ATP molecule to form an $E_1 - ATP\text{-}(Na^+)_3$ complex.

(ii) The $E_1 - ATP\ (Na^+)_3$ complex changes into $E_1 \sim P - (Na^+)_3$, a high-energy intermediate, by using the bound ATP to phosphorylate a specific aspartate residue of the enzyme into a phospho-aspartate residue, and releases ADP into the cytosol.

(iii) The high-energy $E_1 \sim P - (Na^+)_3$, intermediate next changes into a low-energy $E_2–P - (Na^+)_3$, form; this change in the conformation of the enzyme translocates its cation-binding site to the extracellular surface, lowers the Na^+ – affinity of that site, consequently release the Na^+ ions into the extracellular fluid, and changes into $E_2 - P$ intermediate with high K^+ – affinity.

(iv) The cation-binding site of the E_2 – P intermediate now binds to two extracellular K^+ ions to form an E_2–P–$(K^+)_2$ intermediate whose phosphoaspartate residue releases P_i into the ECF; this gives rise to an E_2–$(K^+)_2$ intermediate.

(v) This changes the conformation of the latter into E_1 – $(K^+)_2$ with its cation-binding site translocated on the cytosolic surface, lowering its K^+ – affinity; the two K^+ ions are consequently released in the cytosol and the enzyme resumes its original E_1 conformation.

Functions :

By maintaining transmembrane inward and outward concentratiion gradients of Na^+ and K^+ respectively, the Na^+ – K^+ ATPase principally helps in **(i)** active intestinal and renal tubular absorptions of Na^+ ions from their lumens, **(ii)** maintenance of the resting polarized state and the post-excitation repolarization of neuromuscular membranes, and **(ii)** maintenance of intracellular osmolarity and fluid content.

(b) Ca^{2+} – ATPases :

These are also *p-type ATPases* occurring as integral proteins in plasma membrane, endoplasmic reticulum membrans and muscles sarcoplasmic reticulum menbranes. Acting in a manner similar to the Na^+ – K^+ ATPase, Ca^{2+} – ATPases carry out the extrusion of each pair of Ca^{2+} ions from the cytosol to either the ECF or the ER/SR lumens, against the electrochemieal gradient of Ca^{2+}, utilizing the energy from one ATP bond hydrolyzed in the process. $2Ca^{2+}{}_{in} + ATP + H_2O \rightarrow 2Ca^{2+}{}_{out} + ADP + Pi$. Thus, Ca^{2+} – ATPases help to maintain **(i)** a much higher Ca^{2+} concentration in the ECF in the cytosol, and **(ii)** Ca^{2+} pools in the lumens of endoplasmic and sarcoplasmic reticulums.

3.3.5 Ion gradient-driven active transports

In these cases, endergonic active transports are conducted by coupling them with the exergonic dissipation of transmembranes ion gradients established by processes like mitochondrial electron transport and photosynthetic reactions. You will find below how two such ion gradient-driven active transports are carried out.

(a) ATP-ADP translocator :

This homodimeric transmembrane integral protein of the inner mitochondrial membrane carries out an *electrogenic antiport* by transporting one ATP^{4-} outward and in exhcnage, one ADP^{3-} inward across that membrane. The translocator molecule has a single binding site for an adenine nucleotide (ATP or ADP) at the area of contact between its two peptide subunits—ATP and ADP have to complete with one another for binding to that site. The translocator exists in two alternativc conformations, viz., one with its binding site exposed to the outer mitochondrial compartment adjoining the outer surface of the inner membrane, and the other having its binding site exposed to the mitochondrial matrix. Binding of either ATP or ADP tó the binding site from

any side of the membrane changes the conformation of the translocator, placing its binding site on the other surface of the membrane. Such altermating conformational changes on binding to the adenine nucleotides enable the translocator to carry out their antiport. This antiport is *electrogenic* because of the overall extrusion of one negative charge in the direction of ATP^{4-} movement, and has to be driven by the difference in membrane potential resulting from the inward proton gradient across the inner membrane. The ATP-ADP antiport creates the respective transmembrane gradients of ATP and ADP. It is inhibited by *bongkrekic acid,* a mould antibiotic, and by *atyractyloside,* a plant glycoside, acting respectively from the inner and outer sides of the inner membrane. (See also 2.5.7 for the role of the translocator in regulating oxidative phosphorylation.)

(b) **Na^+ – glucose symport :**

The symport of dietary Na^+ and gloucose, driven by the Na^+ gradient across the luminal membrane of brush-bordered intestinal epithelial cells, helps in their intestinal absorptions from food. The mechanism depends upon three membrane transport systems of those cells.

(i) The **Na^+ – K^+ ATPase** (Na^+ – K^+ pump) of the basal membrane of the cell hydrolyzes a high-energy bond of an ATP molecule and utilizes that energy to bring about an active Na^+ – K^+ antiport transporting three Na^+ ions outward and two K^+ ions inward across the basal plasma membrane (Fig. 3.3). This keeps the intracellular

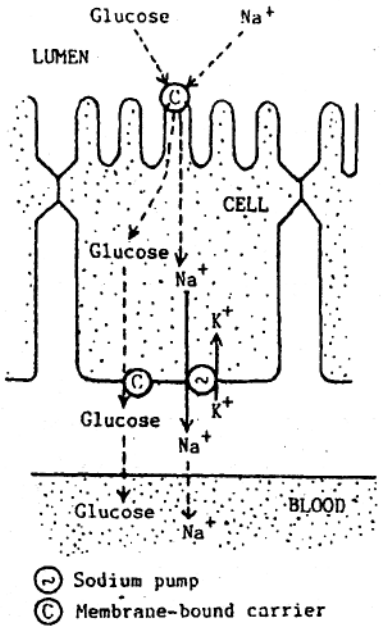


Figure 3.3 : Intestinal Na^+–glucose sumport. [From D. Das, *Biochemistry*, Academic Publishers, 2000]

Na^+ concentration lower than the Na^+ concentration in the intestinal lumen to provide an inward Na^+ gradient from the lumen to the cell interior.

(ii) **A Na^+ glucose cotransporter** of the luminal plasma membrane of the cell utilizes this Na^+ gradient for concentrating both Na^+ and glucose from the lumen on that membrane; for this, luminal Na^+ first binds to the Na^+-binding site of the cotransporter to enhance its glucose-affinity, folowed by the subsequent binding of luminal glucose to the glucose-binding site of the cotransporter. The successive bindings of Na^+ and glucose changes the conformation of the cotransporter; exposing both the binding sites on the cytoplasmic side of the membrane, lowering their affinities for the respective ligands and releasing them in the cytoplasm. The cotransporter thereby regains its original conformation and the substrate-affinities of its binding sites which are again exposed now on the luminal surface of the membrane. *Phlorizin* inhibits this Na^+ -dependent cotransporter by binding to its domains on the extracellular surface of the membrane.

(iii) **A passive mediated glucose transport system** of the basal plasma membrane finally transpoprts glucose across that membrane passively—down the outward concentration gradient of glucose—to release it into the ECF (**facilitated diffusion**).

3.4 Biosynthesis of Urea

Urea is the principal nitrogenous end product of amino acid catabolism in *ureotelic animals* such as mammals, elasmo-branches, terrestrial amphibia and aquatic reptiles. It is the main uninary nonprotein nitrogenous (NPN) waste product in those animals. Most of the urea is synthesized in the liver of ureotelese by the *arginine-urea pathway* which was first outlined by Krebs and Henseleit and has since been considerably elaborated and modified. In this process, ammonia from catabolized amino acids and the amino group of aspartate serve as the sources of the nitrogens of urea while HCO^-_3 (or CO_2) acts as the source of its carbon. Ornithine, another amino acid, participates, in the pathway, but is ultimately obtained back after passing through the successive intermediates, citrulline and arginine, which are also amino acids. The overall reaction of the pathway is summarized below :

$$NH_4^+ + HCO_3^- + 3ATP + \text{aspartate} + 2H_2O \longrightarrow 2ADP + AMP + 4Pi + \text{fumarate} + \text{urea}.$$

The initial reactions of the pathway take place in hepatic *mitochondria* where ammonia, bicarbonate, ornithine and two ATP molecules interact to produce citrulline. The latter is then translocated to the hepatic *cytosol* where aspartate interacts with it to give rise to arginine. Arginine is finally hydrolyzed to ornithine and urea in the cytosol. These steps of the arginine-urea cycle are elaborated below (Fig. 3.4).

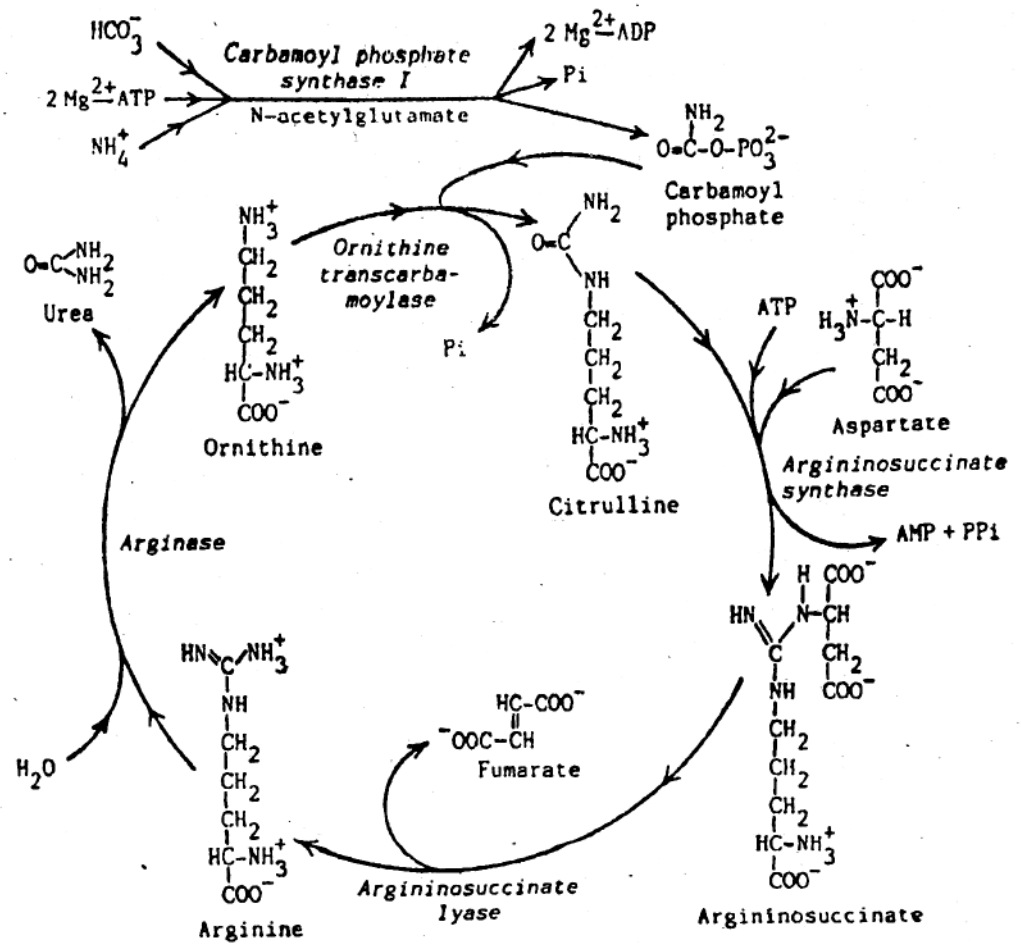


Figure 3.3 : The arninine-urea pathway [From D. Das, *Biochemistry*, Academic Publishers, 2000]

(i) Mitochondrial *carbomoyl phosphate synthase I* condenses free ammonia obtained from amino acid catabolism, with HCO_3^- to form carbomoyl phosphate at the expense of two high-energy bonds of ATP.

$NH_4^+ + HCO_3^-$ ATP → carbamate + ADP + Pi;

ATP + carbamate → carbamoyl phosphate + ADP.

This enzyme is allosterically activated by N-acetyl-glutamate, which enhances its affinity for ATP molecules largely.

(ii) *Ornithine transcarbamoylase,* closely associated with the carbomoyl phosphate synthase I in mitochondria, next transfers the carbamoyl group from carbamoly phosphate to L-ornithine, changing the latter to L-citrulline.

Carbamoyl phosphate + L – ornithine → L-citrulline + Pi

(iii) A *citrulline-ornithine transporter* of the inner mitochondrial membrane then carries but a *passive mediated antiport* of citrulline and ornithine, transferring citrulline outward to the cytosol across that membrane and in exchange transferring ornithine inwards from the cytosol to the mitochondrial matrix.

(iv) In the cytosol, *argininososuccinate synthase* catalyzes the stepwise reaction between L-citrulline and L-aspartate to from L-argininosuccinate at the expense of both high-energy phosphate bonds of an ATP.

ATP + L-citrulline → AMP - citrulline + PPi;

AMP - citrulline + L-aspartate → L-argininosuccinate + AMP.

Thus, amino groups of different amino acids may be transferred by transamination to oxalocetate, changing it to aspartate; the latter then forms argininosuccinate in the arginine-urea pathway.

(v) Another cytosolic enzyme, *argininosuccinate lyase*, next splits L-argininosuccinate into L-arginine and fumarate. The latter may be changed in the TCA cycle to oxaloacetate which may be transminated again to aspartate for use in the arginine urea pathway.

(vi) Finally, *arginase* of hepatic cytosol hydrolyzes L-arginine into ornithine and urea. Ornithine is next translocated back into the mitochondrion by the citrulline-ornithine transporter (see above) and is used again in the arginine-urea pathway.

Energy expenditure :

A total of *four high-energy bonds* of three ATP molecules are spent as follows in synthesizing one urea molecule. (i) Two high-energy bonds, one from each of two ATP molecules are spent in synthesizing carbamoyl phosphate by the action of *carbamoyl phosphate synthase I.* (ii) Two more high-energy bonds of a third ATP molecule are spent in the synthesis of L-argininosuccinate by *argininosuccinate synthase.*

Significnace of urea synthesis :

(i) Highly toxic ammonia, formed by amino acid catabolism, is *detoxicated* in ureoteles by being converted to less toxic urea before the latter can be transported by blood to the kidneys for urinary eliminations. Genetic disorders affecting the arginine-urea pathway produces abnormal rise in blood ammonia and sympotoms of ammonia toxicity.

(ii) Arginine, formed as an intermediate in urea synthesis, constitutes a major nondietary source of this amino acid in ureotelic organisms. Uricoteles such as birds and lizards as well as ammonoteles such as fishes and salamanders do not run the arginine urea pathway and eesentially require the dietary supply of this amino acid.

Regulation of urea synthesis :

(i) The rate of urea synthesis os largely regulated by the *allosteric modulation* of its rate limiting enzyme, *carbamoyl phosphate synthase I.* N-Acetylglutamate binds to the positive allosteric site of the enzyme and enhances its substrate-affinity for ATP to enhance the urea synthesis rate.

(ii) The rate of urea synthesis is decreased by the *feed-back inhibitions* of carbamoyl phosphate synthase I ornithine transcarbamoylase and argininosuccinate systhase by the respective products of their actions, viz., carbamoyl phosphate, citrulline and argininosuccinate.

3.5 Glycogenesis

Glycogenesis is the biosynthesis of glycogen, a homoglycan polysaccharide, from glucose in the *cytoplasm* of many tissue cells. It is carried out most extensively in liver and muscles and to much smaller extents in other tissue cells, except brain cells, erythrocytes and renal cells.

(a) Glucose is firs tphosphorylated to glucose 6-phospahte by either *glucokinase* or *hexokinases* I, II and III, spending a high-energy phosphate bond of ATP (Fig. 3.5). Of these four isozymes, glucokinase predominates in the liver over the other three, and has a higher substrate-specificity but a lower substrate-affinity for glucose. Hexokinases predominate in extrahepatic tissues such as muscles and intestinal mucosa, possess less substrate-specificities but higher substrate-affinities for glucose, compared to glucokinase.

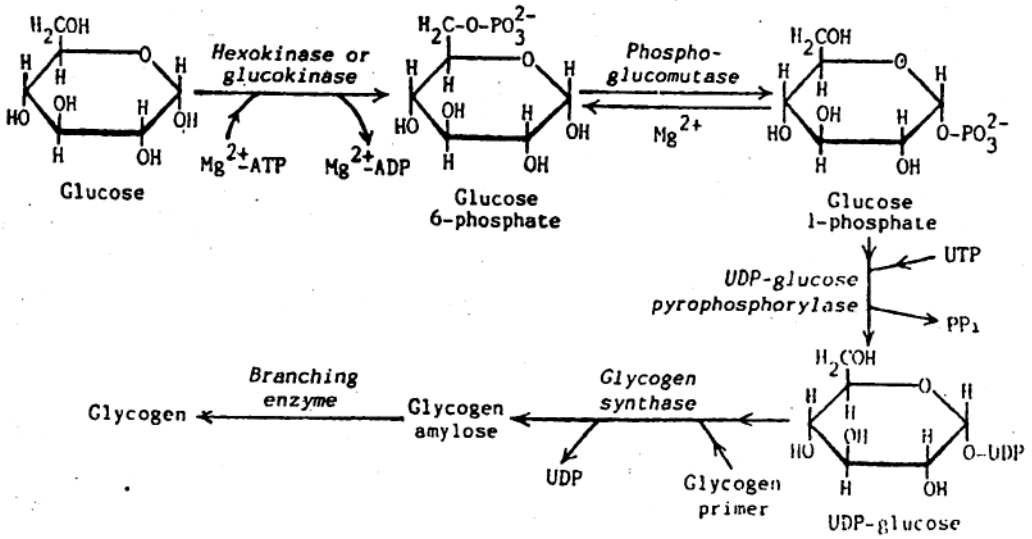


Figure 3.5 : Pathway for glycogenesis. [From D. Das, *Biochemistry*, Academic Publishers, 2000]

(b) *Phosphoglucomutase* next catalyzes the isomerization of glucose 6-phosphate to glucose 1-phosphate through an enzyme-bound glucose, 1, 6-bisphosphate intermediate. The latter is formed by the transfer of a phosphate group from a phosphoserine residue of the enzyme to glucose 6-phosphate, and is subsequently changed to glucose 1-phosphate by donating its C^6–phospahte group to the smae serine residue of the enzyme.

Enz-phosphoserine + glucose 6-phosphate → Enz-serine + glucose 1, 6-bisphosphate;

Enz-serine + glucose 1, 6-biphosphate → Enz-phosphoserine + glucose 1-phosphate

(c) *UDP-glucose pyrophosphorylase* next replaces the two terminal phosphoryl groups of UTP by glucose 1-phosphate, relasing PPi and forming UDP-glucose (Fig.3.5.).

(d) *Glycogen synthase* then transfers the glucose residue of UDP-glucose to the nonreducing end of a pre-existing polyglucose chain (*glycogen primer).* The C^1 of each newly added glucose molecule gets bound by an 1, 4-*glycosidic bond* to the C^4– OH of the terminal glucose residue of the primer molecule. Each such addition extends the polyglucose chain by one glucose residue. Because a poly-glucose chain is thus extended by the successive additions of glucose molecules by identical 1, 4–glycoside bonds only, the product becomes a *glycogen amylose* molecule with unbranched chains.

(e) After ten or more molecules of glucose have been so added to an unbranched polyglucose chain of glycogenamylose, *amylo–1, 4 → 1, 6-transglycosylase* (branching enzyme) transfers the terminal oligosaccharide chain of five or more glucose units from its nonreducing end to the nonreducing end of another such polyglucose chain of the molecule, joining the transferred chain by an 1, 6-glycoside bond to the C^6 – OH of the terminal glucose, of that new chain, resulting in the branching of the latter, as occurs regularly in a glycogen chain (Fig. 3.6).

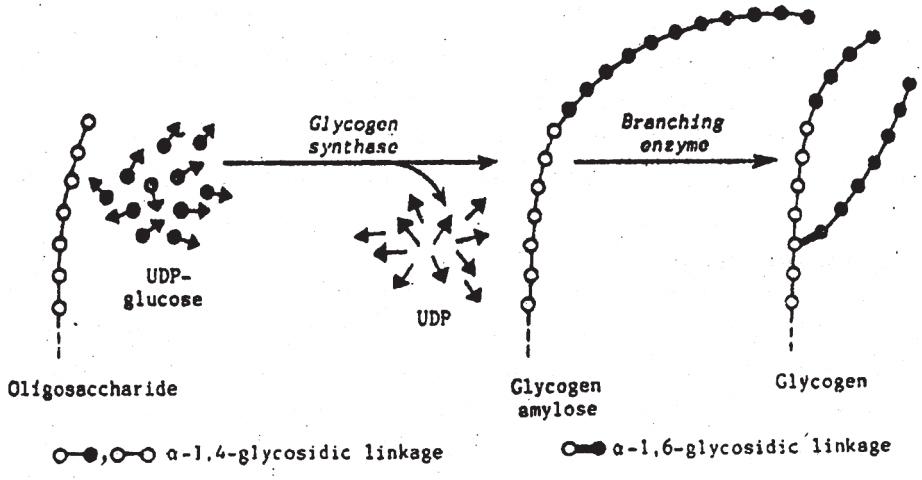


Figure 3.6 : Actions of glycogenasynthase and branchinagenzyme. [From D. Das, *Biochemistry*, Academic Publishers, 2005]

Such successive actions of glycogen synthase and branching enzyme are repeated to form branched glycogen molecules by the polymerization of glucose.

Energetics :

Two high-energy phosphate bonds are spend in adding each glucose residue to the glycogen amylose molecule in the following reactions. **(i)** One ATP bond is hydrolyzed for phosphorylating glucose to glucose 6-phosphate by hexokinase or glucokinase. **(ii)** A second ATP bond is spent in rephosphorylating UDP, released by the action of glycogen synthase from UDP-glucose, to UTP by the action of *nucleoside diphosphokinase.* Thus *two high-energy phosphate bonds* of ATP are spent for extending the polyglucose chain of glycogen amylose by one glucose unit.

Significances of glycogenesis :

(i) Blood glucose that is in excess of the immediate need for energy production is stored in liver and muscles as glycogen for future utilization in energy production and in maintaining the blood sugar.

(ii) Polymerization of glucose to glycogen decreases the bulk of stored glycogen.

(iii) Glycogenesis enables, the cellular storage of large amounts of insoluble glycogen granules without any significant rise in the intracellular osmolarity.

Regulation of glycogenesis :

(i) Insulin *induces* the synthesis of hepatic *glycogen synthase,* the rate-limiting enzyme of glycogenesis, and *glucokinase,* leading to enhanced glycogenesis. This is a long-term regulation of glycogenesis.

(ii) Its short-term regulation depends on *reversible covalent modifications* of *glycogen synthase* by phosphorylation and dephosphorylation. Phosphorylation of specific serine residues of the enzyme to phosphoserine residues by an ATP-dependent *protein Kinase* decreases the activity of the synthase and causes a decline in glycogenesis; dephosphorylation of those phosphoserine residues by a protein phosphatase increase the activity of the synthase with a consequent rise in glycogenesis.

3.6 Gluconeogenesis

Gluconeogenesis is the process of biosynthesis of glucose from noncarbohydrate sources. Gluconeogenesis from metabolites of glycolysis such as pyruvate and lactate takes place mainly in the liver. Moreover, particularly during starvation, many amino acids such as alanine and serine are used in gluconeogenesis in both liver and kidneys. Gluconeogenesis is carried out partly in mitochondria and partly in the cytoplasm.

3.6.1 Significances of gluconeogensis

(i) Lactate and other products of glycolysis, released by contracting muscles and other extrahepatic tissues into the blood, are converted by the liver into glucose which is transported to those tissues by blood for energy production and for replenishing their depleted glycogen stores (*Cori cycle*).

(ii) During fasting or on a poor carbohydrate intake, gluconeogenesis from the catabolites of amino acids from tissue proteins helps to prevent hypoglycemia and to maintain the supply of blood glucose to extrahepatic tissues such as the brain, muscles, heart and erythrocyts.

(iii) Gluconeogenesis serves for the final metabolism of the carbon-skeletons of glycogenic amino acids after their nitrogenous parts have been removed and catabolized.

3.6.2 Gluconeogenesis from lactate

(a) Lactate is first reoxidized to pyruvate by the NAD^+ -dependent *lactate dehydrogenase* in the hepatocyte cytosol.

(b) Pyruvate is next transported into the mitochondrial matrix by a *pyruvate-H^+ symport* carried out by a *Pyravata transporter* of the inner mitochondrial membrane.

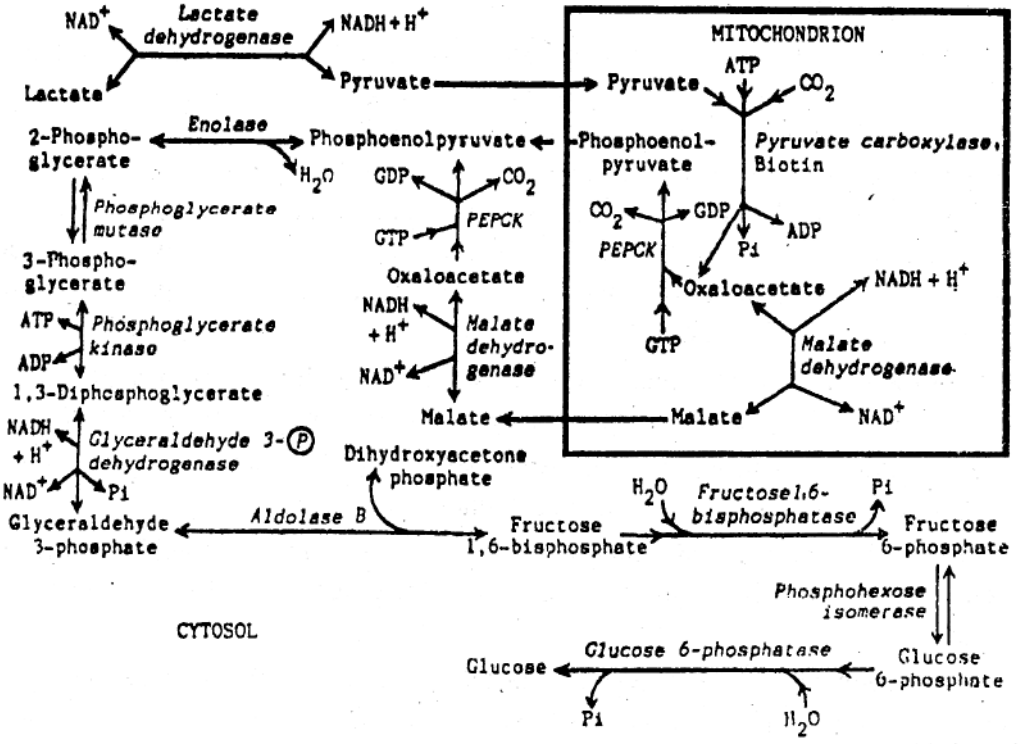


Figure 3.7 : Gluconeogenesis from lactate, [From D. Das, *Biochemistry*, Academic Publishers, 2000]

(c) Biotin-dependent *pyruvate carboxylase* of mitochondrial matrix then carboxylates pyruvate to oxaloacetate, using its biotin prosthetic group, CO_2 and the energy from a high energy bond of ATP (Fig. 3.7).

Enz-biotin + ATP + CO_2 → Enz-carboxybiotin + ADP + Pi,

Enz-carboxybiotin + pyruvate → Enz-Biotin + oxaloacetate

(d) In some animals such as lagomorphs and pigeons, mitochondrial *phosphoemolpyruvate carboxykinase* (PEPCK) first decarboxylates oxalocetate to an enolate intermediate and then phosphorylates the latter into phosphoenolpyruvate (PEP) using a high-phosphate group from GTP; PEP is then transfered to the cytosol by an inner membrane transporter.

Oxaloacetate → CO_2 + enolate intermediate;

GTP + enolate → GDP + PEP.

In rodents, on the contrary, oxaloacetate is reduced by mitochondrial *malate dehydrogenase* and NADH to malate which is translocated to the cytosol by a *malate-HPO_4^{2-} antiport* brought about a *dicarboxylate transporter* of the inner membrane. In the cytosol, malate is next *reoxidized to oxalocetate by cytoplasmic malate dehydrogenase and NAD^+; oxaloacetate is then* converted to PEP by the cytoplasmic *PEPCK,* using GTP.

(e) PEP is next converted to fructose 1, 6-bisphosphate in the cytosol by the successive actions of several glycolytic enzymes (Fig. 3.7) ; **(i)** *enolase* hydrates PEP to 2-phosphoglycerate, **(ii)** the latter is isomerized by *phosphoglycerate mutase* to 3-phosphoglycerate which **(iii)** is phosphorylated by ATP and *phosphoglycerate kinase* to 1, 3-bisphosphoglycerate, **(iv)** the latter is dephosphorylated and reduced by NADH and *glyceraldehyde 3-phosphate dehydrogenase* to give Pi and glyceraldehyde 3-phosphate, and **(v)** the latter is condensed with dihydroxyacetone phosphate by *aldolase* B, giving fructose1, 6-bisphosphate.

(f) Fructose 1, 6-bisphosphate is hydrolyzed into fructose 6-phosphate and Pi by *fructose 1, 6-bisphosphatase,* a specific cytosolic enzyme for gluconeogenesis.

(g) Fructose 6-phosphate is next isomerized by a glycolytic enzyme, *phosphohexose isomerase,* into glucose 6-phosphate.

(h) Finally, microsomal *glucose 6-phosphatase,* another specific enzyme for glucoeogenesis, hydrolyzes glucose 6-phosphate to glucose which may then pass into the blood.

For gluconeogenesis from each lactate (C_3) moleule, **(i)** one high-energy ATP bond is spent during pyruvate carboxylase action, (ii) one high-energy bond of GTP is spent in PECK action, and **(ii)** one high-energy bond of GTP is spent in PEPCK

action, and **(iii)** a second ATP bond is spent during the action of phosphoglycerate kinase. As two lactate molecules are used in forming each glucose (C_6) molecule, a total of *six-energy phosphate bonds* are spent in the gluconeogenesis of one glucose molecule from lactate.

3.6.3 Gluconeogenesis from amino acids

Catabolism of *glycogenic amino acids* such as glycine, alanine, serine, aspartate, glutamate, methionine, valine, cysteine, histidine, arginine, proline and hydroxyproline, and of *glycogenic-ketogenic amino acids* such as phynylalanine, tyrosine, tryptophan, threonine and isoleucine, yield gluconeogenic intermediates like pyruvate, fumarate, α-Ketoglutarate, succinate and oxaloacetate, which are changed to glucose by successive actions of enzymes of carbohydrate metabolism including the glycolytic and specific gluconeogenic enzymes described above (See 3.6.2 and Fig. 3.7). But Ketogenic amino acids, viz., leucine and lysine, do not yield any glucose by gluconeogenesis.

3.7 Glutathione biosynthesis

You may recall that glutathione is a tripeptide, made of glycine, L-glutamate and L-cysteine, and functions as a reducing agent, acts in detoxications and participates in the biosynthesis of pheomelanin pigments of skin. However, unlike the body peptides in general, glutathione cannot be translated by polysomes because its N-terminal peptide bond between glutamate and cysteine is a *γ-peptide bond,* and an α-

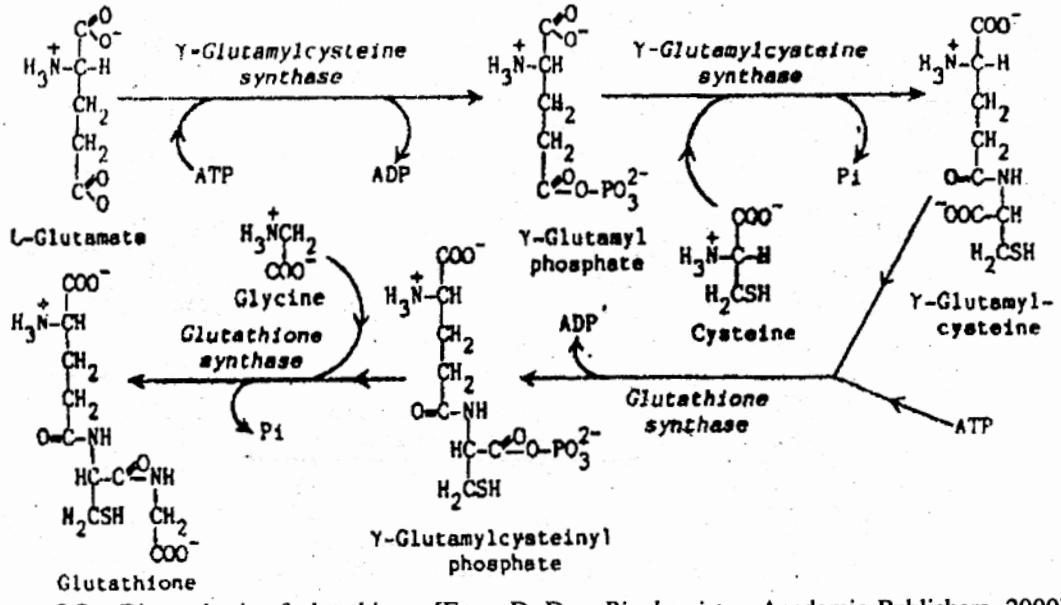


Figure 3.8 : Biosynthesis of glutathione. [From D. Das, *Biochemistry*, Academic Publishers, 2000]

peptide bond. Consequently, glutathione is synthesized by the successive actions of the following cytoplasmic enzymes without the participation of ribosomes, tRNA and mRNA. Each molecule of glutathione is synthesized spending *two high-energy bonds* of ATP.

(i) *γ-Glutamylcysteine synthase* first transfers a phosphate group from ATP to the sidechain γ-carboxyl group of l-glutamate of form γ-glutomyl phosphate, and then replaces that phosphate group by transferring a cysteine molecule in its place, releasing Pi and γ-glutamylcysteine (Fig. 3.8)

(ii) *Gluathione synthase* next phosphorylates the carboxyl group of the cysteine residue of γ-glutamylcysteine using ATP, and then replaces that phosphate group by glycine, releasing Pi and glutathione.

3.8 Norepinephrine and Epinephrine biosynthesis

You are aware that norepinephrine and epinephrine (also called *noradrenaline* and *adrenaline)* are catechdamines released as neurotransmitters at sympathetic axon terminals and secreted from adrenal medulla as its hormones. These two *biogenic amines* are synthesized from L-tyrosine in adrenergic neurons and adrenal chromaffin cells (Fig. 3.9).

(i) *Tyrosme hydroxylase* uses molecular O_2 and the electron-donor *terahydrobioterin* to hydroxylate L-tyrosine to dihydroxyphenylalanine (dopa).

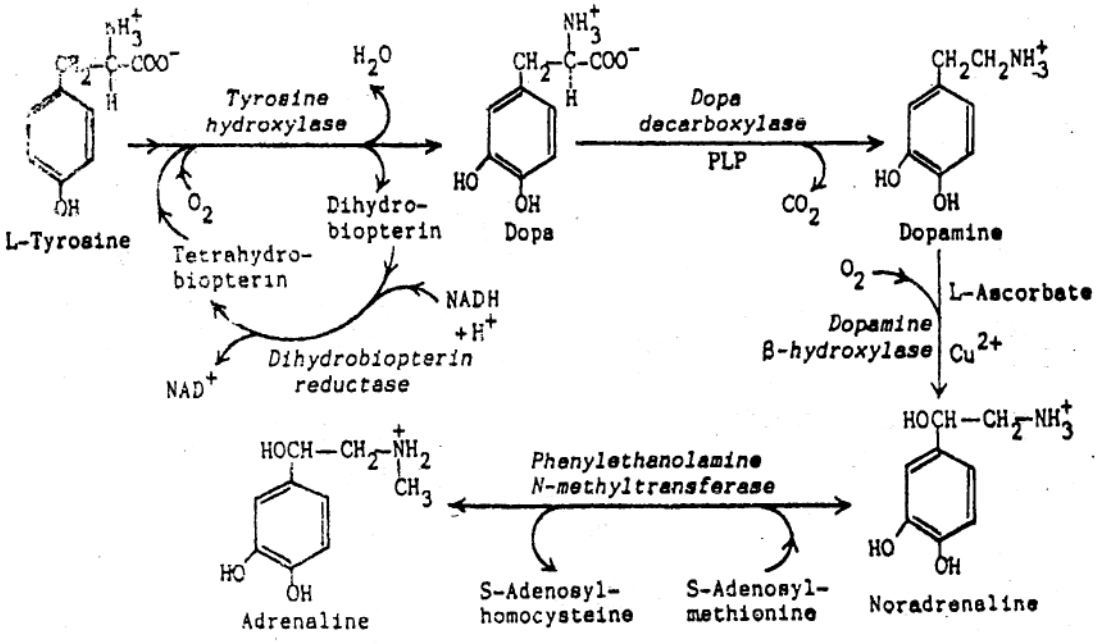


Figure 3.9 : Biosynthesis of catecholamines. [From D. Das, *Biochemistry*, Academic Publishers, 2000]

(ii) *Aromatic L-aminoacid decarboxylase* (dopa decarboxylase), an enzyme with pyridoxal phosphate (PLP) as its prosthetic group, release the carboxyl group of dopa as CO_2 and thus forms dopamine.

(iii) *Dopamine β-hydroxylase,* bearing Cu^+ ion at its active site and using vitamin C as an electron donor cofactor, hydroxylates dopamine to norepinephrine (noradrenaline).

(iv) Finally, *phenylethanolamine N-methyl ferase* transfers the labile methyl group of S-adenosylmethionine ("active" methionine) to the sidechain-NH_2 group of norepinephrine, changing the latter to epinephrine (adrenaline) and the "active" methionine to S-adenosylhomocysteine.

3.9 Serotonin and Melatonin biosynthesis

Serotonin is synthesized from the amino acid L-tryptophan in serotonergic neurons of the hypothalamus, brain-stem, basal ganglia and analgesic neural pathways of the central nervous system and released from their axon terminals. Melatonin is synthesized from serotonin by the pineal gland cells.

(i) *Tryptophan hydroxylase* uses molecular O_2 and the electron-donor *tetrahydrobiopterin* to hydroxylate L-tryptophan to 5-hydroxytryptophan (Fig. 3.10); the sidechain of the latter is next decarboxylated by *aromatic L-amino acid decarboxylase* (5-hydroxytryptophan decarboxylase), an enzyme bearing PLP as its prosthetic group, thus giving rise to serotonin.

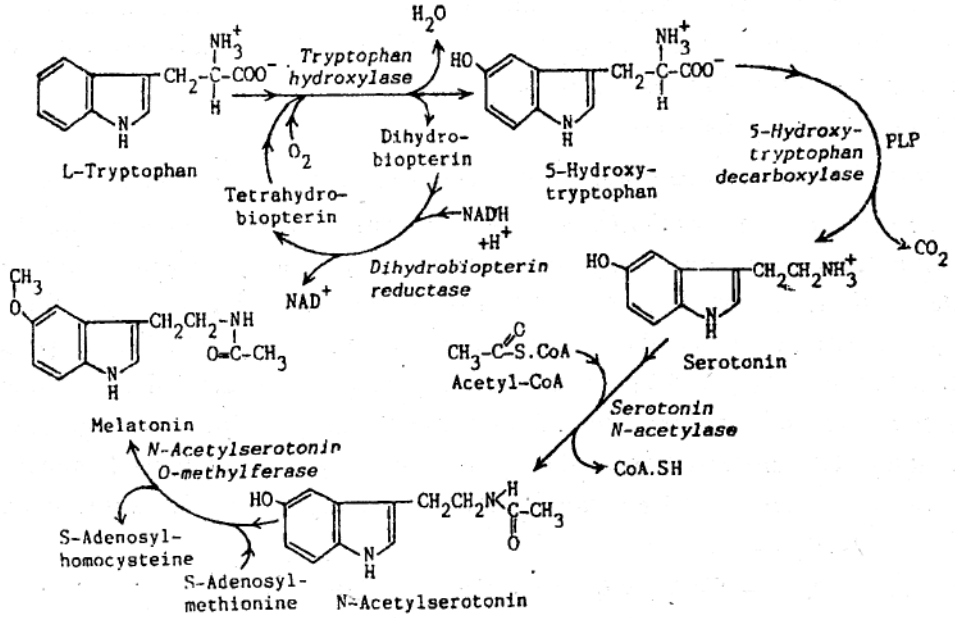


Figure 3.10 : Serotonin and metatonin biuosynthesis [From D. Das, *Biochemistry*, Academic Publishers, 2000]

(ii) *Serotonin N-acetylase* of pineal gland cells transfers the acetyl group of acetyl-CoA to the sidechain—NH_2 group of serotonin to convert the latter to N-acetylserotonin; next, *N-acetylserotonin O-methylferase* transfers the labile methyl group of S-adenosylmethionine ("active" methionine) to C^5 –OH group of N-acetylserotonin, changing them respectively to S-adenosyl homocysteine and melatonin. Melatonin synthesis is photo-regulated.

3.10 Saturated Fatty Acid biosynthesis

In animals such as mammals and birds, tissues like the liver, mammary glands and adipose tissues mostly synthesize palmitic acid (C_{16}) as the basic saturated fatty acid, with the help of two cytosolic multienzyme systems, viz., *fatty acid synthase* and *acetyl-CoA carboxylase.* Longer saturated fatty acids ($>C_{16}$) are subsequently formed by the stepwise addition of C_2 units to palmitate with the help of mainly a *fatty acid elongase* system of smooth ER membranes, while shorter fatty acids ($<C_{16}$) are produced by the stepwise removal of C_2 units from palmitate and other long-chain fatty acids by the mitochondrial β-oxidation enzymes. However, some fatty acid synthase systems of lactating mammary glands are endowed with the ability to turn out short chan fatty acids like butyric and hexanoic acids for secretion in the milk.

3.10.1 Cytoplasmic synthesis of palmitate

The C_2 units (acetyl groups) of acetyl-CoA and the reducing equivalents (H^+ and electron) from the electron-donating coenzyme NADPH constitute the ultimate source materials for synthesizing palmitic acid.

(a) Sources of acetyl-CoA :

Acetyl-CoA used in palmitate synthesis comes from two sources in the cytoplasm.

(i) In *nonruminant animals,* acetyl-CoA formed in the mitochondrion from the oxidative decarboxylation of glycolytic pyruvate, is condensed with oxaloacetate by *citrate synthase* of the TCA cycle to form citrate (Fig. 3.11). When citrate is in excess of the immediate need for its oxidatiion in the TCA cycle for energy production, it is transferred to the cytoplasm by a *citrate-malate antiport* across the inner membrane by a *tricarboxylate transporter* of that membrane and in exchange, malate is simultaneously transferred from the cytoplasm to the mitochondrial matrix. In the cytoplasm, *ATP-citrate lyase* cleaves citrate into acetyl-CoA and oxaloacetate, using ATP and coenzyme A.

Citrate + ATP + CoA → Citryl – CoA + ADP + Pi

Citryl – CoA → oxaloacetate + acetyle – CoA

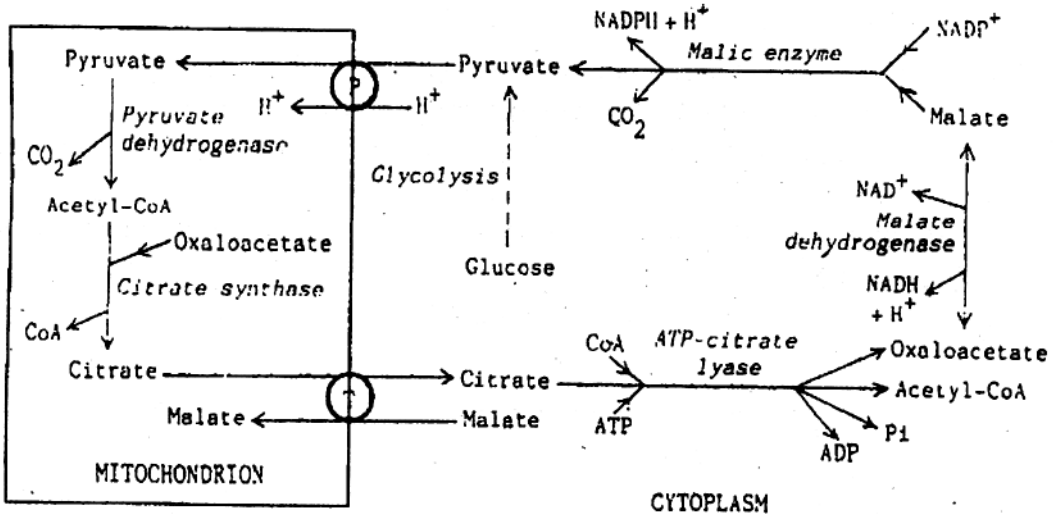


Figure 3.11 : Provision for acetyl-CoA and NADPH in cytoplasm for palmitate synthesis in non ruminants. P : pyruvate transporter. T : tricarboxylate transporter. [From D. Das, *Biochemistry*, Academic Publishers, 2000]

(ii) In *ruminants,* acetic acid absorbed from the rumen as a product of cellulose fermentation is thioesterified with CoA by cytoplasmic *acetyl-CoA synthase (acetatethiokinase)* to form acetyl-CoA at the cost of two high-energy bonds of ATP.

Acetate + ATP + CoA → acetyl – CoA + AMP + PPi

(b) **Sources of NADPH :**

NADPH is generated in the cytoplasm in more than one way for acting as an electorn-donor in palmitate synthesis.

(i) Both in *ruminants and nonruminants,* $NADP^+$ is reduced to NADPH by the actions of *glucose 6-phosphate dehydrogenase* and *6-phosphogluconate dehydrogenase* in the pentose phosphate pathway. This is why that pathway takes place at high rates in lipgenic tissues like adipose tissues.

(ii) In *nonruminant animals,* oxaloacetate obtained by ATP-citrate lyase action on citrate (see above) is reduced by NADH and cytoplasmic *malate dehydrogenase* to form malate; the latter is next decarboxylated and converted to pyruvate by cytoplasmic *malic enzyme* ($NADP^+$ – malate dehydrogenase), with the simultaneous reduction of $NADP^+$ to NADPH (Fig. 3.11).

Malate + $NADP^+$ → CO_2 + pyruvate + NADPH + H^+

(iii) In *ruminants,* citrate is isomerized to isocitrate in the cytoplasm by *aconitase;* isocitrate is next oxidized by $NADP^+$—dependent cytoplasmic *isocitrate dehydrogenase* and then decarboxylated to α-ketoglutarate, simultaneously reducing $NADP^+$ to NADPH.

Isocitrate + $NADP^+$ → CO_2 + NADPH + H^+ + α-Ketoglutarate.

(c) Acetyl-CoA carboxylase action :

Acetyl-CoA carboxylase, a cytoplasmic *multienzyme protein* carboxylates acetyl-CoA to malomyl-CoA, using HCO_3^- and spending an ATP bond (Fig. 3.12). This enzyme system bears a *biotin prosthetic group* on one of its peptide subunits and active sites for *biotin carboxylase* and *transcarboxylase* on two other subunits. The *biotin carboxylase* uses a high-energy ATP bond to carboxylate the biotin prosthetic group into carboxybiotin.

$$ATP + Enz\text{-}biotin + HCO_3^- \rightarrow ADP + Pi + Enz\text{-}carboxybiotin$$

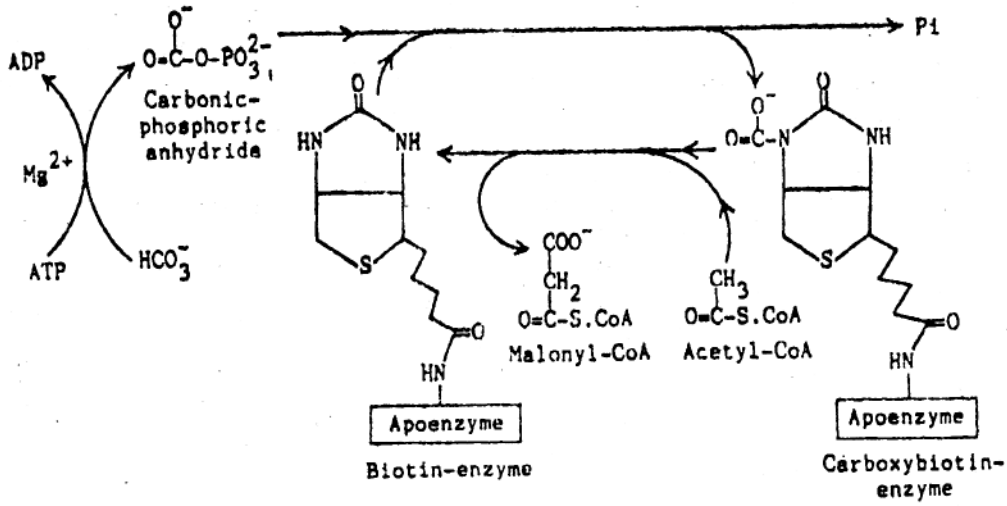


Figure 3.12 : Acetyl-CoA carboxylase action. [From D. Das, *Biochemistry*, Academic Publishers, 2000]

Transcarboxylase then carboxylates acetyl-CoA to malonyl-CoA by the transfer of the carboxylate ion to the substrate from the carboxybiotin prosthetic group.

$$Enz\text{-}carboxybiotin + acetyl\text{-}CoA \rightarrow Enz\text{-}biotin + malonyl - CoA.$$

(d) Falty acid synthase :

Sevel malonyl-CoA molecule and are acetyl-CoA molecule are next used by fatty acid synthase, another cytoplasmic *multienzyme protein,* in synthesizing one palmitic acid molecule through seven cyles of reactions, in mammals and birds. Overall :

$$Acetyl - CoA + 7\ malonyl\text{-}CoA + 14\ NADPH + 14H^+ \rightarrow palmitate + 7CO_2 + 8CoA + 14\ NADP^+ + 6H_2O$$

In these animals, this multienzyme protein exists as an ellipsoid homodimer, its two monomers being noncovalently bound to each other in an antiparallel manner with the N-terminal end of each facing the C-terminal end of the other. Each minomer

consists of N-terminal *domain-I (condensation unit);* a middle *domain-2 (reduction unit),* a C-terminal *domain-3 (palminate release unit),* and these domains face respectively the domains 3, 2 and 1 of the other monomer. The domains of each monomer carry the following active sites : **(i)** domain 1: 3-ketoacyl synthase, malonyl transacylase and acetyltransacylase; **(ii)** domain-2 : 3-hydroxylacyl dehydratase, enoyl reductase, 3-Ketoacyl reductase and acyl-carrier protein (ACP); **(iii)** domain 3 : thioesterase. The 3-ketoacyl synthase active-site carries of cysteine residue with a sidechian SH (Cys-SH) group, the ACP has a 4'-phosphopantetheine prosthetic group with another SH group (Pan-SH), and the Cys-SH of one monomer faces the Pan-SH of the other and vice versa. The two monomers, as long as held together as the homodimer, act as the multienzyme synthase, synthesizing two palmitate molecules simultaneously from the two ends of the dimer; with the coordinated participation of the Cys-SH of each monomer and the Pan-SH of the other. In this process, the enzymes of domain-1 of each monomer and those of domains-2 and 3 of the opposite monomer function together as one unit.

In contrast, bacteria, **euglena** and potato tuber cells carry separate and individual enzymes and the ACP molecule each occurring singly to participate in fatty acid synthesis, instead of existing and acting as a dimeric multienzyme system.

(e) Reactions catalyzed by the multienzyme synthase :

(a) Acetyl transacylase and **malonyl transacylase** of domain-1 of each monomer of the multienzyme complex transfers the acetyl group from an acetyl-CoA and the malonyl group from a malonyl-CoA respectively to the Cys-SH of the same monomer and to the Pan-SH of the other, forming an acetyl-malonyl-enzyme (acyl-malonyl-Enz) complex (Fig.3.13).

(b) 3-Ketoacylsynthase of domain-1 of the first monomer transfers the acetyl group from its Cys-SH to the methylene –C (C^2) of the malonyl group held by the domain-2 Pan-SH of the other monomer, releasing the carboxyl-C of that malonyl group as CO_2 and forming a Pan-S-bound 3-Ketoacyl group on the second monomer.

(c) **3-Ketoacyl reductase** of domain-2 of the second monomer next used NADPH to reduce the Pan-S-bound 3-ketoacyl group to D(–)-3-hydroxyacyl group.

(d) **3-Hydroxyacyl dehydratase** of the same domain-2 then changes the Pan-S-bound D(–)-3-hydroxyacyl group to Δ^2-*trans* enoyl group by releasing H_2O from it.

(e) **Enyl reductase** of the same domain-2 next reduces the Δ^2-double bond in the Pan-S-held enoyl group using NADPH and giving rise to a new Pan-S-held C_4-acyl group at the end of this first cycle of reactions.

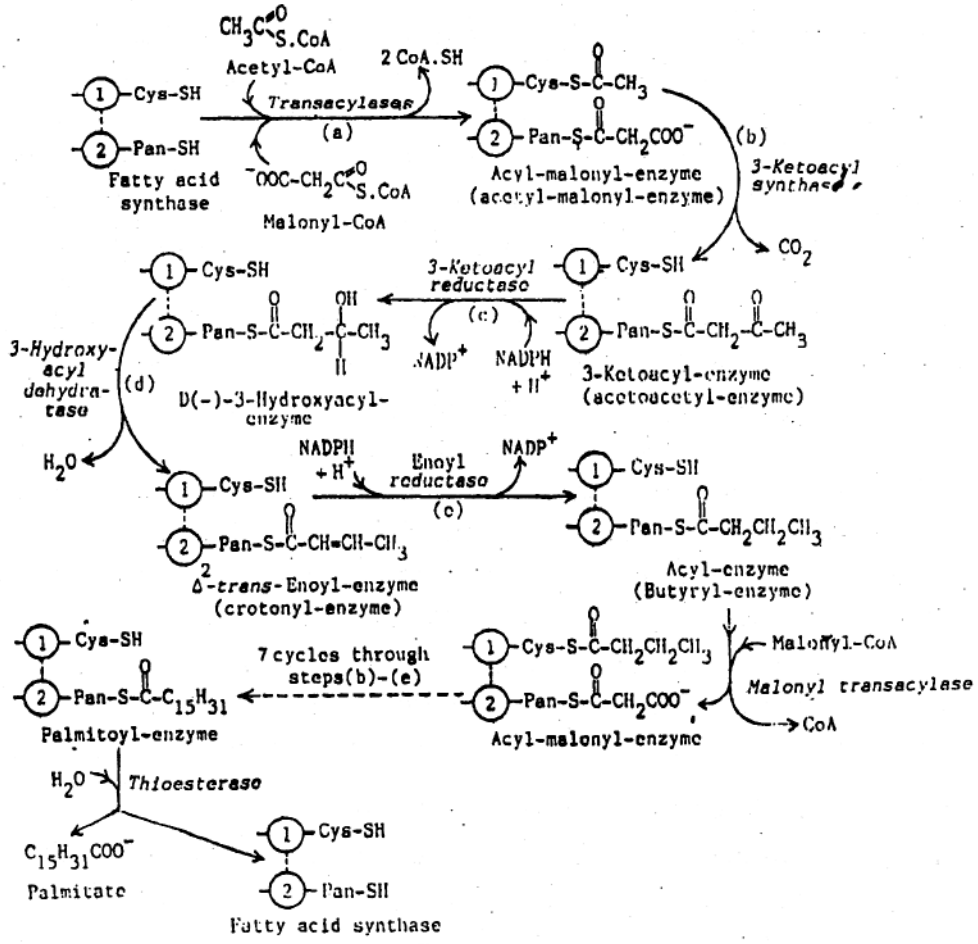


Figure 3.13 : Palmitate synthesis by fatty acid synthesis. [From D. Das, *Biochemistry*, Academic Publishers, 2000]

A second cycle then starts with the transfer of a malonyl group from a fresh malonyl-CoA molecule by *malony transacylase* of domain-1 of the first monomer to the domain-2 Pan-SH group of the second monomer while the Pan-S-held C_4-acyl group gets transferred to the domain-1 Cys-SH of the first monomer. A new acyl-malonyl enzyme complex is thus formed and the second cycle follows through the reactions described in (b) to (e) above.

After seven such cycles, a C_{16}-acyl (palmitoyl) group is formed on the Pan-SH. Now, *thioesterase* of domain-3 of the second monomer hydrolyzes the thioester bond holding the palmitoyl group and releases the latter as palmitic acid. However, lactating mammocytes posses fatty acid synthases having thioesterase action capable of releasing different short and medium-chain fatty acids, instead of palmitic acid, from the Pan-SH. This is how such lower fatty acids occur in the milk.

(f) Energetics of palmitate synthesis :

In synthesizing each palmitate molecule, **(i)** a total of 8 high-energy bonds of ATP are spent by ATP-citrate lyase in cleaving 8 citrate molecules to 8 acetyl-CoA molecules, and **(ii)** 7 more high-energy ATP bonds are spent by acetyl-CoA carboxylase in carboxylating 7 acetyl-CoA molecules to 7 malonyl-CoA molecules. Thus a total of *15 high-energy phosphate bonds* are spent in forming each palmitate molecule.

(g) Regulation of fatty acid synthesis :

Short-term regulation of palmitate synthesis is carried out mainly by regulating *acetyl CoA earboxylase,* the rate-limiting enzyme of the process. (i) *Citrate,* produced by high rates of carbohydrate metabolism when on high-carbohydrate diet, allosterically activates acetyl-CoA carboxylase to raise its Vmax and consequently enhances the availability of malonyl-CoA for palmitate synthesis, thus increasing the rate of the latter. (ii) *Palmitoyl-CoA* and other long-chain acyl-CoA *antagonize the allosteric effect of citrate* on acetyl-CoA carboxylase to decrease palmitate synthesis, when on a high-fat diet. (iii) *Reversible cocalent activation* of acetyl-CoA carboxylase by its *phosphorylation* increases its activity, leading to enhanced palmitate synthesis; *dephosphorylation* produces the opposite effect.

Long-term regulation depends mainly on the *inducing action of insulin* on the synthesis of acetyl-CoA carboxylase, fatty acid synthase, ATP-citrate lyase, and NADPH-generating enzymes such as malic enzyme and the dehydrogenases of the pentose phosphate pathway.

3.10.2 Microsomal elongation of fatty acids

Longer saturated fatty acids ($>C_{16}$) are mostly formed from palmitate by its stepwise elongatin catalyzed by separate individual enzymes of the *microsomal elongase system* of the smooth ER membranes, using C_2 units from malonyl-CoA molecules. You may recall that malonyl-CoA is formed in the cytoplasm by the carboxylation of acetyl-CoA by acetyl-CoA carboxylase; such malonyl-CoA can be used for both the synthesis of palmitate and its elongation to longer fatty acids. However, during their elongation, fatty acids are not carried by any acylcarrier protein nor do the individual enzymes of the elongase system occur as a multienzyme protein. In non-neural tissues, the microsomal elongase system enzymes almost solely catalyze the elongation of palmitic acid (C_{16}) to stearic acid (C_{18}) through the following reactions; but such enzymes of brain microsomes catalyze the elongation of palmitate to much longer ($C_{18} - C_{24}$) fatty acids for brain lipids. Overall reaction in elongating palmitate to sterate is quoted below :

$$\text{Palmitoyl-CoA} + \text{malonyl} - \text{CoA} + 2\text{NADPH} + 2H^+ \rightarrow \text{stearoyl-CoA} + 2\text{NADP}^+ + H_2O$$

(a) Microsomal *3-Ketoacyl-CoA synthase* transfers the palmitoyl group from palmitoyl-CoA to the methylene-C (C^2) of malonyl-CoA, relasing the CoA from palmitoyl-CoA and the carboxyl-C (C^3) of malonyl-CoA as CO_2 (Fig. 3.14). This results in the formation of 3-ketostearoyl-CoA (a 3-ketoacyl-Coa).

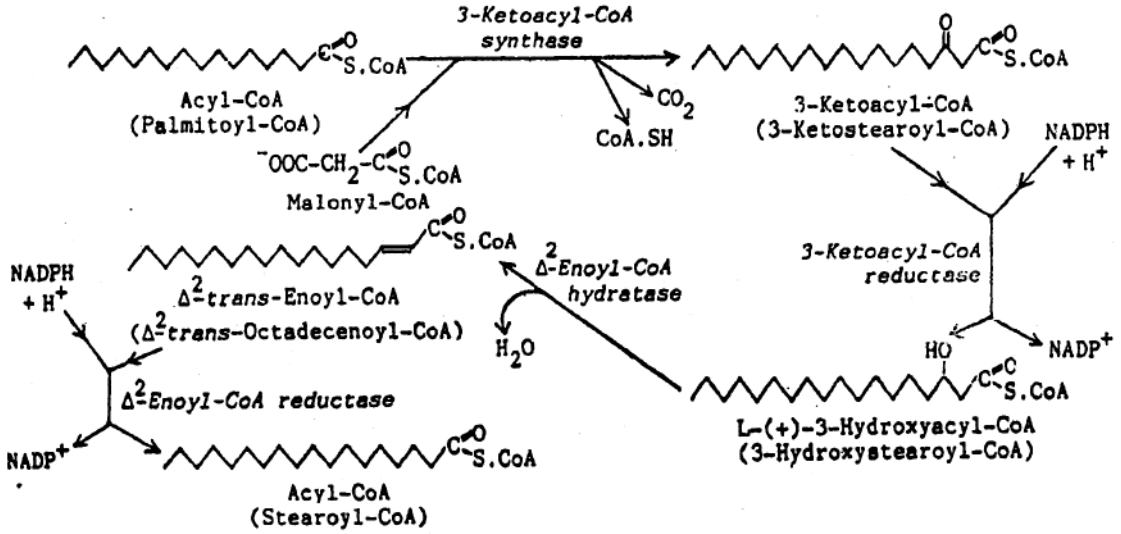


Figure 3.14 : Elongation of polmitoyl-CoA to stearotyl-CoA [From D. Das, *Biochemistry*, Academic Publishers, 2000]

(b) *3-Ketoacyl-CoA reductase* next used NADPH to reduce 3-ketosteraoyl-CoA to L(+)-3-hydroxystearoyl CoA (a 3-hydroxyacyl-CoA).

(c) Δ^2-*Enoyl-CoA hydratase* releases H_2O from L(+)-3-hydroxystearoyl-CoA to give Δ^2-*trans*-octadecenoyl-CoA (a Δ^2-*trans*-enoyl-CoA).

(d) Δ^2-*Enoul-CoA rductase* finally reduces the product of reaction (c) using NADPH, to form a C_2-higher saturated acyl-CoA which, in the present example, is stearoyl-CoA.

3.11 Prostaglandin biosynthesis

Prostaglandins (PG) are monoenoic or polyenoic, hydroxy-substituted and Keto-substituted, cyclic C_{20}-fatty acids. The molecule consists of an unsaturated liner carbon-chain with a saturated C_5-cyclopentase ring in the middle of that chian. Prostaglandins are classified into PG_1, PG_2 and PG_3 series with the subscripts indicating the number of double-bonds in their linear chains. Prostaglandins of each series are sub-classified into PGD, PGE, PGF, etc., according to the positions and types of the substituent groups on their cyclopentane rings (Fig.3.15).

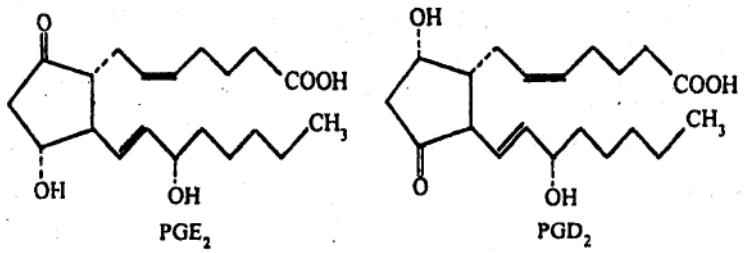


Figure 3.15 : Two prostaglandins.

3.11.1 Cyclo-oxygenase pathway of PG synthesis

Prostaglandins are synthesized by this pathway by enzymes of smooth ER membranes from different C_{20}-polyenoic (polyunsaturated) fatty acids in diverse tissues such as lungs, kidneys, seminal vesicles, vascular endothelia, liver, uterus, brain gastrointestinal tract, thymus, heart, adipocytes and platelets. Prostaglandins of PG_1, PG_2 and PG_3 series, containing respectively 1, 2 and 3 double-bonds in their linear molecular chains, are biosynthesized respectively from dihomo-γ-linotenic acid, arachidonic acid and timnodonic acid, released from membrane phospholipids into the cytoplasm by the action of Ca^{2+} dependent specific *phospholiphases* such as phospholipases A_1, A_2 and C. These three fatty acids released from membranes are C_{20} polyenoic fatty acids bearing respectively 3, 4 and 5 double-bonds. The phospholipase-catalyzed release of these C_{20}-polyenoic acids is the *rate-limiting reaction* of PG biosynthesis—while angiotensin, bradykinin and hypoxia activate the phospolipase to enhance PG synthesis, mepacrine and corticosteroids inhibit the phospholipases and depress PG synthesis. The released C_{20}-polyenoic acids are converted to prostaglandins by the *cyclo-oxygenase pathway* summarized below (Fig. 3.16). The principal enzyme system of this pathway is a microsomal heme-containing multienzymes system, *prostaglandin endoperoxide synthase.*

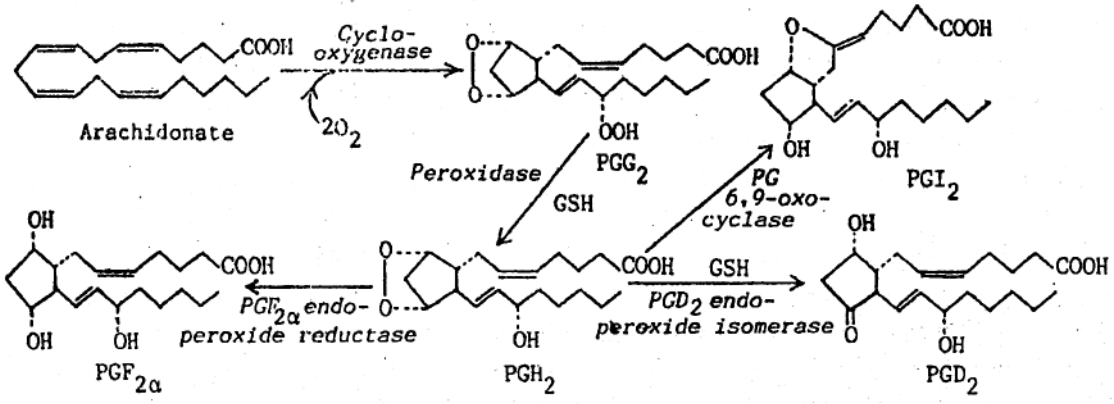


Figure 3.16 : Cyclo-oxygenase pathway for PG_2 series. [From D. Das, *Biochemistry*, Academic Publishers, 2000]

(a) The *fatty acid cyclo-oxygenase* component of the PG endoperoxide synthase system oxidizes and cyclizes the C_{20}-polyenoic acids (e.g., arachidonic acid) into the respective unstable cyclic endoperoxides (e.g. PGG_2 from arachidnate).

(b) The *Peroxidase* component of the PG endoperoxide synthase system changes the endoperoxides (e.g. PGG_2) into the other respective unstable cyclic endoperoxides (e.g., PGH_2) using glutathione as a reducing cofactor and H_2O_2 as an electorn-acceptor.

(c) Microsomal PG *endoperoxide isomerases* may then use glutathione as a cofactor in isomerizing the unstable PG endoperoxides into prostaglandins like PGD_2 and PGE_2.

(d) *Prostaglandin endoperoxide reductases* may then reduce prostaglandins like PGH_2 and PGE_2 into still other prostaglandins, like PGF_{2a} from PGH_2 or PGF_2.

Catecholamines, vasopressin and angiotension enhance PG biosynthesis by increasing PG endoperoxide synthase activity; acetylsalicylate, indomethacin and nonsteroid antiinflammatory agents inactivate the fatty acid cyclo-oxygenase and thereby decrease PG biosynthesis.

3.12 Muscle contraction

There are significant differences between the molecular mechanisms of contrcaction of striated and smooth muscles. These basiclly arise from differences between their muscle proteins. Here, you will read first about muscle proteins of striated muscles and their roles in the mechanism of contraction. You will next read about how the differences in muscle proteins between striated and smooth muscle fibres result in a distinctive mechanism of smooth muscle contraction.

3.12.1 Major muscle proteins

Dark anisotropic and electron-dense A bands of each myofibril of a striated muscle fibre contain *thick protein myofilaments,* made of parallel and staggered molecules of the muscle protein **myosin.** Each hexameric myosin molecule is made of two heavy and four light L peptide chains (Fig. 3.1.7). The α-helical C-terminal segments of the two H chains are entwined into a left-handed coil forming a double-stranded fibrous rod-like domain, while the N-terminal end of each H chain bears a pear-shaped globular S_1 *domain* with *myosin-ATPase* activity—an *essential L chian* (ELC) and a *regulatory L chiain* (RLC) remain associated with the S_1 domain. Trypsin hydrolyzes myosin into a *heavy meromyosin* (HMM) molecule consisting of two S_1 domains connected to a C-terminal rod-like tail of healically coiled fractions of two H chains, and a double-stranded rod-like *light meronyosin* (LMM) molecule consisting of the helically entwined C-terminal ends of the two H chains. Hydrolysis of HMM by papain yields two S_1 *Subfragments* with their associated L chains, and a double-helical rod-like S_2 *subfragment* (Fig. 3.17).

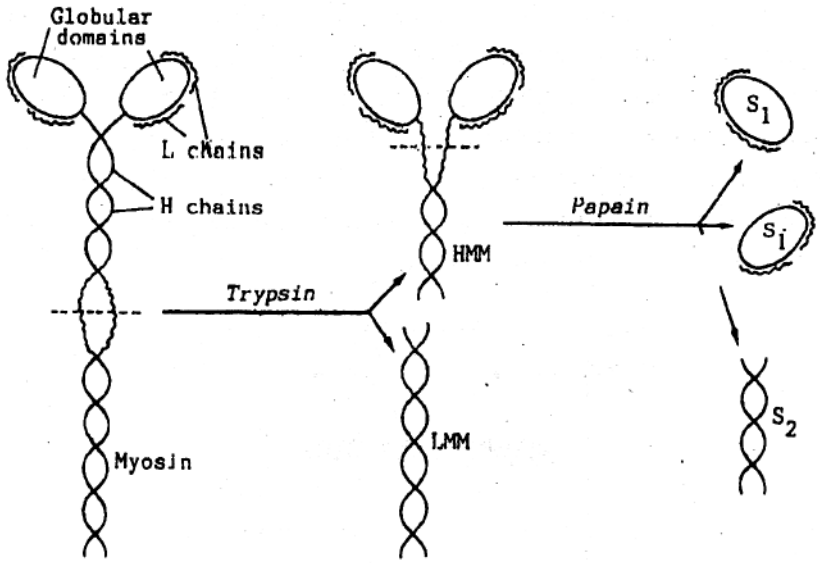


Figure 3.17 : A myolecule and its cleavage frogments. [From D. Das, *Biochemistry*, Academic Publishers, 2000]

In each thick myofilament, many parallel myosin filaments are arranged in two bundles with their N-terminal S_1 domains oriented towards opposite poles at the two margins of the A band and the C-terminal ends of both bundles meeting end-to-end at the less dense H zone in the middle of the A band. Moreover, myosin molecules are staggered both longitudinally and helically along the thick myofilament so that their S_1 domains protrude from the two outer segments of the filament at regular intervals along their and at staggered angles from the filament (Fig. 3.19).

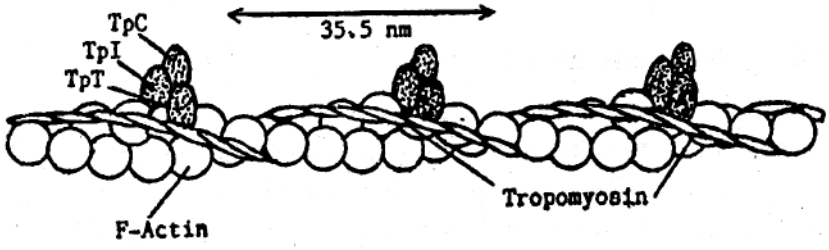


Figure 3.18 : A thin myofilament. [From D. Das, *Biochemistry*, Academic Publishers, 2000]

Light, isotropic I bands consist mainly of parallel, double-stranded, *thin protein myofilaments,* each made of two helically entwined, long and beaded strands of *F-actin polymers* (Fig.3.18). F-actin is formed by the ATP and Mg^{2+} dependent polymerization of soluble *G-actin* molecules at isotonic ionic concentration. Double stranded cables of the protein *tropomyosin* course along the helical central groove of

F-actin double-strands which also remain associated at intervals with three inter-connected globular peptide subunits of *troponin.* The F-actin and tropomyosin double-strands and the troponin subunits together constitute a thin myofilament (Fig. 3.18). F-actin filaments extend from the central Z line of an I band upto some distance into the neighbouring A band and interdigitate with the thick myosin filaments of the latter (Fig. 3.19). F-actin strands acquire an arrow-head pattern along their lengths due to their binding to myosin S_1 heads, giving a polarity to their filaments.

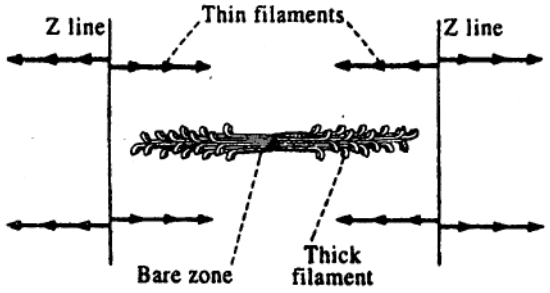


Figure 3.19 : Polarities of thick and thin filaments. [From D. Das, *Biochemistry*, Academic Publishers, 2000]

Peptide subunits of each *troponin system* associated with the F-actin double strand consist of **(i)** *troponin* C (TpC) having Ca^{2+} binding sites, **(ii)** *TpT* binding to tropomyosin and (iii) *Tpl* binding to F-actin and TpC and inhibiting the actin myosin interaction (Fig. 3.18).

Besides the major muscle proteins described abvoe, several minor muscle proteins such as α-actinin, desmin, vimentin, C-protein and M-protein also help in assembling myofilaments in myofibrils.

3.12.2 Sliding filament model

According to this model of muscle contraction, myofilments themselves do not change in length during the contraction or relaxation of the muscle; instead, the muscle fibres are shortened due to the shortening of their *sarcomeres.* Sarcomeres are successive segments of each myofibril and function as the functional units of the latter—each sarcomere extends along the myofibril from one Z line to the next (Fig.3.19). According to this model, thin actin filaments slide during muscle contraction towards the centre of the A band, penetrating deeper between the thick myosin filaments interdigitating with them within that band. Such sliding of actin filaments in between the myosin filaments of the A band draws the Z lines at the two ends of the sarcomere closer to each other, shortening the length of each sarcomere from ~2300 nm in a relaxed myofibril to ~1500nm in its contracted condition and thereby shortening the muscle

fibre. Shortening of the sarcomeres during muscle contraction increases the overlap between the thin actin filaments and the thick myosin filaments in each A band and greatly shortens the lengths of I bands, but does not affect the length of any A band because its length depends solely on the lengths of its myosin filaments slide out from between the interdigitating thick myosin filaments so that the overlap between the two type of filaments decreaes. This enhances the length of each I band considerably and elongates each sarcomere from ~1500 nm in the contracted myofibril to ~2300 nm in the relaxed one, but again leaves the length of the A band unaltered.

3.12.3 Molecular mechanism of muscle contraction

In striated muscles, contraction and relaxation result from an actin-myosin interaction cycle. Contraction of nonstriated muscles, on the contrary, is principally myosin based.

Actin-myosin interaction :

The modern concept of the contraction of a muscle due to the actin-myosin-ATP interaction, and of the ATP-related dissociation of actomyosin (actin-myosin complex) during its relaxation, is based primarily on *Szent-Gyorgi's findings* that **(i)** addition of an action solution to a myosin solutin resulted in an increased viscosity of the latter due to actomyosin formation by actin-myosin interaction, **(ii)** addition of ATP to that mixed solution decreased its viscosity again, indicating the dissociation of actomyosin into actin and myosin, **(iii)** actomyosin threads were found to contract on being immersed in a solution containing Mg^{2+}, K^{+} and ATP, and **(iv)** myosin threads did not contract when dipped into a similar solution.

According to the modern concept of the striated muscles contraction, the attachment of the globular S_1 domains of myosin-heads with the adjacent F-actin strands gives rise to cross-bridges between the myosin and actin filaments, leading to the sliding-in of the thin actin filaments deeper between the thick myosin filaments. This *endergonic* sliding-in of the filaments is accomplished by coupling it with the *exergonic* hydrolysis of ATP by *myosin-ATPase* of the S_1 domains of myosin-heads—the activity of the myosin-ATPase is manifold heightened by the binding of F-actin to myosin. During relaxation, on the contrary, the detachment of the S_1 domains of myosin-heads from F-actin cleaves the cross-bridges between the acting and myosin filaments, leading to the sliding-out of thin actin filaments from between the thick myosin filaments, leading to the sliding-out of thin actin filaments from between the thick myosin filaments. The sequence of reactions in this actin-myosin interaction cycle is described below (Fig.3.20).

(a) During relaxation, *Ca^{2+} –ATPase* of the sarcoplasmic reticulum membrane actively transports Ca^{2+} ions from the sarcoplasm to the SR cisternae where the protein *calsequestrin* binds to the Ca^{2+} ions and sequesters them. This maintains the sarcoplasmic

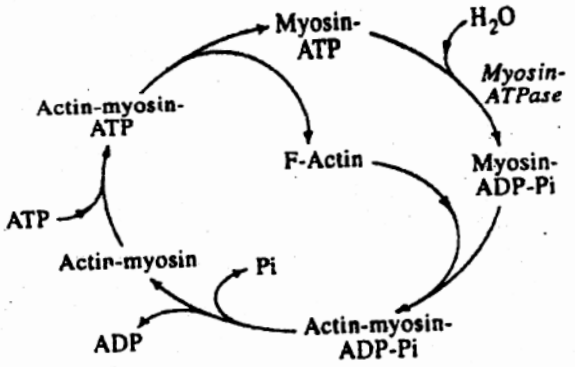


Figure 3.11 : Actin-myosin interaction cycle during contraction and relaxation of striated muscle. [From D. Das, *Biochemistry*, Academic Publishers, 2005]

Ca^{2+}concentration below 10^{-6} M and thus keeps the *trponin* C (TpC) of thin filaments almost free from Ca^{2+} and consequently incapable of promoting any actin-myosin interaction. In other words, this lack of TpC activity permits the *troponin I* (TpI) to continue inhibiting actin-myosin interaction, probably by blocking the S_1-binding sites of F-actin and thereby keeping the latter detached from the myosin S_1 domains. Myosin still binds to ATP which is hydrolyzed by the *myosin-ATPase* of its S_1 domains to ADP and Pi; but in the absence of attachments between F-actin and myosin, most of the ADP and Pi fail to be released from myosin and continue to remain as the high-energy myosin-ADP-Pi complex, retaining the free energy of ATP hydrolysis. In this state, the S_1 headpieces of myosin make an angle of ~90° with the axis of the thick myosin filament and neither actin-myosin interaction nor filament sliding can take place. This continues to prevent the shortening of myofibrils in the relaxed muscle.

(b) When the muscle fibre is stimulated, the depolarization of membranes of its T-tubular system leads to the opening of *ligand-gated Ca^{2+} –release channels* of the SR membrane; the consequent release of Ca^{2+} ions through these opened Ca^{2+} channels into the sarcoplasm from the SR cisternae, where they were so long sequestered by calsequestrin, increases the sarcoplasmic Ca^{2+} concentration above 10^{-5} M, causing the binding of Ca^{2+} to the *troponin C* (TpC) component of F-actin-bound TpC from active TpC-$(Ca^{2+})_4$ comples. The latter allosterically rolls the troponin T-bound *tropomyosin double-stranded cables,* coursing through the central groove of the F-actin filament, to move them deeper into that groove. This makes the S_1-binding sites of F-actin more accessible and enables their electrostatic bonding with the myosin S_1 heads. This counteracts the inhibitory effect of TpI on actin-myosin interaction.

(c) A weak initial binding of F-actin of S_1 headpieces releases Pi from the high-energy myosin-ADP-Pi complex and increases the actin-affinity of myosin S_1 heads. This causes a stronger binding of the two. It leads to a change in the high-energy

conformation of the actin-bound myosin to a low-energy conformation bringing about a rotation of the myosin heads from an angle of 90°–to a 45°–angle with the myosin filament axis (Fig.3.21), releasing ADP from the myosin head and resulting in the formation of actin-myosin complex. The rotation of myosing heads acts as the *power stroke* utilizing the energy released by the conformational change of the actin-bound myosin, and pulls the thin actin filaments to slide them by >7 nm towards the A band centre and between the interdigitating thick myosin filaments.

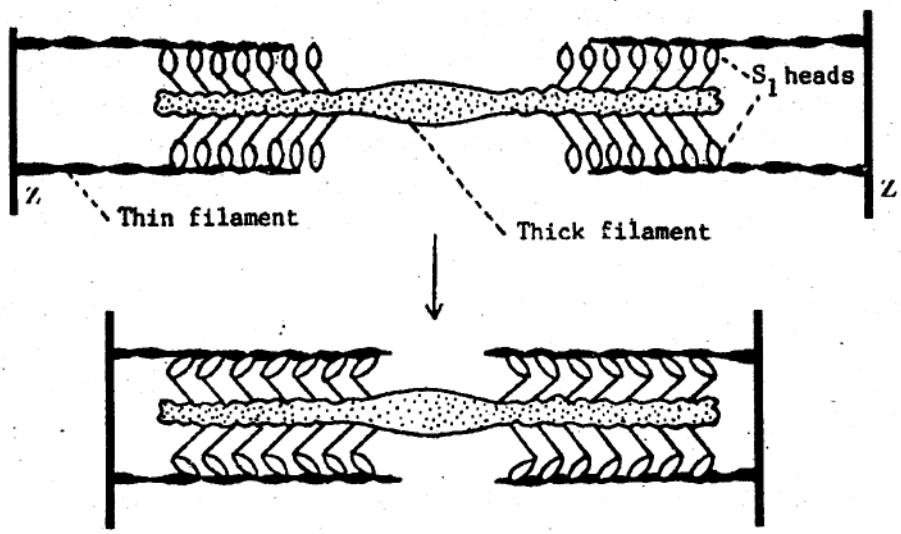


Figure 3.21 : Mechanism of sliding of myofilaments during muscle contraction. [From D. Das, *Biochemistry*, Academic Publishers, 2000]

(d) Next, a fresh ATP molecule binds to the myosin S_1 domain to form an actin-myosin-ATP complex. This weakens the actin-affinity of myosin, detaches the myosin S_1 heads from F-actin, thereby cleaves the cross-bridges between actin and myosin filaments, and make them slide away from each other. This lengthens the myofibril in the relaxed muscle, while myosin is left behind as a myosin-ATP complex.

(e) Myosin-ATPase hydrolyzes the myosin-bound ATP to ADP and Pi, and the resulting myosin-ADP-Pi complex awaits the next cycle of actin-myosin interaction (Fig. 3.20).

3.12.4 Myosin-based contraction

Smooth muscle fibres also depend for their contraction on the interaction between longitudinally oriented thin and thick myofilaments made principally of F-actin and myosin molecules respectively. But smooth muscle fibres differe in many ways from striated muscle fibres with respect to microanatomical structure, chemical constituents

including muscle proteins, and regulatory mechanisms. These make the events in smooth muscle contractions considerably more dependent on myosin.

In smooth muscle cells, myofilaments are not organized into myofibrils; sarcoplasmic reticuluns and T-tubules are rather rudimentary, sarcomeres are not aligned to give a striated appearance, and no troponin system is associated with the F-actin strands; instead of troponin, a protein *caldesmon* remains bound to F-action in the relaxing muscle, inhibiting the actin-myosin interaction—this interaction is promoted in the contracting muscle by ht erelease of caldesmon from actin. Moreover, smooth muscle myosin possesses characteristic light peptide chains, called *p-light chains* or *pL-myosin,* different from the L chains of striated muscle myosin; this pL-myosin exists in an inactive phosphorulated from and an active dephospho form, and plays the role of an inhibitor of actin-myosin interaction in place of troponin I. The sarcoplasm also contains **(i)** a *myosin p-light-chain kinase* (ML CK) which catalyzes the ATP-driven phosphorylation and consequent inhibition of pL-myosn, **(ii)** a Ca^{2+} –binding protein *calmodulin* (CaM) whose Ca^{2+}– bound form ($CaM.4Ca^{2+}$) participates in activating MLCK, and **(iii)** a Ca^{2+}– independent *myosin light-chain phosphatase* which can catalyze the dephosphorylation of phospho-pL-myosin into active dephospho-pL-myosin.

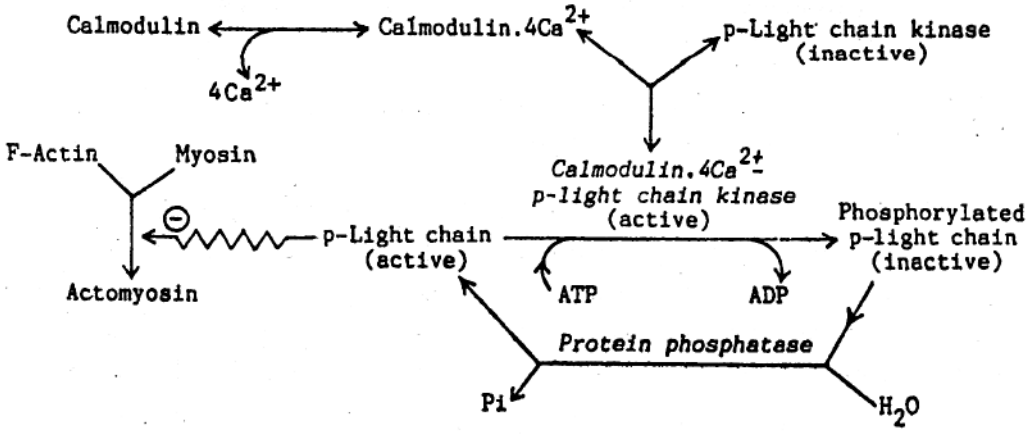


Figure 3.22 : Regulation of smooth muscle contraction by Ca^{2+} ions. [From D. Das, *Biochemistry,* Academic Publishers, 2000]

(a) In the relaxing smooth muscle with a sarcoplasmic Ca^{2+} concentration lower than 10^{-7}M, actin-myosin interaction remains inhibited by the active *dephospho-pL-myosin* as well as the acting-bound *caldesmon.* Antonomic nerve impurses, reaching the smooth muscle cell, depolarizes its plasma membrane; this enhances the Ca^{2+} influx across the latter from the ECF to the sarcoplasm, where *calmodulin* binds to

the inflowing Ca^{2+} to form the CaM.4Ca^{2+} complex (Fig. 3.22). When the sarcoplasmic Ca^{2+} concentration exceeds about 10^{-5} M in this way, optimum amounts of CaM.4Ca^{2+} –molecules bind to the dephospho form of MLCK to activate the latter. The active CaM.4Ca^{2+} -MLCK now catalyzes an ATP-driven phosphorylation of active dephospho-pL-myosin into inactive phosphorylated pL-myosin which can not longer inhibit the actin-myosin interaction. Simultaneously, the CaM.4Ca^{2+} complex may bind to the actin-bound caldesmon and dislodge it from actin. Thus actin-myosin interaction is now freed from the inhibitory effects of both pL-myosin and caldesmon. This leads to the attachment of myosin heads to actin filaments and the sliding of myosin and actin filaments towards each other for muscle contraction.

(b) With the onset relaxation, membrane C^{2+} *–ATPase* actively extrudes Ca^{2+} from the sarcoplasm to lower the sarcoplasmic Ca^{2+} concentration below 10^{-7} M. This releases Ca^{2+} from sarcoplasmic CaM.4Ca^{2+} complex, changing it to calmodulin. **(i)** The lowering of sarcoplasmic CaM.4Ca^{2+} concentration allows the actin-inhibiting protein caldesmon to bind again to actin, thereby inhibiting the actin-myosin interaction. **(ii)** Having lost its Ca^{2+}, calmodulin is released from its binding with MLCK which is thereby inactivated and unable to phosphorylate and to inactive pL-myosin. **(iii)** The Ca^{2+} independent *myosin light chain phosphatase* dephosphorylates and activates the remaining phospho-pL-myosin. **(iv)** The rise in active phospho-pL-myosin due to **(ii)** and **(iii)** enhances its inhibitory action on actin-myosin interaction. All these events lead to the detachment and sliding of actin and myosin filaments away from each other.

3.13 Summary

Active transports are transmembrane transfers fo specific molecules or ions by some membrane proteins, acting as specific transporters or carriers against the electorchemical gradients of the transpported substrates across the membrane. Such transports are endergonic and must be carried out by being coupled with exergonic reactions such as either the exergonic hydrolysis of high energy ATP-bonds by Na^+ – k^+ ATPase or the exergonic dissipation of transmembrane ion-gradients as in cases of mitochondrial ATP-ADP translocase and intestinal Na^+-glucose symport. Active transports may be uniports, symports or antiports, according respectively to whether a single substrate, or two substrates in the same direction or two substrates in opposite directions are transported against their respective electrochemical gradients. Active transport is unidirectional in a specific direction, sigmoid kinetics in some cases of symport and antiport.

In ureofelic animals, the nitrogenous waste-products of amino acid catabolism are converted to urea as the ultimate urinary NPN end product in the hepatic arginine-urea cycle spending four ATP bonds per urea molecule formed.

Glucose absorbed from dietary carbohydrates in the intestine is converted by glycogenesis in liver, muscle and many other tissues to glycogen for storage, spending two high energy ATP bonds per glucose molecule incorporated in glycogen.

Noncarbohydrates, particularly many amino acids and products of glycolysis such as lactate and pyruvate may be converted back to glucose by gluconeogensis in liver and kidneys, Gluconeogenesis from pyruvate or lactate requires the spending of six high-energy phosphate bonds per momecule of glucose synthesized.

Gluathione, a tripeptide acting as an important reducing agent, is synthesized from glutamate, cysteine and glycine at the cost of two ATP bonds in the cytoplasm, but without the participation of polysomes because one of its peptide bonds is a γ-peptide bond instead of an α-peptide bond. Norepinephrine and epinephrine are synthesized in adrenergic neurons and adrenal medulla from the amino acid tyrosine. Tryptophan is used in synthesizing serotonin in serotonergic neurons and melatonin in pineal gland cells.

Saturated fatty acids, plamitic acid in particular, are synthesized in liver, adipocytes and mammocytes by two cytoplasmic multienzyme proteins, viz., acetyl-CoA carboxylase and fatty acid synthase, using acetyl-CoA and NADPH. Such synthesis of a molecule of palmitate involves the expenditure of 15 high energy phosphate bonds. Usually, longer saturated fatty acids are synthesized from palmitate by the stepwise additions of C_2 units with the help of microsomal fatty acid elongase system of enzymes. C_{20}-polyenoic fatty acids, released by phospholipase action on membrane phospholipids, are used in synthesizing prostaglandins by the cyclo-oxygenase pathway with the participation of a midrosomal multienzyme system called the prostaglandin endoperoxide synthase.

According to the sliding filament model, muscle contraction results from the sliding-in of thin actin filaments between thick myosin filaments interdigitating with each other in myofibrils. In striated muscle fibres, the molecular mechanism of muscle contractiion involves and elaborate actin-myosin interaction. In monstriated muscles, contraction is more myosin-based. These mechanisms have been discussed in some detail.

3.11 Terminal questions

1. (a) Give an account of the elongation of palmitate to stearate by smooth ER enzymes.
   (b) Describe how the following specialized products are synthesized from amino acids : (i) glutathione, (ii) epinephrine, (iii) melatonin.
2. (a) Differentiate between active and passive carrier-medialted transports.
   (b) Discuss the characteristics of active transports, including their kinetics.
   (c) Describe the significances of active transports.
3. (a) Describe the synthesis of glucose from lactate, using a suitable flow chart.
   (b) State the significances of gluconeogensis.
   (c) Using a suitable diagram, describe the enzymatic reactions of the principal steps of the cyclo-oxygenase pathway for synthesizing PGD_2- and $PGF_{2\alpha}$ from C_{20}-polyenic acids of membrane.
4. (a) Describe how acetyl-CoA and NADPH are made availabel for fatty acid biosynthesis in the cytoplasm of ruminant na dnonruminant animals.
   (b) Describe a suitable diagram the action of acetyl-AoA carboxylase for palmitate synthesis.
   (c) Give an account of the domains of multienzyme fatty acid synthase of mammals and the enzymatic and nonenzymatic sites in thsoe domains.
   (d) Discuss the short-term and long-term regulations of fatty acid synthesis.
5. (a) What are the significnaces of urea biosynthesis in animals?
   (b) Describe the arginine urea pathway of urea biosynthesis, using a suitable flowchart and mentioning the energetics of the process.
   (c) Discuss the regulation of urea biosynthesis.
6. (a) Describe different active sites in the domains of multienzyme fatty acid synthase.
   (b) With a suitable flow chart discuss the reactions of palmitate synthesis by the enzymic components of fatty acid synthase.
   (c) Mention the energetics of palmitate synthesis.
7. (a) Discuss how glycogenesis is regulated.
   (b) Describe using suitable diagrams the process of glycogenesis, mentioning its energetics.
   (c) How is glycogenesis regulated?

8. (a) Mention different types of ATPases involved in ATP-driven active transports.

   (b) Describe with a suitable flow-chart the action of the $Na^+ - K^+$ ATPase as an ATP-driven active antiport.

   (c) Discuss the role of ATP-ADP translocator in ion gradient-driven active transport.

9. (a) Describe Na^+ -glucose symport with a diagram, as an example of ion gradient driven active transport.

   (b) Describe the sliding filament model of muscle contraction.

10. Discuss how actin-myosin interaction results in the contraction and relaxation of striated muscle fibres.

11. (a) Describe with a flowchart the myosin-based contractions of smooth muscles, regulated by Ca^{2+} ions.

    (b) Write briefly about the following two muscle proteins : (i) troponin system; (ii) myosin.

3.15 Answers

1. (a) See Section 3.10.2.

   (b) See Sections 3.7-3.9.

2. (a) See part (b) of Section 3.2.

   (b) See Section 3.3.1.

   (c) See 3.3.2.

3. (a) Section 3.6.2.

   (b) See Section 3.6.1.

   (c) See Section 3.11.1.

4. (a) See parts (a) and (b) of Section 3.10.1.

   (b) See part (c) of Section 3.10.1.

   (c) See part (d) of Section 3.10.1.

   (d) See part (g) of Section 3.10.1.

5. (a) See relevant part of Section 3.4.

   (b) See relevant parts of Section 3.4.

   (c) See relevant part of Section 3.4.

6. (a) See part (d) of Section 3.10.1.
   (b) See part (e) of Section 3.10.1.
   (c) See part (f) of Section 3.10.1.
7. (a) See relevant part of Section 3.5.
   (b) See relevant parts of Section 3.5.
   (c) See relevant part of Section 3.5.
8. (a) See first paragraph of Section 3.3.4.
   (b) See part (a) of Section 3.3.4.
   (c) See paragraph (a) of Section 3.3.5.
9. (a) See part (b) of Section 3.3.5.
   (b) See Section 3.12.2.
10. (a) See Section 3.12.3.
    (b) See first, second and third paragraph of Section 3.12.1.

Unit 4 ❐ Enzymes

Structure

4.1 Introduction

You are aware that in living organisms, chemical reactions are generally catalyzed by enzymes which are proteins in nature. In this unit, however, you will know about some RNA molecules also, that function as enzymes. This unit will describe how highly reactive but transient complexes are formed by enzyme-substrate interactions for the catalysis. Each encyme can bind to and change only one or a few substrates; different enzymes may possess different types of such substrate specificity. You will learn here about the frequent existence of more than one protein for catalyzing the same reaction.

Therefore of enzyme-catalyzed reactions may be either rectangular hyperbolic functions or sigmoid functions of the substrate concentration, according as they follow the Michaelis-Menten hyperbolic kinetics or the Hill equation for sigmoid kinetics. These will be elaborated in this unit.

Activities of some enzymes may be regulated by reversible addition and removal of specific groups to/from them through the formation/cleavage of covalent bonds. Some inactive proenzymes are irreversible changed to active enzymes by the hydrolytic removal of a part their paptide chain. This unit will also describe the activation or

inhibition of some enzymes by substances called allosteric modulators, binding to specific sites of the enzyme molecule, distinct from the substrate-binding site. Often a metabolic pathway may be augmented or impeded by modulating in various ways one or more rate-limiting enzymes of theat pathway.

Objectives

On reading this unit, you should be able to :

- Understand the transition state theory of enzyme-sybstrate interaction,
- Describe the models of enzyme-substrate interactions,
- Know about the bonds involved in forming enzyme-sybstrate complexes,
- Explain different types of substrate specificities of enzymes,
- Understand the Michaelis-Menten substrate saturation Kinetics,
- Explain the significance and narrate the characteristics of K_m,
- Describe reversible covalent modifications of enzyme,
- Narrate how some proenzymes undergo irreversible convalent activation,
- Learn about ribozymes and isozymes,
- Define rate-limiting enzymes and their characteristics,
- Describe the characteristics of allosteric modulations,
- Understand the sigmoid saturation kinetics and its changes during allosteric modulations,
- Narrate different models of allosterism.

4.2 Enzyme-Substrate Interaction

Like inorganic catalysts, enzymes bind to their substrates to form transient and highly reactive enzyme-substrate complexes (ES Complexes).

4.2.1 ES Complex formation

The action of an enzyme (E) on its substrate (S) is always initiated by the binding of specific groups of the substrate molecule to the sidechains of specific amino acid residues at the *active site* of the enzyme molecule to form the activated but unstable ES complex. The latter would soon dissociate into the unchanged enzyme and the product (P) formed by the catalyzed change of the substrate. That such ES complexes are formed during enzyme actions, has been supported by (i) the *isolation* of some ES complex of glyceraldehyde 3-phosphate dehydrogenase, (ii) the *electron microscopic demonstration* of some ES complexes like those of nucleic acid

polymerases, (iii) the *X-ray crystallography* of some ES complexes such as that of carboxypeptidase A, and (iv) the *spectroscopic detection* of some ES complexes either by nuclear or electorn magnetic resonance (NMR or EMR) spectroscopy or by fluorescence spectroscopy, e.g., the ES complex of prokaryotic tryptophan synthase.

Formation and fate of ES complexes follow different routes in different types of enzymatic reactions. For example :

(a) In *single-substrate reactions* such as those catalyzed by isomerases in isomerizing a single substrate,

$$E + S \rightleftarrows ES \rightarrow E + P$$

(b) In *bisubstrate reactions* between two substrates (S_1 and S_2), there are the following two alternative routes : **(i)** In *single-displacement bisubstrate reactions* such as those catalyzed by phosphotransferases (e.g., glucokinase) or by purine-dependent dehydrogenases (e.g., lactate dehydrogenase), S_1 and S_2 bind successively to E to form a ternary ES_1S_2 complex of the enzyme with both substrates; the products (P_1 and P_2) produced by the enzyme action are then released successively :

$$E + S_1 \rightleftarrows ES_1; ES_1 + S_2 \leftarrow ES_1S_2 \leftrightarrows EP_1P_2;$$

$$EP_1P_2\ EP_2 \rightarrow + P_1; EP_2 + EP.$$

But in (ii) *double-displacement bisubstrate reactions* such as those catalyzed by transaminases, one substrate (say, S_1) binds to the enzyme singly to form the first ES complex (ES_1) which dissociates to release the first product (P_1) and a modified E′ to form the second ES complex ($E'S_2$) which subsequently releases the second product (P_2) and the original form of E :

$$E + S_1 \rightarrow ES_1 \rightarrow E' + P_1; E' + S_2 \rightarrow E'\ S_2 \rightarrow E + P_2.$$

Bonds involves in the formation of ES complexed may be of several types :

(a) *Noncovalent bonds : (i) hydrogen bonds* between polar groups of the substrate and the enzyme, such as that between a uracil residue of RNA and sidechain OH groups of specific serine and threonine residues of pancreatic RNase; (ii) *ionic bonds* between counterionic polar groups of the substrate and the enzyme, such as that between the sidechain NH_4^+ group of a lysine residue of a histone (substrate) and a sidechain COO^- group of the enzyme trypsin; (iii) *hydrophobic interactions* holding the nonpolar groups of the substrate and the enzyme together, (iv) *van der Waals forces of attraction* between both polar and nonpolar groups, acting as dipoles in the substrate and the enzyme.

(b) *Covalent bonds* such as (i) *thioester or thiothemiacetal bonds* between sidechain SH group at the active site of the enzyme and a carboxyl or an aldehyde group of its substrate, as in the ES complex of glyceraldehyde 3-phosphate and glyceraldehyde 3-phosphate dehydrogenase, and (ii) *Schiff bases* between the sidechain NH_2 group of a basic amino acid residue (e.g., lysine) at the active site of the enzyme and a ketonyl group of the substrate, as in the ES complex of fructose 6-phosphate and transaldolase.

Role of enzyme-substrate complexes :

Formation of the ES complex may help in catalysis in various ways.

(i) The ES complex formation may *lower the energy barrier* for the reaction to be catalyzed, by providing alternative reaction paths requiring much less activation energy (vide 4.2.4).

(ii) It may bring the bond to be changed in the substrate very close to and *in proper alignment* with such groups of the enzyme as would participate in the catalysis.

(iii) It may strain the bond to be changed in the substrate to help in its change.

(iv) It may provide functional acidic, basic or other groups like NH_3^+, COO^-, thiol and phenolic OH groups for accepting or donating specific ions or groups during catalysis.

4.2.2 Active sites of enzymes

Several amino acid residues of the enzyme molecule function together to constitute its *active site.* Such amino acid residues may be situated in the peptide chain of the enzyme at distances considrably apart from each other, but have been brought in close proximity and proper steric relations to each other by the coils and folds of the three-dimensional secondary and tertiary structures of the enzyme molecule. Such higher orders of its molecular structure may also form a nonpolar cleft or crevice amidst the peptide chain coils, which is accessible to substrates possessing specific sizes and three dimensional shapes and contains specific amino acid residues of the enzyme with sidechains capable of binding to specific groups of the substrate; this constitutes the *substrate-binding site* in the active site of the enzyme. Specific amino acid sidechains of the binding site and/or of a nearby *catalytic site,* similarly constituted by the three-dimensional form of the enzyme, may catalyze changes of the relevant covalent bond in the enzyme-bound substrate.

Specific amino acid residues must occur at the active site of the enzyme for its action. Thus, *thiol enzymes* (e.g., papain, glyceraldehyde 3-phosphate dehydrogenase and HMG-CoA reductase), *lysine enzymes* (e.g., transaldolase), and *serine enzymes* (e.g., trypsin, chymotrypsin and acetylcholinesterase) must have respectively a cysteine, a lysine and a serine-residue, each at a specific position on the peptide chain and with a free sidechain, at their active sites.

Evidently, the substrate-affinity and the rate of enzyme action may be augmented or diminished if any change in the higher orders of its molecular structure affect the three-dimensional form of its active site. Moreover, the binding of specific nonsubstrate molecules to such a goup of the active, site of an enzyme, as is essentially required to remain free for the enzyme action, inactivates the enzyme; e.g., the binding of iodoacetate to the cystainyl-SH group at the active site any thiol enzymes inhibits the latter.

4.2.3 Models for enzyme-substrate interaction

Two models have been proposed for enzyme-substrate interactions.

(a) ***Fisher's template or lock and key model :***

Fisher proposed that irrespective of the presence or absence of the substrate, the active site of an enzyme exists in the three-dimensional form fully suitable for catalytic activity, and needs no change in that conformation for binding to and changing the substrate. He thus considered that the active site, even when existing by itself free from any substrate, occurs as a *rigid, pre-shaped template* with such three-dimensiional molecular form, size and groups as would readily fit with those of its substrate, therefore, the latter would bind to the active site wihtout any change in the pre-existing three-dimensional form of that site (Fig. 4.1) |——| the binding of the substrate to the active site could be compared with the fitting of a key into the pre-shaped keyhole of a lock (*lock-and-key model).* This would be followed by the enzyme-catalyzed change of the substrate into the product which is released leaving behind the active site in the same unchanged rigid template-like conformation. This model explained the sterospecificity of enzymes, proposing that only one specific stereoisomer of the substrate, but not its other stereoisomers, can fit into and bind to the pre-existing rigid template-like active

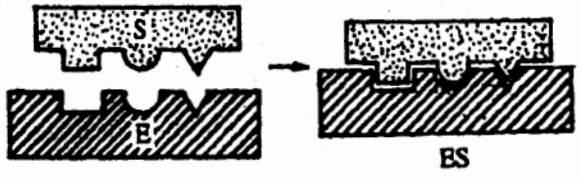


Figure 4.1 : Template model for enzyme-substrate interaction
[From D. Das, Biochemistry, Academic Publishers, 2000]

site. The coenzyme and the substrate would need to bind to the active site in a successive sequence, because the first of such ligands would bind to the rigid template-like active site, providing therby additional groups or binding sites for the binding of the next ligand. Fisher's template model met with no problems in case of the simple form of hyperbolic substrate-saturation kinetics, but could not explain how enzyme-activity changes on the binding of allosteric modulators at sites other than the substrate-binding site if the active site was rigid with no flexibility.

(b) ***Koshaland's induced fit model :***

Koshland proposed that the active site of an enzyme is not a rigid template with its full catalytic activity in its pre-shaped three-dimensional form in which it exists in the absence of the substrate. According to his model, the active site possesses significant *flexibility* and changes its three-dimensional conformation as the substrate molecule comes close to or binds loosely to it to form an initial superficial complex (Fig. 4.2). This initial interaction with substrate-alters the conformation of the active site to a final and much more active form which now binds the substrate more tightly and exerts the full catalytic action on it. This concept of flexibility and substrate-induced change of the active site has been strengthened by the conformational changes detected in many enzymes such as carbomoyl phosphate synthase I (mitochondria) and hexokinase (cytoplasm) on their initial binding to their repective substrates. Moreover, the allosteric modulation of enzymes can be explained by the idea of such conformational changes of the active site when a ligand binds to the allosteric site of such enzymes (vide 4.6.2 and 4.6.3).

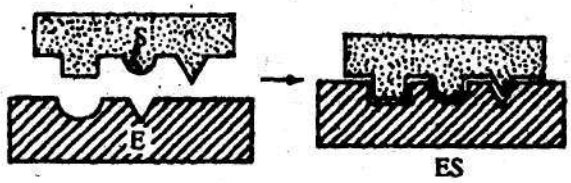


Figure 4.2 : Induced fit model for enzyme-substrate interaction
[From D. Das, Biochemistry, Academic Publishers, 2000]

4.2.4 Transition state theory

For any reaction, the reactants must collide or approach each other with a minimum Kinetic energy called the *activation energy* to overcome the *energy barrier* for forming or cleaving the bonds (*Kinetic theory of reactions).* In any mass of reactants, however, the molecules possess different Kinetic energies, some having the free energy higher than the required activation energy and others falling short of it. A rise of temperature increases the reaction rate by enhancing the kinetic energies of reactant molecules so that a higher number of reactant molecules now have kinetic energies exceeding or equalling the activation energy and can thus overcome the energy barrier for the reaction. According to the *transition state theory* (activated complex theory), the *activation energy* necessary for crossing the energy barrier of a reaction is to be provided by the difference between the free energy of the reactants (A, B) and that of a transient activated intermediate (AB*) formed as a *transition complex* by the initial binding together of the reactants (Fig. 4.3). This transition complex soon dissociates into the products (C, D) if its free energy equals or exceeds the activation energy of the reaction.

$$A + B \rightarrow AB^* \rightarrow C + D$$

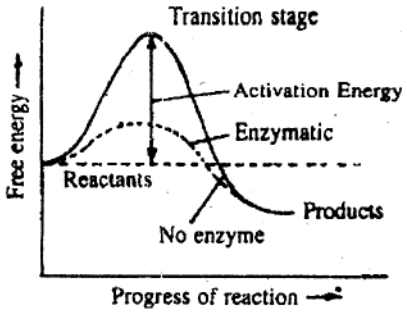


Figure 4.3 : Activation energy of a reaction in presence and absence of an enzyme [From D. Das, Biochemistry, Academic Publishers, 2000]

The transition state theory further proposes that the enzyme catalyzing a reaction binds to the substrate to lower the energy barrier for the reaction probably by providing alternative reaction paths so that much lower activation energy is required. Thus, even at the relatively low temperature in living organisms, the Kinetic energy of a larger number of reactant molecules enables them to form such transition complexes as have free energies equalling or exceeding the lowered activation energy requirement in present of the enzyme, and consequently yield the products at a higher rate.

4.3 Specificities of Enzymes

Each enzyme possesses two types of specificaties, viz., reaction specificity and substrate specificity.

4.3.1 Reaction specificity

Each enzyme catalyzes only a specific type of reaction of its substrate or substraes. For example, creatine phosphokinase, a phosphotransferase, would catalyze only the transfer of a phosphate group from phosphocreatine to ADP or from ATP to creating and no other reaction. Similarly, phosphothexose isomerase, an aldose-ketose isomerase, would carry out only an isomerization reaction between glucose 6-phosphate and fructose 6-phosphate. Because of such reaction specificities, enzymes have been categorized into classes like oxidoreductases, transferases, hydrolases, lyases, isomerases and ligases, according to the specific type of reactions they catalyze.

4.3.2 Substrate specificity

Each enzyme can act upon either a single substrate as a small number of closely related substrates, because the catalytic action of the enzyme requires the substrate to bind to and interact with specific amino acid sidechains and groups at the three-dimensional active site of the enzyme and therefore, to possess specific groups and

three-dimensional form befitting the active structural specifications cannot bind to the active site and consequently cannot be changed by the enzyme. Several types of substrate specificities are described below.

(a) Group specificity :

An enzyme can often act only on a specific type of covalent bond associated or connected with only specific groups or residues in the substrate molecule. For example, gastric pepsin can hydrolyze only such peptide bond which connects the α-COOH group of an aromatic amino acid (Phe, Tyr or Trp) with the α-NH_2 group of another aromatic or dicarboxylic (Glu or Asp) amino acid; intestinal sucrase can hydrolyze such a β-2, 1-glycosidic bond as binds the α-anomeric OH group of C^1 of a glucopyranose the the β-anomeric OH group of C_2 of a fructopyranose. This type of specificity is called the group specificity.]

(b) Chain-length specificity :

Some enzymes can act only on substrates having specific lengths of carbon chains in their molecule. For example, fatty acid thiokinases possess such chain-length specificities for fatty acids; long-chain thiokinases thioesterify only long-chain fatty acids such as palmitic (C_{16}) and stearic (C_{18}) acids, medium-chain thiokinases act on C_4–C_{14} fatty acids, and acetate thiokinase acts similarly on only acetic (C_2) and propionic (C_3) acids.

(c) Geometric or cis-trans sterospecificity :

Some enzymes can act on substrates with double-bonds in their molecular chains, only if the double-bond has a specific steric configuration, either *cis* or *trans*. Such specificity of an enzyme for either *cis* or the *trans* isomer of the substrate, not for both, is called geometric or *cis-trans* stereospecificity. For example, fumarase can act only on the *cis*-isomer, fumaric acid, but not on its *trans*-isomer maleic acid.

(d) D-L stereospecificity

Some enzymes cna act only on the D-stereoisomers of their substrates and not on their L-steroisomers. For example, hexokinases phosphorylate several D-hexoses such ad D-glucose and D-glucosamine, but not their L-stereoisomers. Then, there are other enzymes which act specifically on L-stereoisomers of their substrates, but not on their D-stereoisomers. For example, L-glutamate, but not D-gluamate. Such *D-L stereospecificity* of enzymes indicates that the enzyme and its substrate bind to each other by at least a *three-point binding* so that the enzyme binds to one of the stereoisomers of the substrate; but being a mirror image of that isomer, the other stereoisomer of the latter cannot bind to the active site of the enzyme.

(e) D-l optical specificity :

Enzymes having D-L steresopecificity act also on only one specific optical isomer of the substrate, either its dextrorotatory *(d)* isomer or its levorotary *(l)* isomer, but not on both the optical isomers. This is because the D and L stereoisomers of a substance are also the *d* and *l* optical isomers of each other. For example, L(+)-3-hydroxyacyl-CoA dehydrogenase acts specifically on the dextrorotatory isimers of 3-hydroxacyl-CoA molecules, but not on their levorotatory isomers.

4.4 Michaelis-Menten kinetics

The rate of an enzyme-catalyzed reaction is usually measured and expressed as its *initial velocity* (V_o). The latter is defined as the rate of an enzymatic reaction when very little substrate has yet been converted into the product so that the resultant decline in the molar concentration [S] of the subhstrate has not alterted yet the rate of reaction significantly.

for many enzyme-catalyzed reactions, the initial velocity (V_o) has been found to be a *rectangular hyperbolic function* of the molar concentrations [S] of the substrate or substrates. In other words if the V_0, measured using different molar concentration [S] of a substrate, are plotted against the respective [S] values, the plotted points are found to be distributed along a rectangular hyperbolic curve (Fig. 4.4). This relation between sybstrate concentrations and initial velocities is known as the *Michaelis-Menten rectangular hyperbolic substrate-saturation kinetics* of enzymes. Many of the enzymes catalyzing single-sybstrate, bisubstrate and multisubstrate reactions obey this hyperbolic kinetics. We are, however, restricting ourselves here to single-substrate reactions only, such as those catalyzed by isomerases. A single-substrate reaction, you know, results from the interaction of the enzyme (E) and a single substrate (S).

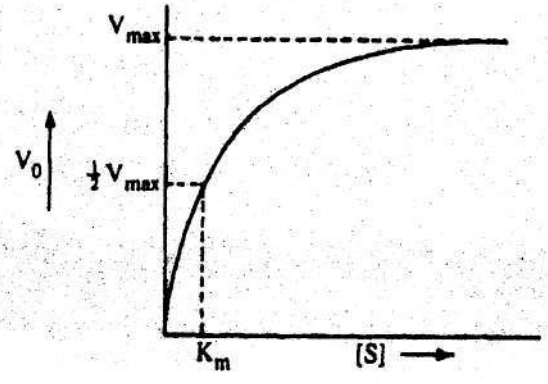


Figure 4.4 : Hyperbolic substrate saturation sainetics
[From D. Das, Biochemistry, Academic Publishers, 2000]

The hyperbolic substrate-saturation plot of a single-substrate reaction shows an *almost linear rise* in the rate of V_0 with the rise in [S] when [S] is low, because at low substrate concentrations, the rises proportionately with the rise in [S]. But at higher substrate concentrations, the *rate of rise in V_o declines progressively* to produce a curved *hyperbolic plot;* this indicates that the enzyme is the *limiting factor* in the enzyme substrate reaction because of a far lower number of enzyme milecules than thsoe of the substrate |——| this leaves less and less numbers of free enzyme molecules available for binding to fresh substrate molecules, with the rise in substrate concentration. Ultimately, at a still higher substrate concentration, the plot reaches a *plateau-like flat summit* because all enzyme molecules are now already saturated with substrate and interact with the substrate at their maximum capacity. The V_o now reaches the maximum level called the *maximum velocity* (V_{max}) which cannot be surpassed by increasing the [S] any further. This hyperbolic kinetics obeys the *Michaelis-Menten equation* which for a single-substrate reaction, may be expressed as follows :

$$V_o = \frac{V_{max}[S]}{K_m + [S]}$$

Where, K_m is the *Michaelis constant* serving as an index of the substrate-affinity, and V_{max} is the maximum rate of enzyme action at saturating substrate concentrations as also an index for the concentration of active enzyme in a tissue.

4.4.1 Michaelis constant

Michaelis constant (K_m) is regularly measured for enzymes to estimate their substrate-affinities, kinetics and their changes due to inhibitions and modulations.

Characteristics of Michaelis constant :

(a) K_m is the molar concentration of the substrate at which the initial velocity (V_o) amounts to half the maximum velocity, i.e., $^1/_2\ V_{max}$. Thus, where $V_o = {}^1/_2\ V_{max}$.

$$V_o = \frac{V_{max}[S]}{K_m + [S]}, \text{ or } K_m = [S]\left(\frac{V_{max}}{V_o} - 1\right)$$

$$\text{or, } K_m = [S]\left(\frac{V_{max}}{\frac{1}{2}V_{max}} - 1\right) = [S]$$

K_m is expressed in moles per litre. It is a constant for each particular enzyme-substrate concentration so long as other conditions like pH, temperature and ionic strengths remain unaltered.

(b) At very low substrate concentrations far below the k_m,

$$[S] \langle\langle K_m, \ \therefore \ K_m + [S] \cong K_m$$

$$\therefore V_o = \frac{V_{max}[S]}{K_m + [S]} = \frac{V_{max}[S]}{K_m} = K[S].$$

Where k is a new constant, being the ratio of V_{max} and k_M for the given enzyme-substrate combination. This indicates that at very low substrate concentrations, V_o of a single-substrate reaction depends only on the molar concentration of the substrate *(a first-order reaction)* and rised lineraly with [S] in the initial part of the hyperbolic plot (Fig. 4.4).

(c) At high substrate concentrations far exceeding the k_m,

$$[S] \rangle\rangle K_m, \ \therefore \ K_m + [S] \cong [S]$$

$$\therefore V_o = \frac{V_{max}[S]}{K_m + [S]} = \frac{V_{max}[S]}{K_m} = V_{max}$$

Thus, when [S] far surpasses the k_m, V_o reaches the V_{max} and is no longer proportional to reactant (substrate) concentration (a **zero-order reaction**). This forms the plateau-like summit of the hyperbolic curve (Fig. 4.4). It follows that with saturating substrate concentrations, usually when $[S] \geq 10^2 \ k_m$, V_{max} depends on the prevailing *molar concentration of the active -enzyme.*

(d) K_m is a measure of the *substrate-affinity* of the enzyme and ranges from about 10^{-7} to about 10^{-1} M. The higher is the K_m, the lower the substrate-affinity; the lower is the k_m, the higher the substrate affinity. This is because the higher is the affinity of enzyme for a substrate, the lower is the substrate concentration required to attain the V_{max} as well as the $\frac{1}{2}V_{max}$.

(e) The substrate-affinity and so, the k_m would vary from substrate to substrate for an enzyme. For example, pyruvate carboxylase, catalyzing a reaction between ATP, pyruvate and HCO_3^- to form oxaloacetate, has progressively higher k_m values for those substrates, viz., 6×10^{-5} M, 4×10^{-4} M and 1×10^{-3} M respectively for them. So, this enzyme possesses its highest affinity for ATP and the lowest affinity for bicarbonate.

(f) K_m as well as substrate-affinity may be changed by any condition that either changes the higher orders of structure of the enzyme protein or alters the ionization of its active site. Thus, K_m may change with changes in temperature, pH and ionic concentration.

(g) K_m of an enzyme is increased by the competitive inhibition of the enzyme, because the competitive or substrate-analogue inhibitor competes with the substrate for binding to the substrate-binding site of the enzyme, decreasing the availability of free substrate-binding sites to the substrate and consequently, diminishing the substrate-

affinity. But K_m remains unchanged in noncompetitive inhibitiion because the noncompetitive inhibitor, not being a substrate-analogue, does not bind to the substrate-binding site of the enzyme and consequently does not affect the substrate-affinity of the latter.

(h) Different proteins, catalyzing the same reaction of identical substrates, are called *isozymes* or *isoenzymes* and differ between themselves in their tissue distributions. Such isozymes of an enzyme differ from each other in their substrate-affinities and so, possess different K_m values for the same substrate. For example, the hexokinase isozyme (glucokinase) of hepatocytes has a hundred-fold higher K_m for glucose than the other hexokinase isozymes of muscles, intestinal cells and brain. Such differences in K_m between isozymes play an important role in conducting the metabolism of a substrate differentially in different tissues.

(i) Allosteric enzymes of the *K-series*, such as phosphofructo-kinase-1, suffer changes of their K_m on the binding of specific ligands (*allosteric modulators*) to their allosteric sites. For example, the binding of some ligands (*activators*) to such sites increases their substrate affinities and consequently decreases the K_m; but the binding of some other ligands *(inhibitors*) decreases the substrate affinity and raises the K_m. On the contrary, the substrate affinity as well as the K_m remains unaltered in M series of allosteric enzymes, such as acetyl-CoA carboxylase, on the binding of allosteric modulators to them.

4.4.2 Linear transformations of Michaelis-Menten equation

For regular measurements of K_m and V_{max} values of enzymes in estimating their substrate-affinities, active enzyme concentrations, and their changes, it is far more convenient to use the following linear forms of the Michaelis-Menten equation than its heperobolic form itself.

Linewaver-Burk double-reciprocal plot :

The Lineweaver-Burk equation is derived from the Michaelis-Menten equation (MME) in the following way :

$$V_o = \frac{V_{max}[S]}{K_m + [S]} \quad or, \quad \frac{1}{V_o} = \frac{K_m + [S]}{V_{max}[S]}$$

$$Or, \quad \frac{1}{V_o} = \frac{K_m}{V_{max}} \times \frac{1}{[S]} + \frac{1}{V_{max}}$$

On scaling $\frac{1}{V_o}$ and $\frac{1}{[S]}$ along the ordinate and the abscissa, respectively, and plotting each experimentally measured $\frac{1}{V_o}$ against the corresponding $\frac{1}{[S]}$, a linear

double-reciprocal plot is obtained (Fig. 4.5). They y-intercept, the negative x-intercept and the slope of the plotted straight line give respectively the $\frac{1}{V_{max}}$, $-\frac{1}{K_m}$ and $\frac{K_m}{V_{max}}$.

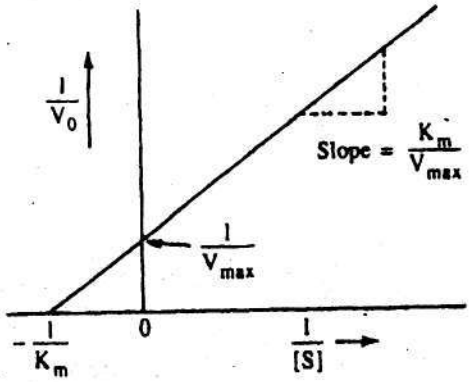


Figure 4.5 : Double reciprocal plot
[From D. Das, Biochemistry, Academic Publishers, 2000]

Eadie-Hofstee plot :

This linear form of the MME is obtained as follows :

$$V_o = \frac{V_{max}[S]}{K_m[S]}, \quad \text{or,} \quad \frac{V_o}{[S]}\left(K_m + [S]\right) = V_{max},$$

$$\frac{V_o}{[S]}K_m + V_o = V_{max}, \therefore\ V_o = V_{max} - \frac{V_o}{[S]}K_m$$

On scaling V_o and $\frac{V_o}{[S]}$ along the ordinate and the abscissa, respectively, and plotting each measure V_o value against the corresponding $\frac{V_o}{[S]}$, a linear *Eadie-*

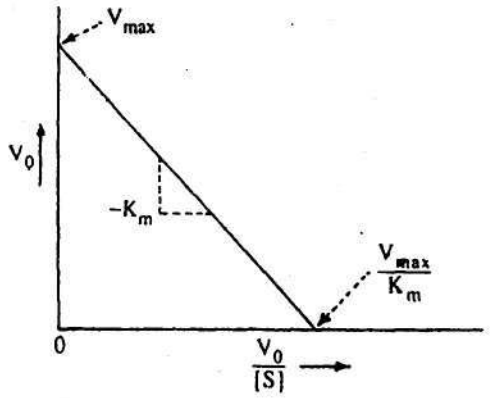


Figure 4.6 : Eadic-Hofstee plot [From D. Das, Biochemistry, Academic Publishers, 2000]

Hofstec plot is obtained (Fig. 4.6). In this plot, they y-intercept, the x-intercept, and the slope give respectively the values of V_{max}, $\frac{V_{max}}{K_m}$ and –Km.

***Wolf-Hanes plot* :**

The linear Wolf-Hanes equation results on multiplying both sides of the Lineweaver Burk equation by [S] :

$$\frac{[S]}{V_o} = \frac{K_m}{V_{max}} + [S] \times \frac{1}{V_{max}}$$

A linear *Wolf-Hames plot* is drawn by scaling $\frac{[S]}{V_o}$ and [S] along the ordinate and the abscissa, respectively, and plotting each measured $\frac{[S]}{V_o}$ against the corresponding [S] value (Fig. 4.7). They intercept, the negative x-intercept and the slope of this linear plot gives respectively the $\frac{K_m}{V_{max}}$, the $-K_m$ and the $\frac{1}{V_{max}}$ values.

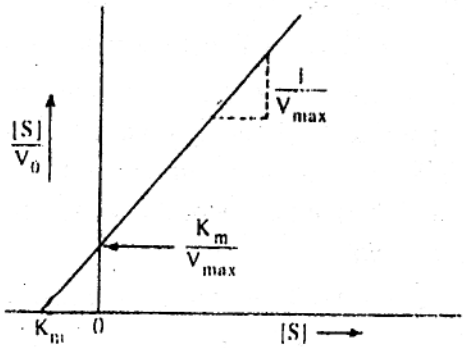


Figure 4.7 : Wolf-Hanes plot [From D. Das, Biochemistry, Academic Publishers, 2000]

4.5 Covalent modifications of enzymes

Covalent modifications of enzymes consist of either addition or removal of specific groups to/from the enzyme molecule by respectively the formation and cleavage of a covalent bond, leading to either an increase or a decrease in the enzyme activity. Covalent modifications may be either irreversible or reversible. They are important mechanisms of *in vivo* regulations of enzyme activities.

4.5.1 Irreversible covalent activation

Stronge protein-hydrolyzing enzymes are in many cases synthesized and secreated as inactive precursor molecules called *proenzymes,* so as to prevent their unintended

action on body-tissue components. These inactive proenzymes reach their intended sites of action in the tissues where specific peptide bonds of their molecules get hydrolyzed under appropriate conditions like pH changes or by the actions of other specific proteases. That results in the removal and release of specific segments of the peptide chain of the proenzyme as one or more inactive peptides called *unmasking substances;* the remaining part of the proenzyme molecule becomes an active enzyme due to consequent donformational changes in its molecule and exposure of specific substrate-binding or catalytic groups at its active site. This change is *irreversible* |——| the activated enzyme, once formed, cannot bind again with the released inactive peptide to form back the original proenzyme. Two examples are cited below :

(a) The gastric protease, pepsin, is secreted in the gastric juice of many vertebrates, including mammals, as an inactive proenzyme called *pepsinogen.* In the gastric lumen, a specific peptide bond of pepsinogen is hydrolyzed either spontaneously by the strong acidic gastric pH, or by the autocatalytic action of already activated pepsin molecules at a pH of ~4.6. This releases an inactive polypeptide of 44 amino acid residues from pepsinogen, changing the latter to pepsin.

$$\text{Pepsinogen} + H_2O \rightarrow \text{Pepsin} + \text{inactive polypeptide.}$$

(b) Similarly, the pancreatic inactive protease *trypsinogen* has to reach the intestinal lument where a specific peptide bond is hydrolyzed either by enteropeptidase of intestinal juice at a pH of 5.2 to 6.0, or by already activated trypsin molecules at a pH near 7.9; this releases an inactive hexapeptide from trypsinogen to change the rest of its molecule to active *trypsin.*

$$\text{Trypsinogen} + H_2O \rightarrow \text{trypsin} + \text{inactive hexapeptide.}$$

4.5.2 Reversible covalent modification

Some enzymes get activated or inhibited by enzyme-catalyzed addition of specific groups such as phosphate and adenylate groups, through the formation of covalent bonds between the target enzyme molecules and the group added. Moreover such a process is reversible |——| the opposite change of the same covalent bond, catalyzed by some other enzyme, results in the opposite effect on the target enzyme activity. Two examples of such reversible covalent modifications are cited below.

(a) *Glycogen phosphorylase* of hepatocytes is activated and inhibited by reversible covalent modifications through phosphorylatiion and dephosphorylation, respectively. *Phosphorylase Kinase* used a high-energy phosphate group of ATP to phosphorylate a specific serine residue of the inactive phosphorylase *(dephosphophosphorylase),* changing it to the active phosphorylase *(phosphophosphorylase).*

$$\text{Dephosphosphosphorylase} + \text{ATP} \rightarrow \text{phosphophosphorylase} + \text{ADP}$$

On the contrary, *protein phosphatase-I* hydrolyzes the phosphoester bond in phosphophosphorylase to release Pi and inactive dephosphophosphorylase.

$$\text{Phosphophosphorylase} + H_2O \rightarrow \text{dephosphophosphorylase} + Pi$$

(b) Active *glycogen synthase* a also undergoes reversible covalent modification; however, phosphorylation inhibits it and dephosphorylation activates it. Active glycogen synthase *a* is inactivated to *glycogen synthase b* by phosphorylation catalyzed by *protein Kinase a* using an ATP.

$$\text{Glycogen synthase } a + ATP \rightarrow \text{glycogen synthase } b + ADP$$

On the contrary, *protein phosphatase-I* hydrolyzes the phosphoester bond in glycogen synthase *b* to dephosphory late it to active glycogen synthase *a,* releasing the phosphate group as Pi.

4.6 Allosteric modulations of enzymes

In addition to the substrate binding site present in the enzymes, some enzymes possess also one or more other specific sites called *allosteric sites* which play important roles in the regulation of such enzymes. Binding of specific low-MW ligands to such allosteric sites either enhances of decreases the activity of these enzymes. Such low-MW ligands, binding with allosteric sites to regulate these enzymes, are known as *allosteric modulators* and their regulatory actions on the enzymes are called *allosteric modulations.* According as such modulators enhance or decrease the enzyme activity, they are respectively called positive allosteric modulators (*allosteric activators*) and negative allosteric modulators *(allosteric inhibitors).*

4.6.1 Characteristics of allosteric modulations

(a) The allosteric site and the active or substrate-binding site *(isosteric site*) are located on different peptide subunits of an oligomeric allosteric enzyme.

(b) Changes in pH and temperature, high ionic concentrations, mercurials, radiations, mutarotations and peptidase actions frequently affect or destroy the allosteric property of an enzyme without affecting its catalytic activity. This indicates that the allosteric and the substrate binding sites of an enzyme are not identical.

(c) Denaturation of the active site of an enzyme is sometimes prevented by the presence of the allosteric modulator, but not by the presence of the substrate itself; this also indicates that the allosteric and the isosteric sites of an enzyme are distinct from each other.

(d) Only specific ligands can bind to the allosteric site to about the allosteric modulation of the enzyme (*modulator-specificity* of allosteric sites).

(e) Some enzymes possess allosteric capable to binding to only one type of modulators, either activators or inhibitors, and do not possess sites for binding to the opposite type of modulators. But some others bear separate allosteric sites, one binding to activators and the other to inhibitors. For example, cytoplasmic carbamoyl-phosphate synthase II possesses both positive and negative allosteric sites which respectively bind to PP-ribose-P and UMP; but mitochondrial carbamoylphosphate synthase I bears only a positive allosteric site where N-acetylglutamate binds to activate the enzyme allosterically, while allosteric inhibition of this enzyme has not been demonstrated.

(f) In some cases, either the substrate or an intermediate of an earlier step of a pathway binds to the positive allosteric site of an enzyme for a subsequent step to activate the latter allosterically *(feed-forward allosteric activation).* An example is the allosteric activation of pyruvate Kinase of glycolysis by fructose 1,6-bisphosphate produced by an earlier step of the same pathway.

(g) Sometimes, a metabolic product of a pathway binds to the negative allosteric site of an enzyme catalyzing one of its earlier steps and inhibit that enzyme allosterically *(product feedback allosteric inhibition).* For example, an end product UMP of the pyrimidine synthesis pathway allosterically inhibits carbamoyl-phosphate synthase II, an initial enzyme of that pathway.

(h) Allosterism results from different types of cooperativities between the ligands binding to an allosteric enzyme. Allosteric inhibitors have a *negative cooperativity* with the substrate of the enzyme; in other words, on binding to the allosteric site, an allosteric inhibitor decreases either the binding of the substrate to the substrate-binding site or the catalytic action of the enzyme on the substrate. On the contrary, on binding to the allosteric site, allosteric activators show a *positive cooperatively* with the substrate, enhancing the enzyme-substrate interaction and catalysis. As the allosteric ligand is different from the substrate, such cooperativity between the two is called *heterotropic allosteric effect.*

(i) According to the nature of change in the allosteric enzyme Kinetics by the heterotropic allosteric effects of modulators, allosteric enzymes are categorized into a K-series (e.g. PFK-1) and a *V or M series* (e.g. acetyl-CoA carboxylase). Allosteric ligands change the substrate-affinities (and so, the K_m values) of K-series of enzymes, but alter the V_{max} of V or M series of enzymes.

(j) The peptide subunits of an allosteric enzyme bear more than one substrate binding site and so, several substrate molecules may successively bind to the substrate-binding sites of an enzyme. This produces the *homotropic effect* of substrates of an

allosteric enzymes due to a *positive cooperativity* between these substrate molecules |——| the binding of one substrate molecule to a substrate-binding site of the enzyme brings about a change in three-dimensional conformation of the latter, promoting the binding of a second substrate molecule to the next substrate-binding site of the same enzyme molecule, with similar repetitions for successive substrate molecules.

(k) Allosteric enzymes obey the *sigmoid substrate-saturation Kinetics,* instead of the Michaelis-Menten hyperbolic Kinetics of nonallosteric enzymes (vide 4.6.2.).

4.6.2 Sigmoid Kenitics of allosteric enzymes

Because of the homotropic effect of substrate molecules of an allosteric enzyme, resulting from the positive cooperativity between them, the initial velocity (V_o) for such an enzyme action is a *sigmoid function* of the molar concentration [S] of the substrate. This *sigmoid substrate saturation Kinetics* conforms to the *Hill equation* given below for a single-substrate reaction :

$$V_o = \frac{V_{max}[S]^n}{K' + [S]^n}$$

where V_o is the initial velocity with a substrate concentration [S], V_{max} is the maximum velocity at the saturating substrate concentration K' is a constant which is distinct from K_m and n is the *Hill coefficient,* serving as a measure of positive cooperativity between the substrate molecules, and dependent on the number of substrate-binding sites of each enzyme molecule as well as the nature and strength of interaction between those sites. On plotting the V_o against the corresponding substrate concentrations, a *sigmoid curve* is produced (Fig. 4.8). If n exceeds 1, a positive cooperativity between the ligands, viz., substrate-substrate or substrate-activator cooperativity, promotes the binding and interaction of the enzyme with more substrate molecules; when n is lower than 1, a negative substrate-inhibitor cooperativity decreases the binding and interaction of substrate molecules with the enzyme; in the absence of any cooperativity between the ligands, n equals 1 and the Kinetics is hyperbolic instead of sigmoid.

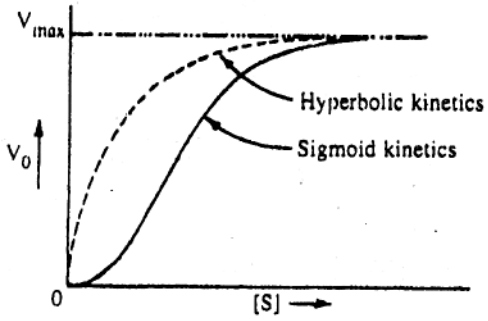


Figure 4.8 : Sigmoid and hyperbolic kinetics [From D. Das, Biochemistry, Academic Publishers, 2000]

The sigmoid plot indicates that Vo rises at a low rate with the rise of [S] at low substrate concentrations; with further rise of [S], V_o rises more steeply with a higher slope due to positive cooperativity between substrate molecules; as [S] approaches the saturating concentration, the curve reaches its flat top summit (V_{max}) because of the nonavailability of free enzyme molecules for interaction with more fresh substrate molecules.

Changes during allosteric modulations :

(a) In *K-series* of enzymes such as PFK-1, binding of an allosteric activator to the positive allosteric site of the enzyme shifts the sigmoid curve to the left making it steeper, and *lowers the* K_m (Fig. 4.9), but leaves the summit and the V_{max} *unchanged.* Binding of an allosteric inhibitor to the negative allosteric site shifts the sigmoid curve to the right making it flatter, and *raises the* K_m, but leaves its summit and V_{max} *unchanged.*

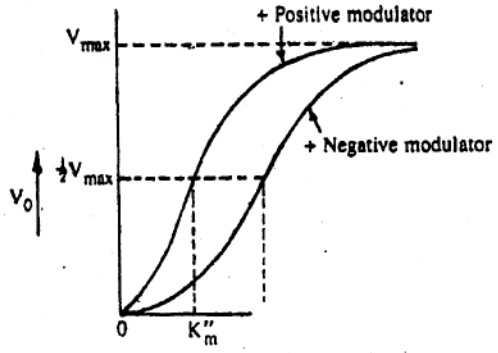


Figure 4.9 : Sigmoid curves of K-series in presence of allosteric modulators
[From D. Das, Biochemistry, Academic Publishers, 2000]

(b) In *M or V series* of enzymes such as acetyle-CoA carboxylase, an allosteric activator shifts the curve to the left and raises its slope and summit, *enhancing the* V_{max} but *not chaniging the* K_m. An allosteric inhibitor shifts the curve to the right, lowers the summit as well as the slope, decreasing the V_{max} but keeping the K_m unaltered (Fig. 4.10).

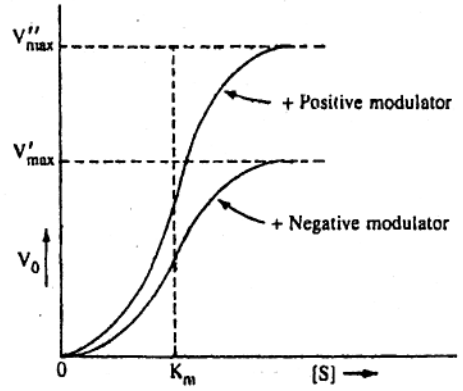


Figure 4.10 : Sigmoid curves of M-series in presence of allosteric modulators
[From D. Das, Biochemistry, Academic Publishers, 2000]

4.6.3 Models of allosterism

Two models have been proposed to explain the molecular basis of allosteric regulation.

(a) *Monod's indirect concerted model :*

This model has been based on the *law of mass action.* Here, an allosteric enzyme is considered to be existing in two molecular forms in a dynamic equilibrium with one another. *(i)* The catalytically *active R form* of the enzyme presents optimally binding sites for the substrate and the allosteric activator; *(ii)* the *inactive T form* has sites binding optimally to the allosteric inhibitor, but is poor in optimal substrate-binding sites.

(i) If a substrate [S] binds to an optimal substrate-binding site of the R form to produce the RS complex, the equilibrium between free R and T forms is disturbed due to a decline in the concentration of free R molecules. To regain that equilibrium, some T molecules change into R molecules, increasing the availability of optimal substrate-binding sites and allowing more substrate molecules to bind to R molecules, consequently enhancing the enzyme sybstrate interaction. This may explain the *homotropic effect* of substrates due to *positive cooperativity* between them.

(ii) The concentration of free R form may also be lowered and the R : T equilibrium may be distrubed on the binding of an activator (A) to the optimal activator sites of R molecules to change them to the RA complex. To restore the R : T equilibrium, some T molecules now change into the R form, increasing the availability of optimal substrate-binding sites and the enzyme-substrate interaction. Thus, a *positive cooperativity* between substrate and activator *results in a heterotropic allosteric activation.*

(iii) An inhibitor (I), on the contrary, binds to the optimal inhibitor-binding site of T form of the enzyme, lowering the concentration of free T molecules and therby disturbing the R : T equilibrium. To restore the latter, some R molecules change into the T form, decreasing the availability of optimal substrate-binding sites and consequently lowering the rate of enzyme-substrate interactin. Thus, a *negative cooperativity* between the substrate and the inhibitor brings about a *heterotropic allosteric inhibition.*

(b) Koshland's direct sequential model :

Instead of assuming two distinct R and T forms of allosteric enzymes, Koshland has based his model on the concept of a *flexible conformation of enzyme molecules,* just as he proposed his induced fit model of enzyme-sybstrate interaction. He proposes that an allosteric enzyme molecule has only *partially accessible binding sites* for

substrates and allosteric modulators, and its three-midensional conformation changes on the initial binding of any of these ligands to the respective partially accessile binding sites, resulting in an altered accessibility of the substrate to the substrate-binding site.

(i) When the substrate binds initially to the partially accessible substrate-binding site of a peptide subunit of an enzyme molecule, conformational changes occur first in that subunit, but subsequently spread sequentially over other subunits also, making the partially accessible substrate binding sites of successive subunits more accessible to the substrate. So, progressively more substrate molecules, *bind sequentially* to those binding sites, leading to increased enzyme-substrate interaction and a *homotropic activation* of the enzyme in a *domino-type manner.*

(ii) Binding of an allosteric activator to the partially accessible binding site of a subunit of an enzyme molecule similarly changes the three-dimensional conformation of that subunit, which spreads over the other subunits in a domino-type effect to increase the accessibility of the substrate-binding sites on successive peptide units. This causes a *heterotropic allosteric activation* by increasing the enzyme substrate interaction in successive units of the enzyme molecule.

(iii) On the contrary, binding of an allosteric inhibitor to its partially accessible binding site on one subunit of an enzyme molecule produces such conformational changes, first in that subunit and then spreading sequentially to other subunits of the enzyme molecule, as to decrease the accessibility of its successive subunits to the substrate molecules in a domino-type effect. This brings about the *heterotropic allsoteric inhibition* of the enzyme.

4.7 Isozymes

In many cases, the same enzymic catalysis of an identical reactiion of the same subtrates is carried out by more than one different protein. Such enzymes are called *isozymes* or *isoenzymes;* e.g., lactate dehydrogenase isozymes of chicken cariac and striated muscles, liver and bhrain; alkaline phosphatase isozymes of mouse small intestine; hexokinase isozymes of liver, brain and muscle; aldolase isozymes of muscle, liver and brain. For example, different lactate dehydrogenase isozymes catalyze the oxidation of lactate to pyruvate as well as the reduction of pyruvate in different tissues. Different isozymes are differentially distributed in different tissues and at different locations in the same cell. For example, hexokinase I, II and III occur mainly in extrahepatic tissues which hexokinase IV (glucokinase) occurs almost solely in the liver. Isozymes of the same enzyme differ in their physicochemical properties such as the types and combinations of peptide subunits in their molecules,

MW, sedimentation co-efficients, isolectric pH, thermolability and chemolability; so, they may be separated from each other by methods of protein separatiion such as ion-exchange chromatography, gel electrophoresis and isoelectrophoresis. They also differ in biological properties such as substrate-specificaties, substrate-affinities, K_m, optimum pH and optimum temperature, in the regulations by allosteric modulations, induction and repression, as also in their distributions in tissues. For example, hexokinase IV (glucokinase) of hepatocytes is far more substrate-specific for glucose but has a higher K_m than hexokinases I, II and III of extrahepatic tissues; lactate dehydrogenase (LDH) isozymes of different tissues are tetramers of H and M peptide subunits in different combinations; acid phosphatase isozymes of bones, liver and pancreas are ethanol-stable and tartrate-stable, while the isozyme of prostate and prostatic fluids is both tartrate-labile and elthanol-labile.

Differences between isozymes in their tissue distributions, kinetic properties and regulations enable the living organism to regulate metabolic activities differentially in different tissues, according to their metabolic needs. For example, hexokinases I, II and III of extrapatic tissues can continue their actions and help in glucose uptake by the tissues even when the blood sugar is low; but hexokinase IV of hepatocytes, because of its far higher K_m for glucose, suspends its action on glucose and thus keeps hepatic glucose uptake on a hold whenever the blood sugar declines.

4.8 Ribozymes

Ribozymes are *RNA molecules* with enzyme activities. They have been discovered in a few unicellular organisms, such as the *RNase P* of the bacterium *Escherichium coli,* and the intron segment of the pre-rRNA transcript of the protistan *Tetrahymena thermophila.* Some ribozyme molecules are made of both RNA and protein parts, but even then have their catalytic activity in the RNA part and not in the protein, and continue to function as an enzyme even after being isolated from the protein part. During post-transcriptional processing of some pre-tRNA and pre-rRNA transcripts of the organisms mentioned above, ribozymes have been shown to cleave the transcripts at specific sites to remove from the latter some inactive oligonucleotide segments as would not participate in the subsquent protein translation; the segments of the RNA transcript, flanking the cleaved segment on its two sides, are then *spliced* (joined) together to from the functional RNA molecule. The intron of the pre-rRNA transcript of *T. thermophila* is a ribozyme which follows the Michaelis-Menten hyperbolic substrate saturatiion Kinetics.

4.9 Rate-Limiting Enzymes

A *rate-limiting or commited enzyme* is such an enzyme of a metabolic pathway having more than one reaction step, as would determine the overall rate and direction of that pathway. In other words, the rate of a metabolic pathway may be regulated by regulating the activity of its rate-limiting enzyme by processes such as allosteric midulations, reversible covalent modifications, feedback inhibitions and competitive inhibitions, or by regulating the syntehsis of that enzyme by induction and repression. Rate limiting enzymes should have the following characteristics.

(a) The rate-limiting enzyme should have the *highest* K_m compared to the K_m values of other enzymes of that pathway. In other words, at saturating substrate concentrations for all those enzymes, the velocity of the committed step, catalyzed by the rate-limiting one, would be the lowest because of its lowest substrate affinity.

(b) The rate-limiting enzyme should catalyze the reaction of an *early step* in the metabolic pathway so that by regulating it, the rate of almost the entire pathway may be controlled, avoiding the unnecessary wastage of reactants and the unwanted accumulation of intermediates of preceeding steps.

(c) The rate-limiting enzyme should catalyze the reaction of the committed step in a *single specific direction* only and thus determine the overall direction of the pathway. The reverse reaction of that step should be carried out by a separate enzyme only.

(d) The rate-limiting enzyme has to be a *regulated enzyme* so that the rate of the metabolic pathway may be controlled by the activation inhibition, induction or repression of that enzyme.

Some rate-limiting enzymes are mentiioned below with the respective relevant pathways within parentheses : *(i)* mitochondrial carbamoyl phosphate synthase I (urea biosynthesis), *(ii)* cytoplasmic carbamoyl phosphate synthase II (pyrimidine biosynthesis), *(iii)* glucose 6-phosphate dehydrogenase (pentose phosphate pathway), *(iv)* phosphofructokinase-1 (glycolysis), (v) α-Ketoglutarate dehydrogenase (TCA cycle), *(vi)* glycogen synthase (glycogenesis), *(vii)* PEP Carboxykinase (gluconeogenesis), *(viii)* acetyl-CoA Carboxylas (fatty acid biosynthesis); *(ix)* HMG-CoA reductase (sterol synthesis); *(x)* AmLev synthase (porphyrin synthesis).

4.10 Summary

The active site of enzyme is constituted by some of its amino acid residues brought close together and in proper steric relations to each other by the coiled higher orders of structure of the enzyme molecule. It has a nonpolar cleft, accessible to the

substrate and containing specific amino acid sidechains forming the substrate-binding site. Specific amino acid residues at the active site must bear free sidechains for binding to and changing the substrate; for example, serine enzyme such as trypsin must possess a specific serine residue at the active site with its sidechain free.

The substrate binds to the substrate-binding site of the enzyme to form a highly reactive enzyme-sybstrate complex by noncovalent ionic, hydrogen, hydrophobic and van der Waals bonds or by covalent bonds such as thioester, thiohemiacetal and Schiff base bonds. The formation of the ES complex helps in catalysis by providing alternative reaction paths requiring much lower activation energy for the reaction. Fisher's template model proposes that the active site of the enzyme has a preformed, rigid, template-like conformation and readily fits without any change to the substrate, binding to the latter and catalyzing its change. But Koshaland proposes in his induced fit model that the active site is somewhat flexible |——| the initial loose binding of the substrate to that site induces its change into a more active and fully catalytic form.

Each enzyme can catalyze only a specific type of reaction of its substrate. Moreover, each enzyme can bind to and act on only one or a few substrate(s) bearing specific three-dimensional structural forms, specific groups and bonds, specific chain-lengths or specific configurations of double-bonds. Accordingly, enzymes show D-L stereospecificity, group specificity, **cis-trans** specificity, chan-length specificity and **d-l** optical specificity.

The initial velocity (V_o) of an enzymatic reaction initially rises linearly with the molar concentration [S] of the substrate, but its rate of rise declines progressively with subsequent further rises in [S], and ultimately reaches a maximum velocity (V_{max}), not surpassed by any further rise in [S]. Plotting the V_o against the corresponding [S] gives a rectangular hyperbolic curve obeying the Mischaelis-Menten equation, known as the rectangular hyperbolic substrate saturation kinetics; for a single-substrate reactiion, $V_o = (V_{max}\ [S]\ [S]/\ (K_m + [S])$, where K_m is the Michaelis constant. K_m is defined as that molar concentration of the substrate at which .... V_{max} is reached. K_m is a measure of the substrate-affinity of the enzyme |——| the lower the K_m, the higher is the substrate-affinity, while the higher the K_m, the lower is the substrate affinity. K_m varies from substrate to substrate for the same enzyme, and also from isozyme of the same enzyme. K_m is increased by competitive inhibition. It is also changed by allosteric modulatiions of K-series of allosteric enzymes. Three linear transformations of the Michel-Menten hyperbolic plot, viz., Lineweaver-Burk plot, Eadie-Hofstee plot and Wolf-Hanes plot, have also been described in this unit.

Specific proteases of pH changes may cause irreversible covalent activations of some inactive proenzymes into active proteases by the hydrolytic removal of specific

segments of their peptide chains. Some enzymes are reversible activated or inhibited by enzyme-catalyzed addition or removal of phosphate or adenylate groups.

Some enzymes undergo activation or inhibitiion when specific ligands (allosteric modulators) bind to their allosteric sites, distinct from their substrate binding sites. Such allosteric enzymes obey Hill's sigmoid saturatin Kinetics : $V_o = (V_{max}\ [S]^n) / (K' + [S]^n)$ where K′ is a constant and n is the Hill coefficient serving as a measure of cooperativity between different ligands of the enzyme. Plotting V_o against [S] gives a sigmoid curve indicating such cooperativity between ligands. Allosteric enzymes bind to more than one substrate molecule on as many substrate-binding sites |——| binding of each substrate molecule promotes the binding of subsequent substrate molecules and their catalytic changes due to positive cooperativity between the substrates (homotropic effect). Binding of an allosteric activator or an allosteric inhibitor to their respective allosteric sites shifts the sigmoid curve respectively to the left and the right, because of respectively their positive and negative cooperativities with the substrate (heterotrophic effects). Allosteric activators decrease the K_m of the K-series of allosteric enzymes, and increase the V_{max} of the M-series of enzymes.

Allosteric inhibitors increase the K_m of K-enzymes and decrease the V_{max} of M-enzymes.

Isozymes are different proteins catalyzing the same reaction of the same substrates. Isozymes of an enzyme differe from each other in physicochemical properties such as substrate specificity, K_m , allosteric modulations, induction and repression.

Ribozymes are RNA molecules having enzyme activiteis. A few ribozymes, occurring in prokaryotes and unicellular eukaryotes, catalyze post-transcriptional modifications of their pre-RNA transcripts.

A rate-limiting enzyme of a metabolic pathway is a regulated enzyme catalyzing the reactiion of an initial step of the pathway in a single specific directin, having lower substrate-affinity than other enzymes of that pathwaym and determinign the overall rate of the entire pathway.

4.11 Terminal questions

1. (a) Describe the process of irreversible covalent activation, using the examples of two gastro-intestinal enzymes.
   (b) Explain how enzymes may be subjected to reversible covalent modifications by phosphorylation and dephosphorylation, describing such modulations of two suitable enzymes.
2. (a) Discuss the characteristics of allowsteric modulations.

(b) Describe the sigmoid saturation Kinetics, quoting the Hill equation.

(c) Mention the changes in the K- and M-series of enzymes during allosteric modulations.

3. (a) What is Michaelis Constant? Discuss its characteristics.

(b) How is the K_m indicated by the double-reciprocal, Eadie-Hofstee and Wolf-Hanes equations and their plots?

(c) What are rate-limiting enzymes? Mention the characteristics they should possess.

4. (a) Mention the bonds involved in forming enzyme-substrate complexes.

(b) Describe the transition state theory, using it to explain how the formation of the ES complex may help in catalysis.

(c) Discuss two contesting models for the formatin of ES complexes, indicating the superiority of one of them to the other.

5. (a) Describe the Michalis-Menten hyperbolic Kinetics of enzymes, quoting the Michaelis-Menten equation for a single-substrate reaction.

(b) Write how the Michaelis-Menten equation may be linearly transformed into the Lineweaver-Burk equatin, the Eadie-Hofstee equation and the Wolf-Hanes equatiion, quoting their respective equations.

(c) Describe the sigmoid saturation kinetics of enzymes, quoting the Hill equation. What is the Hill coefficient?

6. (a) Explain D-L stereospecificity and optical specificity of enzymes with an exmaple of each.

(b) Mention some experiemntal evidences in support of the formation of ES complexes.

(c) Differentiate between ES formations in singe-displacement and double-displacement bi-substrate reactions.

7. (a) Discuss two contesting models for the molecular basis of allosteric effects.

(b) Explain the cooperativities between allosteric modulators and substrate.

(c) What is homotropic effect in allosterism?

8. Write notes on the following :

(a) *cis-trans* stereospecificity and group specificity of enzymes.

(b) Isozymes

(c) K- and M-series of enzymes

(d) Ribozymes

(e) Induced fit model

(f) Indirect concerted model

4.12 Answers

1. (a) See Section 4.5.1.
   (b) See Sections 4.5.2.
2. (a) See section 4.6.1.
   (b) See Section 4.6.2.
   (c) See last part of Section 4.6.2
3. (a) See Section 4.4.1.
   (b) See Section 4.4.2.
   (c) See Section 4.9.
4. (a) See relevant part of Section 4.2.1.
   (b) See Section 4.2.4.
   (c) See Section 4.2.3.
5. (a) See Section 4.4.
   (b) See Section 4.4.2.
   (c) See Section 4.6.2.
6. (a) See paragraph (d) and (e) of Section 4.3.2.
   (b) See first paragraph of Section 4.2.1
   (c) See paragraph (b) of Section 4.2.1.
7. (a) See Section 4.6.3.
   (b) See paragraph (h) of Section 4.6.1.
   (c) See paragraph (j) of Section 4.6.1.
8. (a) See paragraph (a) and (c) of Section 4.3.2.
   (b) See Section 4.7.
   (c) See the last part of Section 4.6.2.
   (d) See Section 4.8.
   (e) See paragraph (b) of Section 4.2.3.
   (f) See paragraph (a) of Section 4.6.3.

Unit 5 ❐ Induction, Repression and Translation

Structure

5.1 Introduction

You will read in this how the synthesis of enzymes is regulated by controlling the transcription of the mRNA molecules coding for them. In that context, you will get to learn about the operon concept of Monod and Jacoo, how it explains induction and repression of enzyme synthesis in prokaryotes, and also about similar regulations of enzyme synthesis in animals.

You will then proceed to read about the process of polysomal translation of proteins in animals, and how the nascent proteins turned out by translation undergo subsequent modifications into their final forms.

Objectives

Reading of this unit should enable you to :

- Understand the concept of operons in prokaryotes,
- Distinguish between inducible, repressible and constitutive enzymes.
- Describe the lac and ara operons of *E.coli* as examples of inducible catabolic operons,
- Describe the trp operon of *E.coli* as a repressible synthetic operon.
- Narrate induction and repression of enzymes in animals and other eukaryotes.
- Describe how amino acids get bound to the respective tRNA molecules before being incorporated in the peptide chain.

- Discuss the roles of polysomes in protein translation.
- Understand the events in eukaryotic translation.
- Narrate how a peptide chain is initiated, extended and terminated during its translation in animals.
- Describe different types of post-translational modifications of the translated protein.

5.2 Induction and Repression of enzymes

Induction and repression are processes for regulating the synthesis of enzymes as also of nonezymatic proteins. Many of the *regulated enzymes,* particularly the rate-limiting one, of various metabolic pathways are regulated by induction and/or repression.

Induction is the augmentation of enzyme synthesis by means of an increase in the gene transcription of the mRNA coding for that enzyme. It is brought by the effect of specific substances, called *inducers,* on the mRNA transcription; such an inducer, or some other molecule called a *gratuitous inducer* differing from the substrate but often possessing structural similarity with the latter. Synthesis of *inducible enzymes* ordinarily either remain suspended or take place at very low and insignificant rates; in presence of specific inducer(s), transcription of the mRNA bearing the genetic code for the relevant enzyme is manifold enhanced, augmenting thereby the translation of the latter. Many of the regulated enzymes of catabolic pathways are inducible enzymes.

Repression is the diminution or suspension of enzyme synthesis, brought about by the suppression of gene transcription of the mRNA bearing the genetic code for the relevant enzyme. It is brought about by the action of specific metabolites or products, called *repressors,* on the transcription of the particular mRNA. Syntheses of *repressible enzymes* ordinarily continue uninterrupted because of continuous transcription of their respective mRNAs; but in presence of specific repressor(s), gene transcription of the mRNA coding for the relevant enzyme gets suspended, leading to a suspension of translation of the latter. Many regulated enzyme of synthetic (anabolic) pathways are repressible enzymes. Such an enzyme for an initial step of the pathway is frequently repressed with the help of an intermediate of a subsequent step or the ultimate product of that pathway (*fee-back repression*). This ensures the prevention of unnecessary production and accumulation of metabolic intermediates and products the pathway. Sometimes, an initial enzyme of a pathway, which branches into several subsequent pathways, is repressed only in presence of intermediates of products of more than one such branches (*multivalent feed-back repression*).

There are many enzymes are synthesized continuously independent of and free from the effect of any repressor. These are called *constitutive enzymes.* Such constitutivity of enzymes frequently results from genetic mutations that may affect the bindings of repressors to their binding sites on the DNA sense strand.

Induction and repression may regulate enzyme synthesis in both prokaryotes and eukaryotes. But for the sake of simple descriptions of these phenomena, discussions about them are mainly restricted here to prokaryotes.

5.2.1 Operon concept

Monod and Jacob propounded the operon concept to explain the gene transcription of a prokaryotic *polycistronic mRNA.* According to this concept, the polycistronic mRNA is transcribe using as the template, called an *operon* of the DNA sense strand. The operon is described as such a specific segment of the DNA sense strand as consists of (*i*) a cluster of several *structural genes,* each carrying the genetic message for synthesizing a specific enzyme or one of its peptide subunits, (*ii*) an *operator gene* (O gene) adjoining the cluster of structural genes immediately on its 3′ or upstream side and operating the transcription of all those structural genes into a *polycistronic mRNA* strand bearing genetic codes for translating the respective peptides, and (*iii*) a *promoter site* (P site) adjoining the 3′ and of the O gene as also partly overlapping the latter and providing and initial binding site for RNA polymerase on the DNA sense strand (Fig. 5.1). In transcribing the DNA template of the operon, RNA polymerase bound initially to the P site, slides downstream in the 3′ → 5′ direction along the operator gene and then along the cluster of structural genes, transcribing the latter into a polycistronic mRNA strand, until it reaches the 5′ or down stream end of the operon and gets released from the DNA strand. Different segments of the polycistronic mRNA strand would subsequently code for the respective peptides during their translation.

A *regulator gene* is located on the same DNA strand that carries the operon, and either just next to or far away from the 3′ end of the later, but it is not a part of the operon. The regulator gene transcribes a *repressor mRNA* which in turn translates either an active *repressor* protein or an inactive *aporepressor,* that would participate in preventing the RNA polymerase from transcribing the structural genes of the operon (see below) and thus repress all the enzymes coded by those genes.

Three operons of *E. coli* are described below.

(*a*) ***Lac operon :***

The lac operon is an *inducible operon* of *E.coli* for transcribing a polycistronic mRNA bearing the genetic codes for three *enzymes of the lactose-catabolizing pathways* in the microbe (Fig. 5.1) A *promoter site* (P site) occurs at the 3′ or

upstream end of the operon, partly overlaps the 3′ end of the operator gene (O gene; next to it, and provides the initial binding site for RNA polymerase. On the 5′ or downstream side of the O gene, there occurs a cluster of three successive *structural genes,* viz, Z, Y and A genes in the 3′ → 5′ order, and coding respectively for β-galactosidase, galactoside permease and thiogalactoside transacetylase. A regulator gene (I gene) is situated immediately on the 3′ or upstream side of the P site, but is not considered as a part of the lac operon. The I gene transcribes a *repressor mRNA* which codes for an active tetrameric *repressor protein.* The latter directly binds to the O gene and prevents the RNA polymerase, bound to the P site, from moving across the O gene to the structural genes of the operon. So, the structural genes fail to be transcribed into any polycistronic mRNA which could code for and translate the three enzymes. Thus, all three enzymes remain *repressed.*

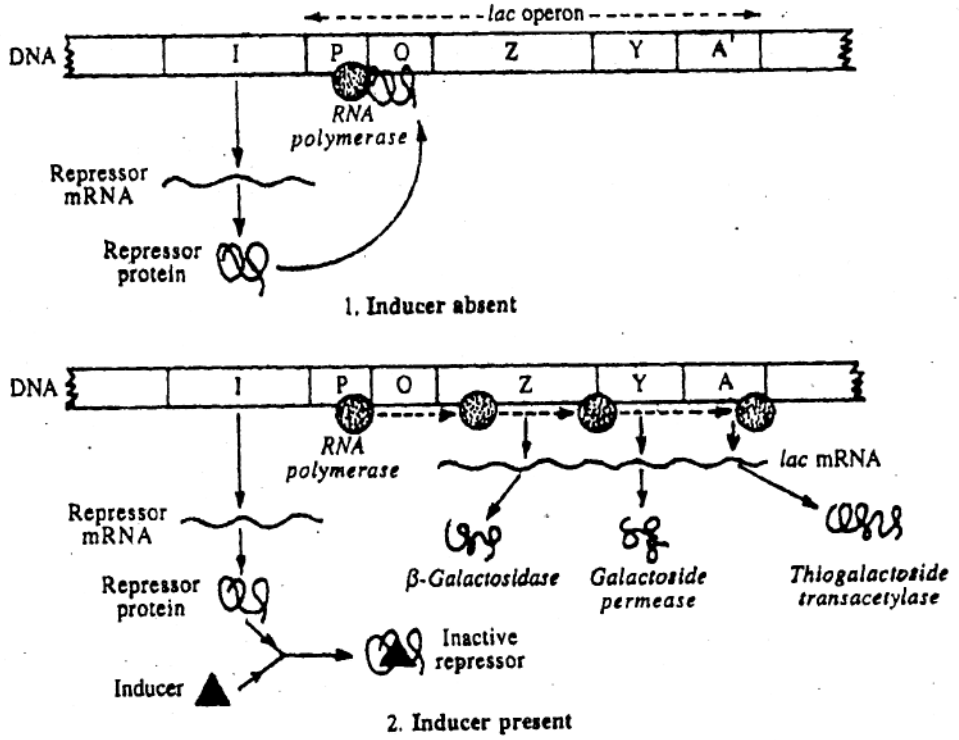


Figure 5.1 : Induction of Lac Operon [From D. Das, Biochemistry, Academic Publishers, 2000]

But in presence of lactose in the medium, small amounts of it enter the cell where it is changed into *1, 6-allolactose* by transglycosylation. This 1, 6-allolactose binds to the operator-bound repressor molecular to get it released from the O gene, and also to free repressor molecule to get it released from the O gene, and also to free repressor molecules to prevent their fresh binding to the O gene. This now allows the RNA

polymerase to bind effectively to the P site as also to slide across the O gene to reach and transcribe the structural genes into the polycistronic mRNA. The latter then starts translating the three enzymes coded by it. The 1, 6-allolactose acts as a *natural inducer* for the lac operon to bring about the *induction* of all three enzymes; in addition, the lac operon may also be induced by *gratuitous inducers* such ad β-methylgalactoside. Induction of lac operon serves as an example of *co-ordinate induction* of more than one enzyme by the effect of an inducer on a single operator gene.

(b) Ara operon :

This is another *inducible catabolic operon* of *E.coli* for transcribing a polycistronic mRNA coding for three *arabinose-catabolizing enzymes.* The operon consists of an *operator gene* (ara O gene), an adjoining *promoter* site (ara I), and just beside them, a cluster of three *structural genes,* viz., araB, araA and araD genes bearing genetic codes for the respective enzymes. A *regulator gene* (araC), situated on the other side of the ara O gene, transcribes a *repressor mRNA* which codes for a *repressor* protein. In the absense of arabinose in the cell, the repressor binds to the ara O gene and represses all three enzymes by blocking the RNA polymerase-catalysed transcription of the polycis ronic mRNA by the structural genes. Arabinose acts as an *inducer* |—| it changes the conformation of repressor, thereby enhancing RNA polymerase activity instead of blocking it. This induces the synthesis of all the three enzymes.

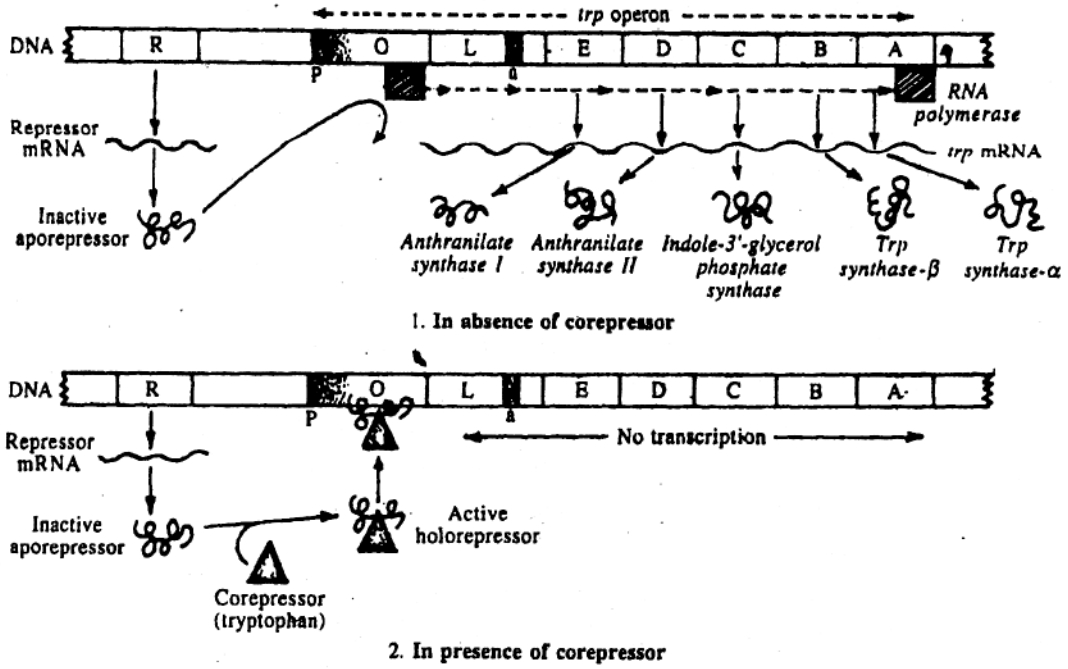


Figure 5.2 : Repression of trp operon [From D. Das, Biochemistry, Academic Publishers, 2000]

(c) Trp operon :

It is *repressible operon* of *E.coli* for transcribing a polycistronic mRNA coding for five peptides which participate in *tryptophan synthesis* in the organism (Fig. 5.2). From its 3 end to its 5 end (upstream to downstream) this operon consists of (*i*) an *operator* (O) *gene* partly overlapping a *promoter site* (P) that provides an initial binding site to RNA polymerase, (*ii*) a *leader gene* (L gene) on the 5′ side of O gene and tanscribing a 5′ -leader sequence in the polycistronic mRNA to be transcribed by the structural genes, and (*iii*) five *structural genes,* viz, E, D, C, B and A gens, in the 3′ → 5′ direction along the DNA sense strand for tanscribing the polycistronic mRNA coding respectively for anthranilate synthases I and II, indole-3′-glycerophosphate synthase, and Trp synthases β and α |——| these five peptides would later combine with each other to from three enzymes for tryptophan synthesis.

A *regular gene* (R gene) occurs far away the 3′ or upstream end of the O gene of the operon on the same DNA sense strand. It trancribes a *repressor mRNA* which codes for an inactive dimeric *apopressor protein,* unable to bind as such to the O gene for bringing about repression. Consequently, so long as the tryptophan concentration is low in the cell, all five peptides continue to be translated; but when significantly high amounts of tryptophan occur in the cell, tryptophan binds as the *corepressor* prosthetic group to the inactive aporepressor to form an active *holorepressor* which in turn binds to the O gene, blocking the RNA-catalyzed transaction of structural genes and repressing all the peptides coded by them. Repression of trp operon is a case of *coordinate repression* of more than one enzyme or peptide by the effect of the repressor on a operator gene. It is also an example of feed-back repression.

5.2.2 Induction and repression in eukaryotes

In animals and other eukaryotes, enzymes as well as non-enzymatic proteins are coded by *monocistronic mRNAs,* each of which bears the genetic code for a single specific protein or peptide. So the concept of operons, each with a cluster of several structural genes transcribed simultaneously into a polycistronic mRNA does not hold good for animals and other eukaryotes. Still, induction and repression are important mechanisms for regulating enzyme synthesis in eukaryotes also; however, in the later, induction or repression affects each monocistronic mRNA individually and separately and thus controls the synthesis of individual enzymes independent of each other.

Several hormones such as corticosteroids, sex hormones and thyroid hormones as well as 1,25-dihydroxycholecalciferol affect the synthesis of many enzymes in animals by induction or repression of specific genes. Many catabolic enzymes of animals such as phospho-fructokinase-1, glucose 6-phosphate dehydrogenase, 6-phosphogluconate dehydrogenase, adipose tissue lapse and tryptophan 2-3,-

dioxygenase are inducible enzymes; many anabolic enzymes such as HMG-CoA reductase, PEP carboxykinase, PRPP synthase and fructose, 1,6-bisphosphatase are repressible enzymes.

The repression of ALA synthase, an enzyme of the porphyrin synthesis pathway in animals is an interesting example of both *aporrepressor-corepressor combination* and *feed-back repression.* This enzyme is repressed by an active *holorepressor* formed by the binding of the end product here as a *corepressor* to an inactive *aporepressore* protein coded by the relevant repressor mRNA.

Syntheses of Ca^{2+} binding and Ca^{2+} carrier proteins or enzymes are induced by 1, 25-*dihydroxychole-calciferol,* an active derivative of vitamin D_3, in intestinal, renal and bone cells. On entering such target cells, 1,25-$(OH)_2D_3$ binds to a specific *mobile receptor protein* in the cytoplasm to form an active 1,25-$(OH)_2D_3$-receptor complex. The latter is translocated to the nucleus where it binds to some nuclear proteins bound to specific genes and induces the transcriptions of those genes to monocistronic mRNA molecules coding for specific Ca^{2+} -binding proteins (e.g. intestinal Ca^{2+} -ATPase) to enhance the synthesis of such proteins.

Similarly, on entering renal tubular and salivary duct cells, the mineralocorticoid hormone *aldosterone* binds to specific soluble but inactive mobile receptor proteins to form active steroid-receptor complexes. The steroid-receptor complex then binds to the *hormone responsive element* (HRE) on the 3′ or upstream side of the promoter site of a specific gene in the DNA strand (Fig. 5.3) and induces the transcription of the gene into a monocistronic mRNA coding for a specific Na^+-transporter for sodium pump mechanism. In a similar way, the glucocorticoid hormone *cortisol* induces the synthesis of some key gluconeogenic enzymes such as PEP carboxykinase and fructose, 1, 6-bisphosphatase, the rate-limiting enzyme glucose 6-phosphate dehydrogenase of the pentose phosphate pathway, and the rate-limiting enzyme adipose tissue lipase for lipolysis in adipocytes.

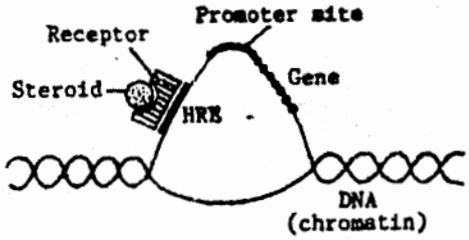


Figure 5.3 : Interaction of steroid-receptor complex with HRE on the 3' side of promoter site of a gene for induction [From D. Das, Biochemistry, Academic Publishers, 2000]

Induction of dihydroorotate dehydrogenase of the pyrimidine biosynthesis pathway by carbamoyl aspartate, an intermediate of an earlier step of the pathway, is an

example of *feed-back forward induction* in mammals, on the contrary, rpression of HMG-CoA reductase of sterol synthesis by the final product cholesterol of the same pathway is a *feed-back repression.*

5.3 Translation of Proteins

Translation is the polymerization of amino acids into a peptide chain by forming peptide bonds between the α-COOH and α-NH_2 groups of successive amino acids arranged serially on the mRNA strand obeying the codon sequence in the genetic code for that peptide. In conformity with the sequence of codons from the 5′ to the 3′ end of the genetic code, the message contained in the latter is *translated* with the help of tRNA molecules and ribosome particles into the amino acid sequence from the N-terminal end to the C-terminal end of the synthesized peptide. A peptide chain being translated grows in its N-terminal → C-terminal direction while the ribosome translates it by moving in the 5′ → 3′ direction along the mRNA. Animals and other eukaryotes translate their cytoplasmic proteins in the cytoplasm, some mitochondrial proteins in mitochondria, and membrane proteins and exportable (secretory) proteins on the membrane of rough Endoplasmic Reticulum (ER).

5.3.1 Formation of amino acyl-tRNA

Each *amino acyl-tRNA synthetase* first catalyzes the binding of the α-COOH group of a specific amino acid by an ester bond to either the 3′-OH or the 2′-OH group of the 3′-terminal ribose residue of either a specific tRNA molecule or one of several specific tRNAs (*cognate tRNAs*) at the cost of *two high-energy bonds* of ATP, producing an amino acyl-tRNA complex (Fig. 5.4).

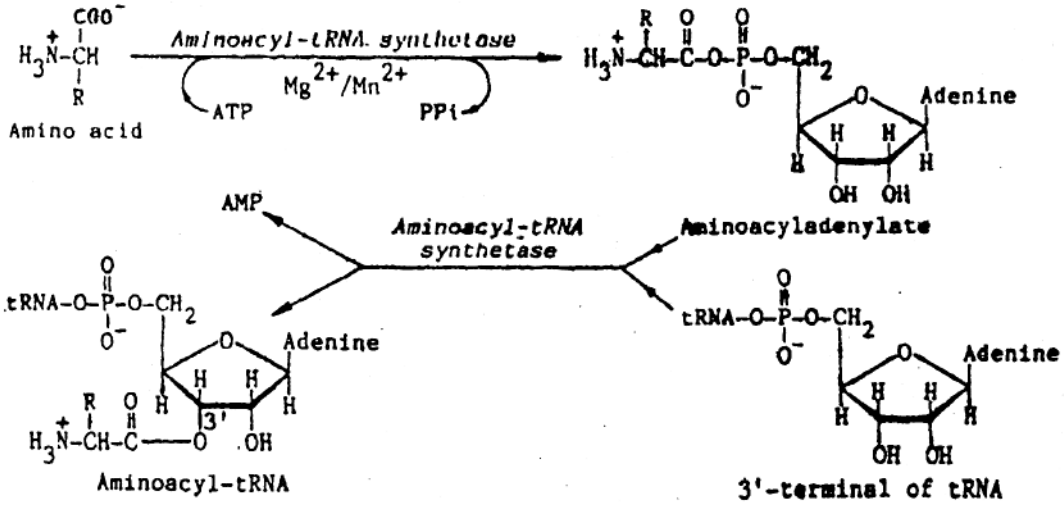


Figure 5.4 : Formation of aminoacyl-tRNA [From D. Das, Biochemistry, Academic Publishers, 2000]

$$\text{Amino acid} + \text{ATP} \rightleftarrows \text{Amino acyl-adenylate} + \text{PPi}$$

$$\text{Amino acyl-adenylate} + \text{tRNA} \rightarrow \text{Amino acyl-tRNA} + \text{AMP}$$

$$\text{Amino acid} + \text{ATP} + \text{tRNA} \rightarrow \text{Amino acyl-tRNA} + \text{AMP} + \text{PPi}$$

The PPi is hydrolyzed by *inorganic pyrophosphatase* to two Pi molecules.

5.3.2 Polysomes in translation

From the cytoplasm, first a 40S ribosomal submit and next a 60S submit come and bind to an mRNA strand over the 5′-codon (first codon) of its genetic code to constitute an eukaryotic 80S ribosome particle on the mRNA (Fig. 5.6). The ribosome would then move in the 5′ → 3′ direction along the mRNA strand, extending the growing peptide chain by incorporating successive amino acids in that chain while passing over the successive codons. However, a number of ribosomes bind successively to the same mRNA strand and flow along it with small intervening gaps, translating as many peptide chains. This combination of the mRNA and several moving ribosomes is called a *polysome* or *polyribosome.*

An *aminoacyl or A site* and a *peptidyl of P site* have been constitued in the three-dimesional form of each ribosome when the latter has been formed on the mRNA by the binding together of its 40S and 60S ribosomal subunits (see above). These two sites can hold respectively an incoming aminocyl-tRNA and a peptidyl-tRNA formed at the preceding step of translation (Fig. 5.5). Each movement of ribosome down the mRNA places its vacant A site opposite the next codon yet to be translated, while its P site comes to lie over the preceding codon and is occupied by the peptidyl-tRNA still held by the ribosome. Next an aminoacyl-tRNA, bearing anticodon bases

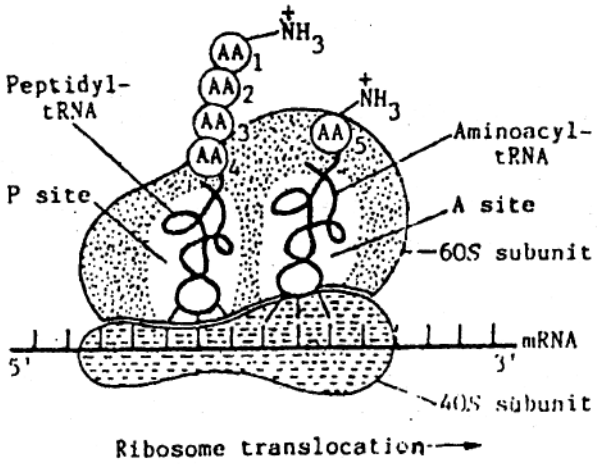


Figure 5.5 : The P and A sites of eukaryotic ribosome [From D. Das, Biochemistry, Academic Publishers, 2000]

complementary to the bases of the codon under the A site of the ribosome, gets hydrogen-bonded to the latter codon and comes to occupy the A site. The peptidyl group of the P-site peptidyl -tRNA is now transferred from the latter by a ribosomal *peptidyl transferase* to the α-amino group of the new A-site amino acyl-tRNA and a peptide bond is formed between the α-carbonyl carbon of the transferred peptidyl group and the α-amino nitrogen of the A-site amino acyl-tRNA. This extends the peptidyl group, now held by the ribosomal A site as the new peptidyl-tRNA, by one amino acid residue, while leaving the deacylated tRNA still in the P site (Fig. 5.7). A ribosome-dependent GT Pause, called *translocase*, next shifts the ribosome in the 5′ → 3′ direction along the mRNA towards the next 3′-codon on the latter. This ejects the deacylated tRNA from its P site which now gets occupied by the new peptidyl-tRNA still hydrogen-bonded to the preceding codon, makes the ribosomal A site vacant and places the latter over the next codon (Fig. 5.7). A new aminoacyl-tRNA now enters the vacant A site and gets hydrogen-bonded to the anticodon bases of the codon under that site. This starts the next cycle of adding another amino acid to the growing peptidyl group the events outlined above. Such events are repeated till the ribosome reaches the 3′-end of the genetic code. Then, the nascent peptide is released from the ribosomal which also falls apart from the mRNA strand in the form of two ribosomal subunits.

Because each tRNA can bind to a specific amino acid only, and also gets H-bonded to a specific codon depending on the base-complementarity between the latter and the tRNA anticodon, amino acids are brought to he ribosome and get peptide-bonded to each other in the Ne-terminal → C-terminal order in the peptide, as determined by the 5′ → 3′ order of codons in the relevant genetic code. So, the latter determines the amino acid sequence or primary structure of the "nascent" peptide.

5.3.3 Events in eukaryotic translation

Peptides are translated in eukaryotes including animals with the participation polysomes and many nonribosomal eukaryotic factors, some of the latter acting as GTP Pauses when bound to ribosomes (ribosome-dependent GT Pases).

1. *Peptide chain initiation :*

(*a*) Nonribosomal *eukaryotic initiation factors* or eIF proteins, viz., eIF-3 and eIF-4C, interact with 80S ribosome released from mRNA to dissociate the ribosome into a 60S ribosomal subunit and a 40S-subunit, eIF-3-4 Complex (Fig. 5.6).

(*b*) *Methionyl-tRNA* (Met-$tRNA_i^{Met}$), eIF-2, GTP and 40S.EIF-3-4C complex bind successively to each other to give an *entry complex.*

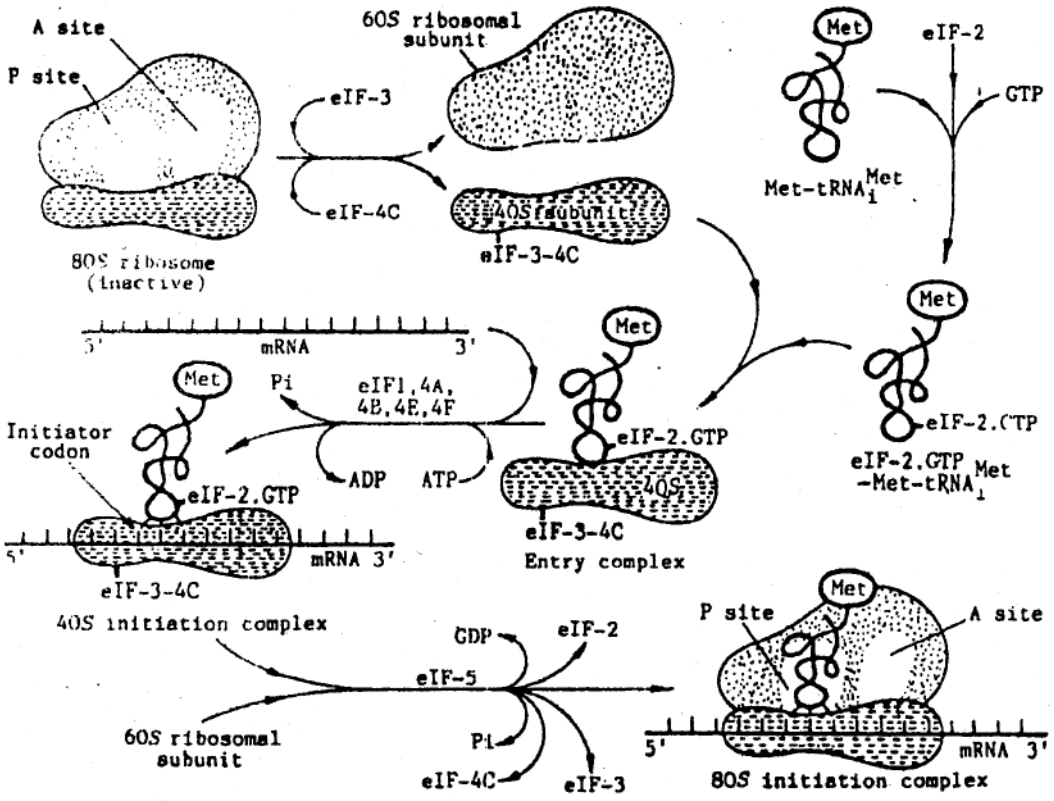


Figure 5.6 : Initiation of eukaryotic translation [From D. Das, Biochemistry, Academic Publishers, 2000]

(*c*) At the cost of high-energy bond of ATP and with the participation of eIF 1, 4A, 4B, 4E and 4F, the entry complex binds initially near the 5′-end of the mRNA stand to form a 40S *initiation complex;* the latter, then moves in the downstream or 5′ → 3′ direction along the mRNA strand to reach the *chain initiator codon* (AUG). On the latter, and gets hydrogen bonded to it by codon-anticondon reaction between the codon bases and the anticodon bases of $tRNA_i^{Met}$ in the 40S initiation complex.

(*d*) eIF-5 next helps in binding a 60S ribosomal subunit to the mRNA-bound 40S initiation complex to form an mRNA-bound 80S *ribosome* with its vacant A site opposite the second codon of the genetic code and its P site accommodating the AUG-bound methionyl-$tRNA_i^{Met}$. All the eIF molecules leave the ribosome during this process. *A ribosome-dependent GTP ase* almost similtaneously hydrolysis the GTP, so long bound to the inititation complex, to release GDP and Pi.

2. *Peptide chain elongation :*

(*a*) Stepwise interactions between a nonribosomal *eukaryotic elongation factor* (eEF1), a GTP molecule and a new amino acyl-tRNA, having base complementarity

of its anticodon with the second codon of the genetic code, release Pi and an eEF-1. GDP complex and result in the codon-anticodon hydrogen-bonding of the second codon with the new aminocyl-tRNA which now gets accommodated in the A site of the ribosome (Fig. 5.7).

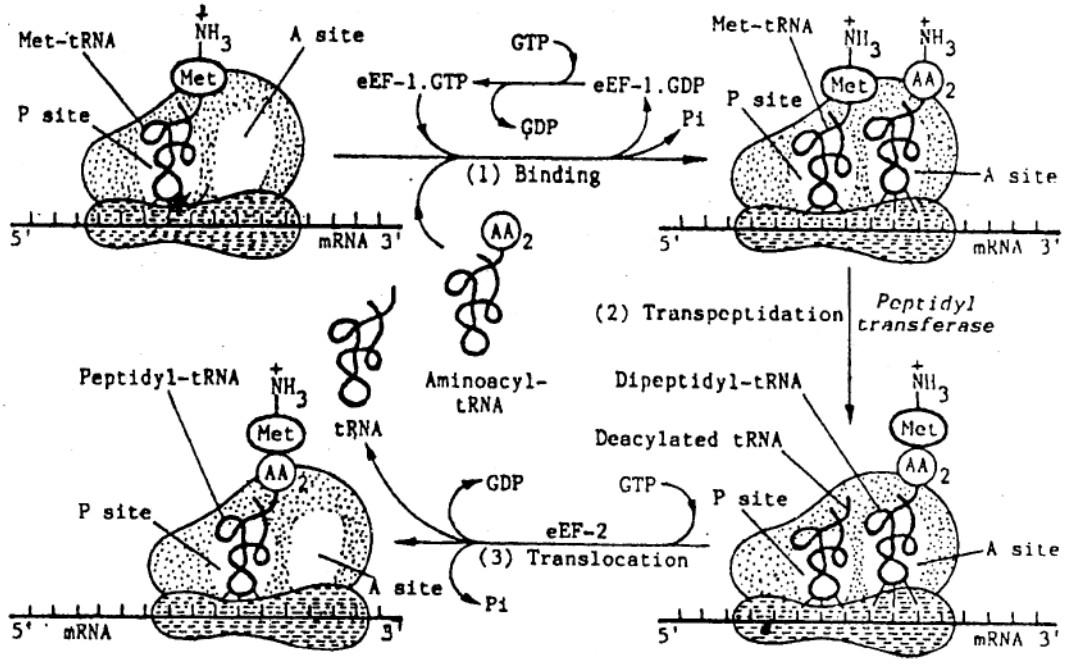


Figure 5.7 : Elongation of peptide chain with a second amino-acid [From D. Das, Biochemistry, Academic Publishers, 2000]

(*b*) *A peptidyl transferase* of the 60S ribosomal subunit next transfers the first amino acid (methionine) from the Met-$tRNA_i^{Met}$ at the P-site to the α-amino group of the amino acid held by the A-site aminoacyl-tRNA, forming a peptide bond between the α-carbonyl carbon of the transferred methionine and the α-amino nitrogen of the amino acid of the A site. This changes the A-site aminoacyl-tRNA into a dipeptidyl-tRNA at the A site while the P-site tRNA becomes deacylated.

(*c*) *A ribosome-dependent GTPase,* called the eEF-2 or *translocase,* hydrolyzes a GTP into GDP and Pi, and uses the released energy in catalyzing the downstream or 5′ → 3′ translocation of the 80S ribosome along the mRNA by one codon. This ejects the deacylated tRNA from the P site, shifts the dipeptidyl-tRNA still bound to the second codon from the A site to the vacated P site, and places the A site thus emptied over the next (third) codon.

This is followed by the next elongation cycle starting with the entry of a new

aminoacyl-tRNA into the vacated A site, as described in step (*a*) above. Subsequent reactions of the elongation cycle then follow.

Repetitions of such elongation cycles elongate the peptidyl-tRNA by one amino acid restidu percycle until the translocation of the ribosome places its A site against a *chain-terminator codon* at the 3′ end of the genetic code of the mRNA.

3. *Peptide chain termination :*

No tRNA can identify or get hydrogen-bonded to any of the chain-terminator codosn, viz., UAA, UAG and UGA. So, when the A site of the ribosome hase

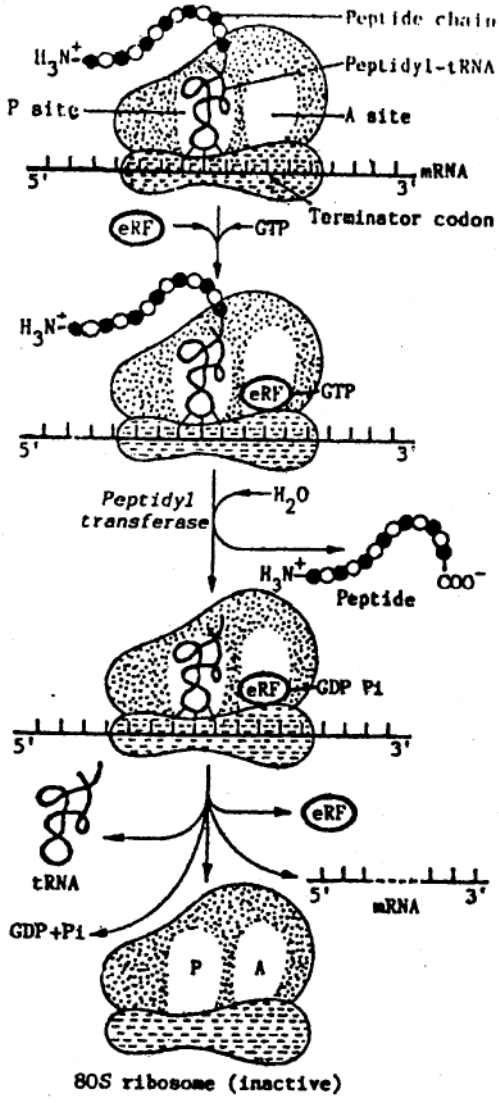


Figure 5.8 : Termination of peptide translation [From D. Das, Biochemistry, Academic Publishers, 2000]

reached a chain-terminator codon, a nonribosomal but ribosome dependent GTP ase called the *eukaryotic release factor* (eRF) binds, along with a GTP, to the chain-terminator codon, instead of any aminoacyl-tRNA (Fig. 5.8). The eRF hydrolyzes the GTP to GDP and Pi, and that energy is utilized by the ribosomal *peptidyl transferase* to release the peptide from the peptidyl-tRNA, along with deacylated tRNA and the eRF from the mRNA.

5.4 Post-Translational Modifications

Before and after its release from the ribosome from the ribosome, an "nascent" peptide is often considerably modified by different process.

5.4.1 Proteolytic cleavages

Specific amino acids or segments of peptides are very frequently removed from the nascent peptide chain by the hydrolysis of specific peptide bonds in the latter.

(*a*) The *N-terminal methionine residue,* with which the translation of each peptide was initiated, is removed either single or as an N-terminal peptide segment from the nascent peptide by peptidase-catalyzed hydrolysis of specific peptide bonds.

(*b*) Some peptide hormones, translated as a single *polyprotein* macromolecule, are released as separate hormone molecules in different endocrine tissues by the post-translationsl hydrolysis of such a polyprotein by tissu-specific peptidase. For example, *pro-opiomelampcprtin* (POMC) translated in both anterior and intermediate lones of pituitary is hydrolyzed at specific peptide bonds in those tisses to yield mainly ACTH and liptropins in the anterior pituitary, but to give α-and β-MSH, endorphins, lipotropins and corticotropin-like intermediate lobe peptide in the intermediate lobe.

(*c*) The short, *N-terminal signal or leader sequence* of nascent exportable and membrane proteins, translated by polysomes on the rough ER membrane, is hydrolyzed away by a microsomal *signal peptidase* even before the protein is released from its ribosome. For example, hydrolysis by signal peptidase removes the N-terminal leader sequence from *preproparathormone* (prepro PTH) initially translated, changing it to shoter prohormone (proPTH).

(*d*) After removal of the signal sequence from some nascent preproteins, additional terminal amino acid sequence may by further removed by other peptidase. For example, after proPTH has been formed by microsomal signal peptidase action on prepro PTH, the N-terminal hexapeptide segment of proPTH is hydrolyzed away by clipase *B* of the Golgi membrane to change proPTH to PTH.

5.4.2 Modification of amino acids

Some unusual amino acids such as pyroglutamate, γ-carbox glutamate, 4-hydrozyproline, phosphoserine and tyrosine 4-O-sulfate, though present in many proteins, are not directly coded by any codon. Such amino acid residues are produced in the protein by post-translational modification of specific coded amino acids already inserted in the protein during its translation.

(*a*) The N-terminal glutamate (Glu) residue of some peptide hormones, such as gastrin TRH and GnRH, is cyclized by *γ–glutamyl cyclase* into a *pyroglytamate* (pGlu) residue through a peptide bonding between its α-NH_2 and γ-COOH groups (Fig. 5.9).

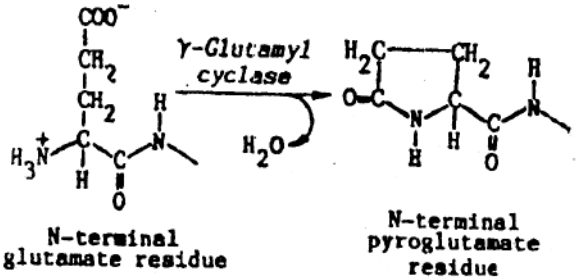


Figure 5.9 : Cyclization of Glu to pGlu residue [From D. Das, Biochemistry, Academic Publishers, 2000]

(*b*) Specific glumate residuess (Glu) of some Ca^{2+} binding proteins such as prothrombin are carboxylated in many tissues like lever and bones at the sidechain γ-carbon by microsomal *protein carboxylase,* modifying those resides to γ-*carboxyglutamate* (Gla) residues with the help vitamin K (Fig. 5.10).

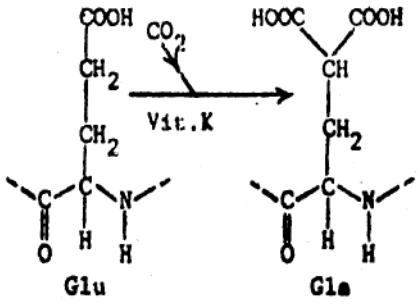


Figure 5.10 : Carbodylation of Glu to Gla residue [From D. Das, Biochemistry, Academic Publishers, 2000]

(*c*) C-terminal *amino acid amide* residues like glycin-amide and prolinamid are produced in many peptides and proteins such as gastrin, oxytocin, vasopressin and calcitonin by the amidation of their C-terminal amino acids. For example, a *peptidylflycine α-amidating monooxygenase* oxidizes the C-terminal glycine residue of a protein to form a double-bond between its α-carbon and α-amino nitrogen; this

double-bond is subsequently hydrolyzed to release glyxylic acid, leaving the nitrogen of glycine as the carboxamide group of the next amino acid; the latter thus becomes the new C-*terminal amino acid amide* residue such as valinamide, glycinamide and prolinamide (Fig. 5.11). The reaction uses vitamin C as a cofactor.

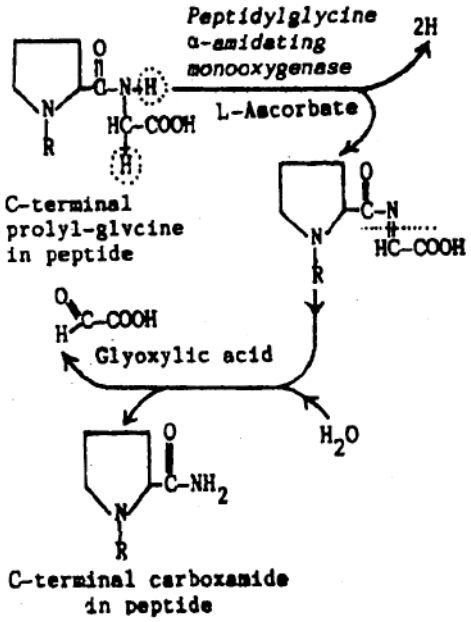


Figure 5.11 : Amidation of peptidylglycine [From D. Das, Biochemistry, Academic Publishers, 2000]

(*d*) The sidechain OH groups of specific tyrosine residues of some proteins, e.g. Tyr^{33} of cholecystokinin (CCK-39), get sulfated to form *tyrosine 4-0-sulfate* residues with the help of adenosine- 3′-phosphate -3′-phosphosulfate (active sulfate).

(*e*) The sidechain OH groups of specific tyrosine residues of some proteins like insulin receptos, and of specific serine residues of some other proteins like glycogen phosphorylase get phosphorylated by *protein Kinases* and ATP, changing then into *phosphotyrosine* and *phosphoserine* residues, respectively.

(*f*) Specific proline and lysine residues of some proteins like procollagen are hydroxylated by enzymes such as *procollagen prolyl hydroxylase* and *procollagen lysyl hydroxylase,* changing them respectively to *4-hydroxyproline* and *5-hydroxylysine* residues (Fig. 5. 12). The reaction uses yitamin C as a cofactor.

(*g*) *Methlyferases* transger the labile methul group of S-adenosylmethionine (active methionine) to the side chain amino groups of specific lysine, histidine and arginine residues modifying them respectively to methyllysine, N^3-methylhisitidine and ε-N-methylarginine residues.

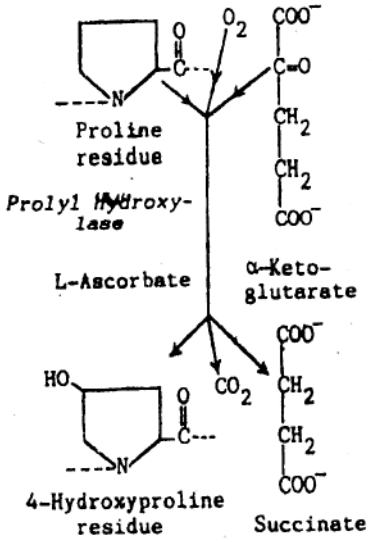


Figure 5.12 : Hydroxylation of proline to hydroxybroline residue [From D. Das, Biochemistry, Academic Publishers, 2000]

(*h*) The sudechain ε-NH_2 group of specific lysine residues in collagens and elastins may be oxidatively deaminated by *lysyl oxidase* into ε-aldehyde group, thus modifying them to *allysine* residues (Fig, 5.13.). Lysine, allysine and histidine residues of different peptide chains may then be cross-linked and further modified to residues like desmosine, allysine aldol, lysinonorleucine and aldol histidine.

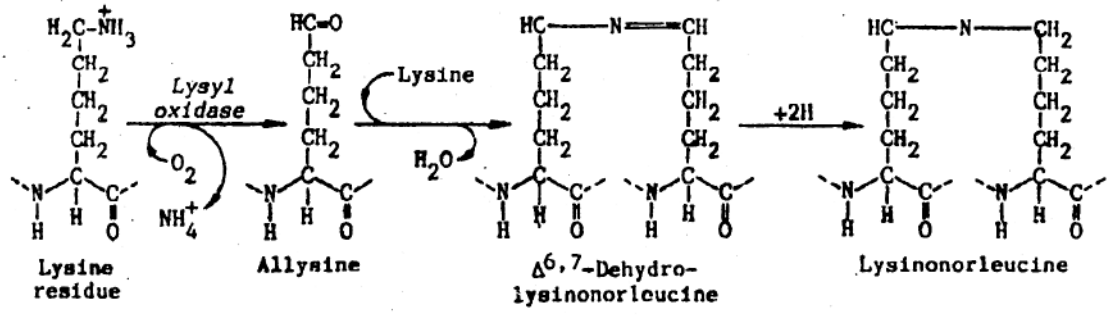


Figure 5.13 : Formation of lysine derivatives n peptides [From D. Das, Biochemistry, Academic Publishers, 2000]

5.4.3 Terminal additions

The peptide chain, left after post-translation removal of the N-terminal methionine and other specific amino acid sequences, may be extended at either of its ends by the addition of new amino acids. For example, the peptide chain of tubulin may be

extended by adding a tyrosine molecule at its C-terminal end by ATP-dependent *tblin-tyrosine ligase.*

5.4.4 Glycosylation

Using *dolichol phosphate* and *nucleotide-sugar complexes* such as UDP-glucose, UDP N-acteylglucosamine. GDP-mannose and CMP-sialic acid, microsomal and Golgi-membrane *glycosyltransferases* transfer monosaccharides and heteroglycans to many membrane and exportable proteins to form their oligosaccharide prosthetic groups, bond to the serine, therenine, arginine or hydroxylysine residues of those proteins by N-or O-glycosidic bonds.

5.5 Summary

Operon is such a segment of prokaryotic DNA that acts as a unit transcribing polycistronic mRNA with genetic codes for more than one peptide. From its 3′ to 5′ end, the opeon consists of a promoter (P) site for the initial binding of RNA polymerase, an operator (O) gene operating the operon, and a cluster of more than one structural genes transcribing together a polycistronic mRNA strand. The operon is regulated by a regulator gene either close to or far away from the 5′ -end of the operon. Lac operon, ara operon and trp operon of *E. coli* have been been briefly described in this unit.

Induction is the augmentation of the gene transcription of an mRNA leading to the enhanced translation of the peptide code by it. Repression is the decline or suspension of such transcription. In inducible operons such as the lac operon, the reglator gene transcribes a repressor mRNA that tranlates an active repressor; the latter ordinarily remains bound to the O gene to block the passage of the P site-bound RNA polymerase past the O gene and along the successive structural genes; the transcription of the latter into a polycistronic mRNA thus remains repressed. Binding of specific ligands (indcers) to the repressor inactivates the latter and allows the RNA polymerase to pass along the O gene and then along the structural genes, inducing their transcription into the polycistromic mRNA that translates all the petides coded by it.

For repressibe operons like the trp operon, the repressor protein (aporepressor) is inactive by itself so that the structural genes ordinarily continue transcribing the corresponding plocistronic mRNA and all the coded peptides continue to be translated, But the binding of a specific ligand (corepressor) to the aporepressor changes it to an active holorepressor which can now bind to the O gene and repress the polycistronic mRNA and the peptides coded by it.

Induction and repression of many ekaryotic enzymes have also been described in this unit although the operon concept is not applicable to eukaryotes.

Polysomes consist of several successive 80S eukaryotic ribosomes passing along each mRNA strand, each of those ribosome translating a peptide molecule by incorporating amino acids in the latter in a specific sequence according to the sequence of codons along the mRNA. For this, each amino acid binds to a specific tRNA to form an amino acyl-tRNA which binds to a specific codon of the mRNA according to the base-complimentarity between the codon and the anticodon of the tRNA. Each ribosome in the polysome assembly has an A site for receiving a new incoming amino acyl-tRNA and a P site holding the already formed peptidyl-tRNA complex of the preceding step. The transfer of the peptidyl group from the latter to the amino acyl group of A-site tRNA by a ribosomal peptidyl transferase at each step elongates the peptidyl group by one amino acid residue. A GTP-dependent translocase shifts the ribosome after each such step by one codon along the $5' \rightarrow 3'$ direction on the mRNA. Initiation, elongation and termination of translation depend also on specific nonribosomal proteins called eukaryotic initiation, elongation and termination factors, respectively. The energy for the endergonic process of translation is supplied by the exergonic hydrolysis of GTP and ATP molecules.

The translated 'nascent' peptide is subsequently modified into its final form by various post-translational modifications, such as modifications of some amino acid residues into unusual amino acids like hydroxyproline, allysine, γ-carboxyglumate, methylhistidine, pyroglutamate and prolinamide, hydrolytic removal of N-terminal methionine residue and N-terminal leader sequence, glycosylation of exportable and membrane proteins, and terminal addition of new amino acids at either and of the peptide chain.

5.6 Terminal questions

1. (a) Give an account of the role of polysomes in translating peptides.
   (b) Describes with the help of suitable diagrams how γ-carboxyglutamate and pyroglutamate residues are formed in a translated peptide chain.
2. (a) Describe the lac operon of *E. coil* and its induction, with a suitable diagram.
   (b) Using a flowchart, explain the action of an amino acy-tRNA synthetase.
   (c) Describe the initiation of translation of a peptide chain in eukaryotes with the help of diagram, mentioing the participation of ribosomal subunits and nonribosomal eukaryotic factors in the process.

3. (a) Describe the operon concept of Monod and Jacob.
   (b) Using a suitable diagram, describe the trp operon of *E.Coli.* and explain its repression.
   (c) Write how corticosteroids induce specific proteins in animals.
4. (a) Describe the events in the initiation and elongation of peptide chains in eukaryotes. Using suitable diagrams and mentioning the roles of ribosome, and nonribosomal eukaryotic factors in those events.
   (b) Discuss how the 'nascent' peptide chain undegoes post-translation modifications by different proteolytic cleavages.
5. (a) Discuss the phenomena of induction and repression in eukaryotes with examples.
   (b) Describe with the help of a suitable diagram how the trp operon is repressed in *E. coli* to illustrate prokaryotic repression in terms of the operon concept.
   (c) State what you understand by gratuitous inducers, coordinate induction, feed back repession and coordinate repression, citing one example of each.
6. (a) What are the roles of peptidyl transferase and translocase in eukaryotic translation?
   (b) Narrate the events in the initiation and termination of peptide translation in eukaryotes, with suitable diagrams.
   (c) What is the fate the N-terminal methionine residue of translated 'nascent' peptides in eukaryotes?
7. (a) Describe the post-translational modifications of amino acid residues in a 'nascent' peptide to give rise to unusual amino acid residues, with suitable examples and diagrams.
   (b) Narrate with the help of suitable diagram how a peptide chain is elongated in each cycle of its translation in eukaryotes.

5.7 Answers

1. (a) See Section 5.3.2
   (b) See paragraphs (a) and (b) of Section 5.4.2
2. (a) See paragraph (a) of Section 5.2.1
   (b) See Section 5.3.1
   (c) See No. 1 of Section 5.3.3

3. (a) See the first two paragraphs of Section 5.2.1
   (b) See paragraph (c) of 5.2.1
   (c) See the penultimate paragraph of Section 5.2.2
4. (a) See Nos. 1 and 2 of Section 5.3.3
   (b) See Section 5.4.1
5. (a) See Section 5.2.2
   (b) See paragraph (c) of Section 5.2.1
   (c) See second and third paragraphs of Section 5.2, last two sentences of paragraph (a) of Section 5.2.1, and last two sentences of paragraph (c) of Section 5.2.1.
6. (a) See paragraphs (b) and (c) under No. 2 of Section 5.3.3
   (b) See Nos. 1 and 3 of Section 5.3.3
   (c) See paragraph (a) Section 5.4.1
7. (a) See Section 5.4.2.
   (b) See No. 2 of Section 5.3.3

Unit 6 ❐ Adipose Tissue Profile

Structure

6.1 Introduction

You will read in this unit about different roles of adipose tissues in the metabolism of lipids. You have already learnt in Unit 1 about the beta-oxidation of fatty acids. Here, you will read about how fatty acids are used in adipocytes in synthesizing triacylglycerol fats for storage in the adipore tissue, how fatty acids are also used in synthesizing phospholipids, how the adipose tissue fat is hydrolyzed to release the fatty acids, and how they are released in the blood for their mobilization to other tissues for oxidation.

You will also read how lipoproteins are synthesized using fatty acids, fats, phosphopids, cholesterol and proteins. Fats and fatty acids are mobilized in blood and lymph in the form of lipoproteins. You will know about their circlation and metabolism. You will learn how cholesterol bound to two forms of lipoproteins of low-density and high-density, has different fates in circulation, and its relation to atherosclerosis. You will also read how the nonavailability of phospholipids and lipoproteins too much lipolysis in adipoacytes, and toxic effects of chemicals on the liver may lead to a high accumulation of fat in the liver.

Objectives

After reading this unit, you should be able to :

- Describe how triacylglycerols are synthesized from fatty acids for storage as fats in adipocytes.
- Narrate the synthesis of phospholipids from fatty acids.
- Understand the process and regulation of lipolysis of adipose tissue fat to fatty acids.
- Understand the composition and functions of different types of plasma lipoproteins.
- Narrate the circulation and fate of different plasma lipoproteins.
- Understand the relations between LDL-cholesterol, HDL-cholesterol and atherosclerosis,
- Explain how fatty liver may result due to decreased mobilization of fat from the liver because of varios factors.

6.2. Triacylglycerol Synthesis

Fatty acids, dietary as well as endogenous, are stored principally as fats in the adipose tissue of the body. For such storage, fatty acids are converted by smooth ER enzymes, and to a lesser extent by mitochondrial enzymes, into tracylglycerol fats in adipocytes and subsequently stored in those cells. Adipocytes occur mainly in the subcutaneous adipose tissue, mesentery, omentum and peritoneum of abdominal cavity, mammary gland and muscles.

Significance of triacylglycerol synthesis :

Whenever energy-producing nutrients, viz., fatty acids and carbohydrates, are available in the body in excess of their immediate requirement in energy production, they are converted to triacylgylcerols which are not much mobilization in that form and instead, are stored as fat in adipose tissues.

(*i*) When on high-fat diet, the dietary fatty acids in excess of their immediate metabolic need are converted to triacylglycerols and stored as adipore tissue fat for future use in energy production.

(*ii*) The mobilization of dietary fatty acids from the liver to the adipose tissues, where they are changed to fats for storage, decreases the fat-load of the liver, maintains its numerous other normal functions and prevents the formation of fatty liver (vide 6.6.).

(*iii*) When on a high-carbohydrate dite, citrate is formed as a TCA cycle intermediate from the glycolysis-endrproduct pyruvate, in excess of the need for its immediate aerobic oxidation. This excess amount of citrate is translocated to the

cytoplasm where it yields acetyl-CoA used in synthesizing fatty acids for forming adipose tissue fat. Moreover, glycerol 3-phosphae, used in tri-acylglycerol synthesis, is also largely formed from glycolytic intermediates like dihydroxyacetone phosphate. So, triacylglycerol synthesis also helps to convert the surplus carbohydrate metabolites into storage fats to be used subsequently for energy production when needed.

(*iv*) Hibernating mammals like polar bears have high rate of triacylglycerol synthesis from dietary nutrients such as fatty acids in the pre-hibernation period, for utilizing the stored fat for energy during hibernation. Similar is the case of nesting female birds preparatory to their nesting. Aquatic mammals such as whales, seals and walruses also have a high rate of riacylglycerol synthesis to form thick fatty outer coats of the body like the blubber of whales which acts as a thermoinsulator.

6.2.1 Thioesterification of fatty acids

Since fatty acids can participate in triacylglycerol synthesisi as acyl-CoA thioesters only, thioesterification of fatty acids must precede their use in the process. *Fatty acid thiokinases,* enzymes showing *chain-length specificites,* catalyze the thioesterifications of fatty acids of specific chain-length |——| some acting only on long-chain fatty acids such as palmtic (C_{16}) and steric (C_{18}) acids, some others only on medium-chain fatty acids (C_4-C_{12}) like caproic (C_6), octanoic (C_8) and capric (C_{10}) acids, and still others on short-chain (C_2-C_3) acids such as acetic (C_2) and propionic (C_3) acids. A thiokinase first transfers the adenyl (AMP) group of an ATP to the carboxyl group of the fatty acids, forming an acyladenylate (acyl-AMP) intermediate and releasing PPi, and next replaces the adenylate group of that intermediate by a coenzyme A molecule, forming acyl-CoA and realing AMP (Fig. 6.1). The PPi released is immediately hydrolyzed by *morganic pyrophosphatase* into two Pi molecules Overall :

Fatty acid + ATP + CoA + H_2O → acyl-CoA + AMP + 2Pi

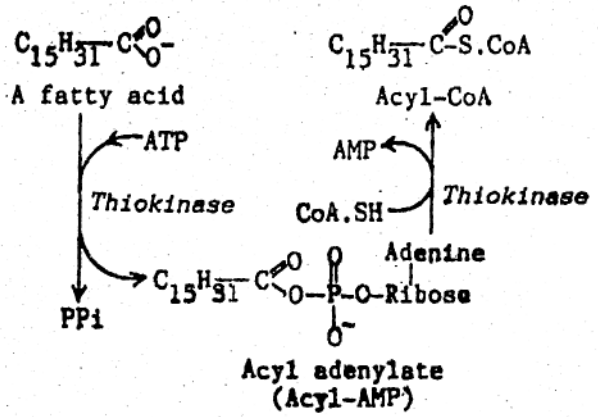


Figure 6.1 : Action of fatty acid this kinase

6.2.2 Glycerophosphate pathway

Acyl-CoA molecules are mostly converted to triacylglycrols by this pathway in the liver, mammary glands, adipore tissue and muscles, using smooth ER enzymes.

(*i*) In the adipocytes of mammary glands and in hepatocytes, cytoplasmic *glycerol kinase* phosphorylates glycerol to glycerol 3-phosphate, using ATP (Fig. 6.2). But adipocycles of subcutaneous fats, abdominal fats and muscle fats as well as enterocytes of intestine form glycerol 3-phosphate by reducing glycolytic dihydroxyacetone phosphate with NADH and cytoplasmic *dihydroxyacetone phosphate dehydrogenase.* The two alternatives are summarized below.

Glycerol + ATP → Glycerol 3-phosphate + ADP.

dihydroxyacetone phosphate + NADH + H → Glycerol 3-phosphate + NAD.

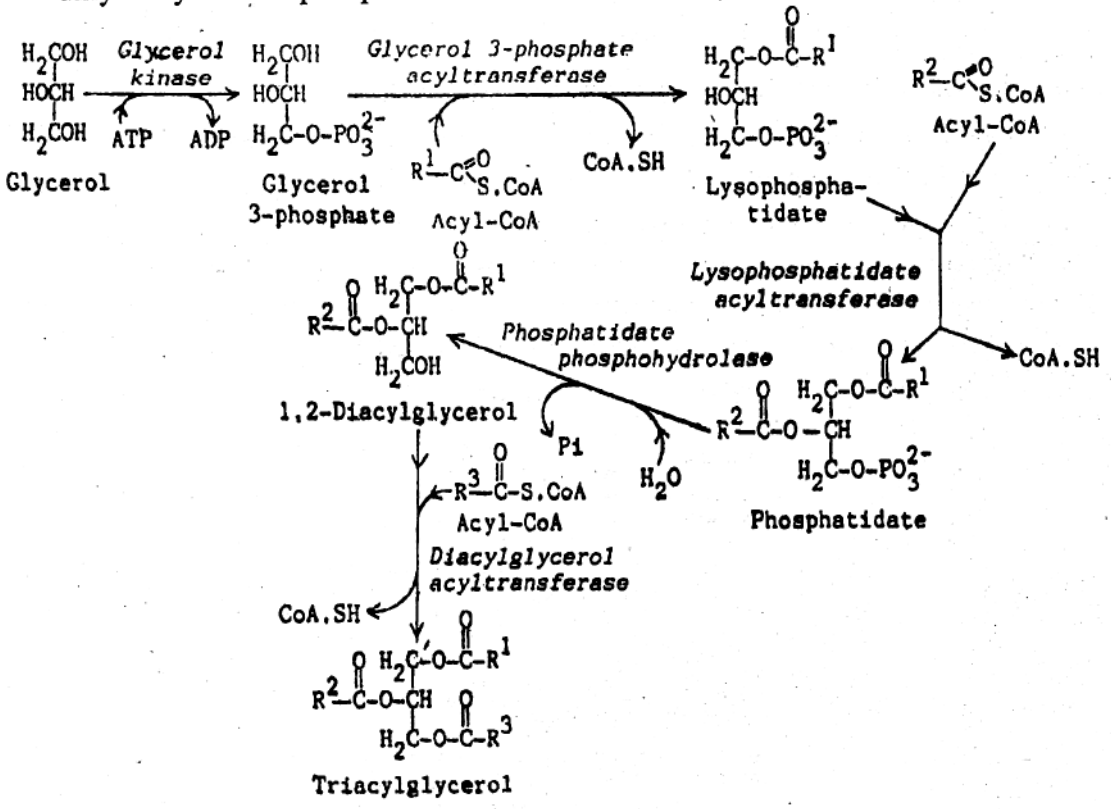


Figure 6.2 : Glycerophosphate pathway of triacylglycerol synthesis

(*ii*) The acyl group of a long-chained *saturated* acyl-CoA molecule (palimitoyl-CoA in mose cases) in next transferred by *glycerol 3-phosphate acyl-transferase* to the C^1 of glycerol 3-phosphate, changing the latter to lysophosphatidate and relasing coenzyme A.

(*iii*) The acyl-group of a long-chain *unsaturated* acyl-CoA (mostly mono-unsaturated oleyl-CoA and in many cases, polyenoic linoleyl-, linolenyl-and arachidonyl-CoA) is next transferred by **lysophosphatidate acyltransferase** to the C^2 of lysophosphatidate, changing the latter to phosphstidate and releasing another CoA.

(*iv*) The phosphate group of phosphatidate in released by the hydrolysis of its phosphoester bond with C^3 of the phosphatidate by *phosphatidate phosphohydrolase* to change the latter to 1, 2-diacylglycerol.

(*v*) The acyl group of another long-chain **saturated** acyl-CoA molecule is next transferred to the C^3 of 1,2-deacylglycerol by **diacylglycerol acyltransferase,** changing the latter to a triacylglycerol and releasing another CoA.

Overall : Glycerol 3-phosphate + 3acyl-CoA + HO → triacylglycerol + 3CoA + Pi

6.2.3 Dihydroxyacetone phosphate pathway

This is an alternative pathway for the microsomal and peroxisomal synthesis of triacyglycerol by subcutaneous, abdominal and muscle adipocytes.

(*i*) Dihydroxyacetone phosphate, produced in glycolysis, is acylated to 1-acyldihydroxyacetone phosphate by the esterification of its C^1 – OH group with an acyl group from an acyl-CoA molecule by *dihydroxyacetone phosphate acyltransferase* (Fig. 6.3). The coenzyme A of acyl-CoA is released free.

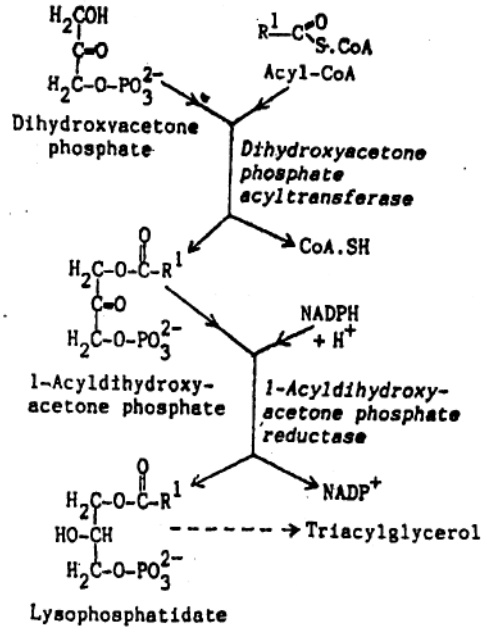


Figure 6.3 : Dihydroxyacetone phosphate pathway for triacylglycerol synthsis [From D. Das, Biochemistry, Academic Publishers, 2000]

(*ii*) 1-Acyldihydroxyacetone phosphate is next reduced to lysophosphatidate by receiving electrons from NADPH under the action of *I-acyldihydrozyacetone phosphate reductase.*

(*iii*) Lysophosphatidate is next changed to triacylglycerol through the successive reactions described in paragraphs (*iii*), (*iv*) and (*v*) of the glycerophsphate pathway (wide 6.2.2).

Overall : Dihydroxyacetone phosphate + 3acyl-CoA + NADPH + H^+ + H_3O → triacylglycerol + 3CoA + $NADP^+$ + Pi

6.3 Adipose Tissue Lipolysis

In diabetes, carbohydrate deprivation or starvation, there happens a dearth of carbohydrate metabolites for immediate energy production in tissues. In such conditions, the stored adipose tissue fats need to be hydrolyzed into free fatty acids for tissue oxidation for energy. This is accomplished in adipocytes by the successive hydrolysis of the stored triacylglycerols by three lipases into free fatty acids and glycerol, followed by the *mobilization of the fatty acids* to other tissues through the blood and their *enhanced oxidation.*

(*i*) Triacylglycerol molecules of adipocytes are first hydrolyzed, each into a free fatty acid and a diacylglycerol, by the *adipose tissue triacylglycerol lipase,* also known as the *hormone-sensitive triacylglycerol lipase* because of its regulation by different hormones (Fig. 6.4.).

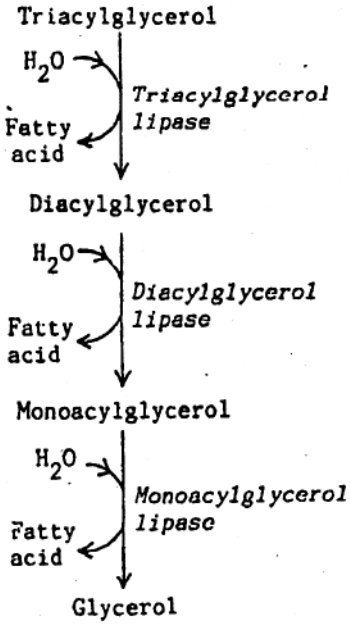


Figure 6.4 : Adipose tissue lipolysis [From D. Das, Biochemistry, Academic Publishers, 2000]

(*ii*) Diacylglycerol is next hydrolyzed by *diacylgtycerol lipase* into another fatty acid and a monoacylglycerol.

(*iii*) Finally, monoacylglycerol is hydrolyzed into glycerol and a third fatty acid by *monoacylglycerol lipase.*

The fatty acids thus released from adipose tisssue may then pass through blood to other tissues. Adipocytes maintain two somewhat distinct *metabolic pools* of fatty acids. One of these pools consists of fatty acids collected by the adipocyte after their release by the action of membrane *lipoprotein lipase* on plasma chylomicron and VLDL particles coming respectively from liver and intestine. The sources of the fatty acids of this pool thus consist of either the absorbed dietary acylglycerols and fatty acids, or the acylglycerols synthesized in enterocytes from dietary fatty acids. The immediate fate of the fatty acids of this pool is mostly the synthesis of triacylglycerols and their storage as adipose tissue fats. The second fatty acid pool in adipocytes consists of fatty acids, released by lipolysis of the stored triacylglycerols and destined to be caried by blood as *albumin-fatty acid complexes* to other tissues, mostly for oxidation.

Regulation of adipose tissue lipolysis :

The rate of adipose tissues lipolysis is enhanced during starvation, carbohydrate deprivation and stress, and is decreased when on high-carbohydrate or high-fat diets. Such regulation is carried out mainly by regulating the *hormone-sensitive triancylglycerol lipase,* the *rate-limiting enzyme* of adipose tissue lipolysis.

(*a*) Fall in blood sugar owing to poor-carbohydrate diets or starvation stimulates secretions of glucagon, adrenaline and glucocoricoids. Conditions of stress also enhance adrenaline and glucocorticoid secretions. (*i*) *Glucocorticoids* such as cortisol induce the synthesis of triacylglycerol lipase and consequently enhance adipose tissue lipolysis. (*ii*) *Glucagon and adrenaline* increase *cyclic AMP* in adipocytes, which activites a *protein kinase;* the latter uses ATP to phosphorylate inactive triacylglycerol lipase *b* to *active triacylglycerol lipase a,* thus enhancing adipose tissue lipolysis. Such lipolytic hormones, therefore, mobilize more fatty acids from the stored fat to raise the plama free fatty level in fasting, carbohydrate deprivation or stress.

(*b*) *Growth hormone and thyroid hormones* also enhance lipolysis and blood fatty acid level. The former brings about a rise in cAMP in the adipocyte by inducing some proteins involved in its formation, while the latter hormones facilitate the lipolytic actions of other hormones.

(*c*) High fat intake increases the plasma FFA level, which inhibits adenylate cyclase in adipocytes, thus decreases the formation and concentration of cAMP in those cells and consequently decreases adipose tissue lipolysis.

(*d*) The rise in blood sugar, when on high-carbohydrate diets, stimulates the secretion of insulin, an *anti-lipolytic* hormone. *Insuline (i)* decrease cAMP formation in adipocytes by inhibiting adenylate cyclase, and (*ii*) activates *phosphodiesterase* which hydrolyzes cAMP to inactive 5'–AMP; both these actions lower cAMP concentration in adipocytes and consequently decrease the action of adipose tissue lipase. (*iii*) Moreover, insulin inhibits the induction of the latter by glucocorticoids. All these lead to a decline in adipose tissue lipolysis and in plasma FFA.

6.4 Phospholipid Synthesis

Cells such as hepatocytes entercytes and adipocytes, playing some roles in fat mobilization into the blood, synthesize significant amounts of phospholipids for helping in the transport of nonpolar lipids in aqueous body fluids. These phospholipids are synthesized by enzymes of smooth ER membranes.

6.4.1 Phosphatidylcholines and phosphatidylethanolamines

(*a*) Choline is phosphorylated to phosphocholine (PC) by *choline Kinase,* using ATP (Fig. 6.5). Similarly, ethanolamine is phosphorylated to phosphoethanolamine (PE) by *ethanolamine kinase,* using ATP.

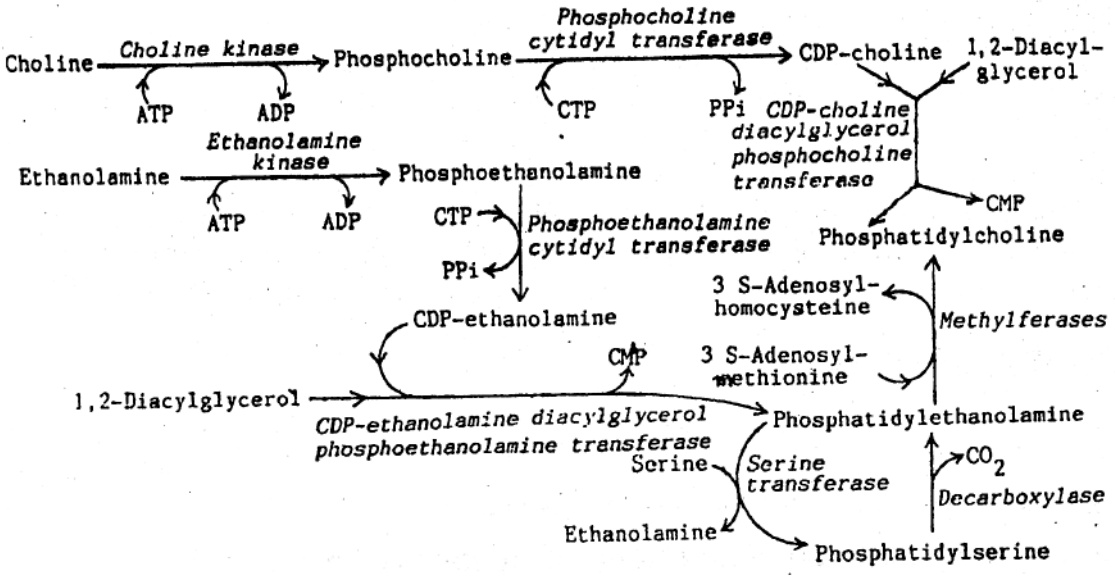


Figure 6.5 : Synthesis of phosphatidylcholines, phosphatidyl ethanolamines and phosphatidyl serines.

(*b*) Cytidylic acid (AMP) is transferred from cytidine triphosphate (CTP) to phosphocholine by *PC cytidyltransferase,* and to phosphoethanolamine by *PE cytidyl*

transferase, changing them respectively to cytidyldiphosphocholine (CDP-choline) and CDP-ethanolamine, with the release of PPi in each case.

(*c*) *CDP-ethanlamine diacylglycerol phosphoethanolamine transferase* next transfers phosphoethanolamine from CDP-ethanolamine to, 1, 2-diacylglycerol, producing phosphatidyl ethanolamine and releasing CMP. Similarly, *CDP-choline diacylglycerol phosphocholine transferase* transfers phosphocholine from CDP-choline to, 1, 2-diacylglycerol, changing the latter to phosphatidylcholine and releasing CMP; alternatively, phosphatidylcholine may also be formed from phosphatidylethanolamine by *methylferase*-catalyzed transfer of three labile methyl groups from three successive S-adenosyl–methionine molecules to the ethanolamine residue of phosphatidylethanolamine.

6.4.2 Phosphatidylserines

Phosphatidylserine may be formed from phosphatidylethanolamine by a *serine transferase* replacing the ethanol amine residue of the latter with serine.

6.4.3 Significance in lipid mobilization

Phospholipids have *amphipathic molecules,* each with a polar headgroup and two nonpolar hydrocarbon tails, and occur in different amounts (5–30%) in different types of lipoprotein particles. Because of their amphipathic molecules, a single layer of phospholipd molecules forms a *surface monolayer* over the lipid core of each lipoprotein particle in plasma and lymph (Fig. 6.6). In that surface monolayer, the monpolar tails of each phospholipid molecule are directed towards the nonpolar lipid core of the lipoprotein particle while its polar headgroup is oriented towards the aqueous phase adjoining the outer surface of the particle. Such orientation phospholipid molecules in the monolayer at the water-lipid interface enables nonpolar lipids to be transported in the core of lipoprotein particles in plasma and lymph. So, phospholipids are *essential for fat mobilization* in aqueous body fluids.

6.5 Lipoproteins

Lipoproteins are the principal vehicles for the mobilization and transport of nonpolar lipids such as triacylglycerols, cholesterol esters and free fatty acids in aqueous body fluids such as plasma and lymph.

6.5.1 Structure and composition of lipoprotein particles

A plasma lipoprotein particle may be smaller than 6 nm and even more than 800 nm in diameter (Fig. 6.6). Each such particle has a ~2 nm thick *surface monolayer* of amphipathic lipids such as phospholipids, glycolipids and free cholesterol. These amphipathic lipid molecules from a single layer on the outer surface of the particle, with ther polar head-groups oriented on the water-adjoining surface of the particle

and their nonpolar tails or groups towards the nonpolar **central core** of the particle. The nonpolar lipids in the central core consist largely of traicylglycerols and cholesteryl esters. Each particle bears on its surfaces some easily extractable apolipoproteins which can be easily transferred to other lipoprotein particles; they are called *peripheral apolipoproteins;* e.g. apo-A, apo-C, apo-D and apo-E. Some others like apo-B100 are deeply embedded in the surface monolayer of the particle, reachin its central lipid core, and are called *integral apolipoproteins.*

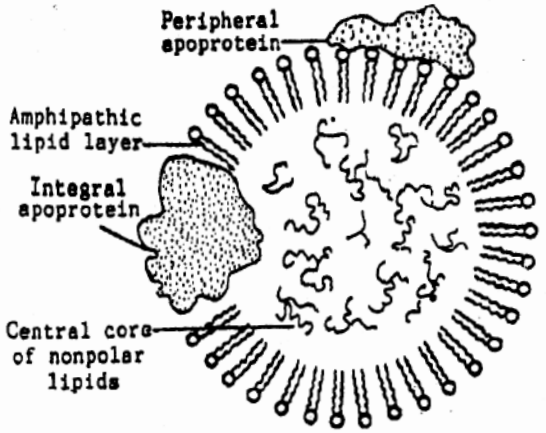


Figure 6.6 : Structure of a lipoprotein particle [From D. Das, Biochemistry, Academic Publishers, 2000]

Lipoproteins possess different densities |——| the higher the lipid percentage, the lower is the density and the larger is the diameter. According to their separation by utlracentrifugation, lipoproteins are classified into chylomicrons, very low-density lipoproteins, (VLDL), low-density lipoproteins (LDL), intermediate-density lipoproteins (IDL) and high-density lipoproteins (HDL). Compositions of different types of lipoproteins are smmarized in Table 6.1.

Table 6.1 Compositions of different types of lipoproteins

	Chylomicron	VLDL	IDL	LDL	HDL
Diameter (nm)	80-100	30-85	25-34	20-25	6-20
Density (g/cm^3)	<0.95	0.96-1.006	1.006-1.019	1.020-1.063	1.063-1.210
Proteins (%)	1.5-2.5	8-10	16-20	21-25	43-55
Triacylglycerols (%)	85-90	50-60	20-24	8-10	4-6
Phospholipids (%)	6-9	16-18	22-24	15-20	20-40
Cholesterylesters (%)	3-4	11-15	30-32	36-40	10-12
Free cholesterol (%)	1-2	4-10	7-8	7-10	3-5
Apoproteins	A-I, A-II, A-IV B-48, C-I, C-II C-III, E	B-100, C-I C-II, C-III, E	B-100, C-III C-III, E	B-100	A-I, A-II, A-IV C-I, C-II, C-III D, E

6.5.2 Functions of plasma lipoproteins

(a) Chylomicrons :

They transport mainly *triacylglycerols and* small amounts of cholesteryl esters, phospholipids and free cholesterol from the interstine to the live, adipose tissue and muscles. The carry both dietary traiacylglycerols and triacylglycerols synthesized in enterocytes from the intestine to the *adipose tissue* for *storage as fats, and to some extent to muscles* also. They are also involved in delivering *exogenous* (dietary) *cholesterol* and its esters to hepatocytes.

(b) VLDL :

They mobilize mainly *endogenous triacylglycerols,* synthesized in hepatocytes, from the liver to the *adipose tissue* in particular for storage as fats; smaller amounts may also be transported to other extrahepatic tissues like *muscles.* High-carbohydrate diets, high insulin : glucagon ratio or any other factor raising the triacylglycerol synthesis in hepatocytes enhances the synthesis of VLDL in the liver to accelerate the transfer of endogenous triacylglycerols from hepatocytes to adipocytes.

(c) LDL :

They transport mainly the *endogenous cholesterol* and its esters, synthesized by hepatocytes, and some *exogenous* (dietary) *cholesterol* also to *extrahepatic tissues* such as muscles glands, connective and lymphoid tissues. So, high plasma LDL may lead to cholesterol deposition in the vascular walls of such tissues.

(d) BHDL :

The serve mainly as *cholesterol scavangers* by transporting cholesterol and cholesteryl esters back from peripheral extrahepatic tissues to liver, diminishing thereby the possibility of cholesterol plaque fromation in peripheral vascular walls. HDL particles also collec *apo-A, apo-C* and *apo-E* proteins from other types of lipoprotein particles, recycle some of those apolipoproteins by transferring them to VLDL and chylomicrons, and return the remainder to the liver.

6.5.3 Metabolism of plasma lipoproteins

(a) Chylomicrons and VLDL :

Polysomes on the rough ER membranes of hepatocytes and enterocytes translate B-100 and B-4 apolipoproteins repectively. They are subsequently incorporated into 'nascent' lipoprotein particles assembled from triacylglycerols, cholesterol and phospholipids in the smooth ER of those cells. After the glycosylation of the apolipoproteins in the Golgi cisternae, secretory vesicles contaning the 'nascent'

lipoproteins are budded of from those cistermae and translocated to the cell membrane, where the vesicles fuse with the latter to release 'nascent' VLDL and chylomicrons respectively from hepatocytes and enterocytes. VLDL particles thereafter enter the blood by passing through the fenestrated walls of hepatic sinusoids; chylomicrons first enter the lymph and through it, pass into the blood. In the blood, these 'nascent' lipoproteins mature to VLDL and chylomicrons by receiving apo-C and apo-E proteins from circulating HDL particles (Fig. 6.7).

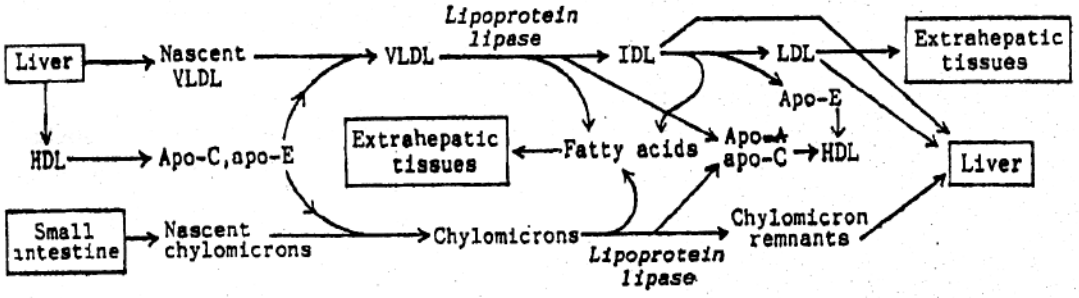


Figure 6.7 : Metabolic fates of Chylomicrons and VLDL [From D. Das, Biochemistry, Academic Publishers, 2000]

In course of their circulation in the blood, VLDL and chylomicron particles come to stick to the capillary endothelial membrane of adipose tissue, mammary glands, cardiac and striated muscles, lungs, renal medulla and other extrahepatic tissues. *Lipoprotein lipase* of the endothelial membrane then hydrolyzes the triacylglycerols of adherent lipoprotein particles through di-and mono-acylglycerols to fatty acids and glycerol. Most of the released fatty acids are taken up by extrahepatic tissue cells, the *adipose tissue* ordinarily taking up the major bulk of those fatty acids for the synthesis and storage of triacylglycerols. The differential uptake of fatty acids from these lipoproteins by different extrahepatic tissues at different times, is accomplished largely by *regulating the lipprotein lipase*. Phospholipids and apo-CII proteins of VLDL and chylomicrons activate this enzyme for its short-term regulation; its long-term regulation depends on the induction of its synthesis by insulin. Lipoproein lipase activity rises in adipocytes after food intake and declines on starvation; but starvation augments its activity in striated and cardiac muscles. So, adipocytes decrease their uptake of fatty acids from lipoproteins during starvation, sparing them for muscles while the latter still continue such uptake from lipoproteins at a higher rate due to enhanced lipoprotein lipase activity there. Prolonged high calory-intake brings about a significant rise in adipose tissue fat due to the induction of the lipase by insulin secreted in that condition.

Lipoprotein lipase action on VLDL and chylomicron particles decreases |—| halves — their triacylgylcerol percentages and diameters, the enhances the percentages

of cholesterol and cholesteryl esters about tow-fold, also increase the phospholipid percentage to about 1.5 times, and transfers some of their apo-A and apo-C proteins to the plasma HDL particles. VLDL and *chylomicron remnants* which get released from the end thelial membrane to reenter circulation.

The circulating chylomicron remmants ultimately bind by their apo-E proteins to *remmant-receptors* (E-receptors) on the hepaocyte membrane, thereby get concentrated on the latter, and are taken up by *adsorptive pinocytosis* into hepatocytes for further metabolism of triacylglycerols, cholesterol and its esters. Some IDL particles are also removed from the circulation in the same way.

(b) Low-density lipoproteins (LDL) :

Most of the circulating IDL particles lose their apo-E proteins by their transfer to HDL particles. They also lose some triacylglycerols to hepatocytes during their intermittent bindings to the E-respectors on the hepatocyte membrane (Fig. 6.7). These change the IDL particles to *LDL particles,* possessing somewhat smaller diameters, higher densities, lower percentages of total lipids, triacylglycerols and phospholipids, but *higher cholesteryl ester percentages* than IDL particles. As LDL circulates in the blood, it acts a *vehicle for delivering cholesterol,* particularly to extrahepatic tissues such as lymphoid, muscular and vascular connective tissues. Most circulating LDL particles get gradually concentrated on those tissues cells by the binding of their *B-100 apoproteins* to specific *LDL receptors* (B-100 receptors) on the cell membrane and are then internalized into those cells by *adsorptive pinocytosis.* The remaining circulating LDL particles may be similarly pinocytized by hepatic cells, and their cholesterol is mostly excerted in bile as cholesterol, its esters, and bile acids formed from them in hepatocytes.

(c) High-density lipoproteins (HDL) :

HDL particles are formed in hepaocytes and enterocytes by the polysomal

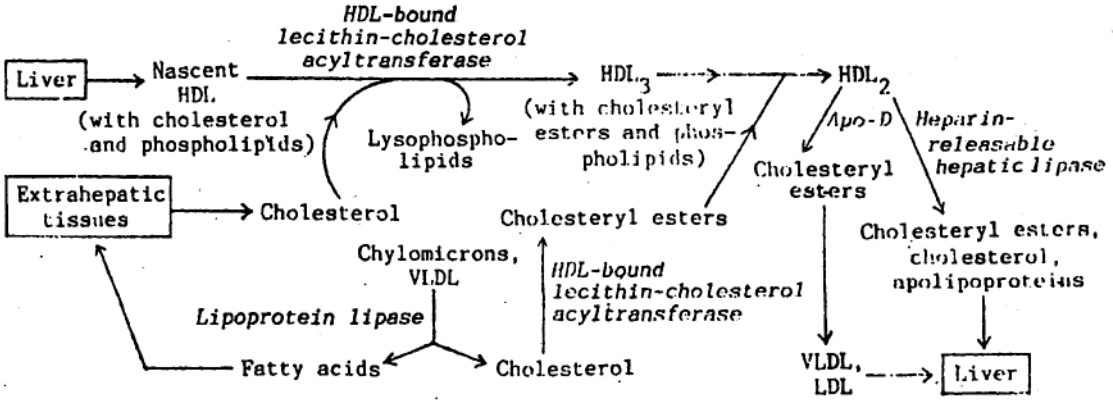


Figure 6.8 : Metabolic fates of HDL [From D. Das, Biochemistry, Academic Publishers, 2000]

translation of A, C and E apoproteins on their rough ER membranes and their incorporation into the 'nascent' assembly of triacylglycerols, phospholipids, cholesterol and its esters, made in the smooth ER (Fig. 6.8). The 'nascent' HDL particle thus constituted is usually a discoid or elliptical particle composed largely of bilayer of phospholipids and cholesterol along with some triacylglycerols and the apoproteins.

As the 'nascent' discoid HDL particles circulate in the blood, their A-I and C-I apolipoproteins activate plasma *lecithin-cholesterol acyltransferase* to transfer unsaturated acyl groups |——| particulary linoleyl goups |——| from the phosphatidylcholine molecules of HDL to the cholesterol molecules of its surface monolayer, esterifying amphipathic cholesterol into nonpolar cholesterylesters, mainly cholesteryl linoleate; phosphaticholine is simultaneously converted to lysophosphatidylcholine by the LCAT action.

Cholesterol + phosphatidylcholine → cholesteryl enter + lysophosphatilylchaline

While lysophosphatidylcholine formed in this reaction is released from HDL to the plasma, the change of amphipathic cholesterol to nonpolar cholesteryl ester makes the latter sink from the surface monolayer of the HDL particle to its nonpolar central core. Such reactions progressively change the elliptical or discoid 'nascent' HDL particles into more and more spherical HDL_3 particles with an amphipathic lipid-protein monolayer enclosing a nonpolar central lipid core. Simultaneously, the decline in cholesterol concentration in the surface monolayer leads to the progressive entry of free cholesterol from the plasma, other circulating lipoprotein particles and extrahepatic tissues into the surface monolayer of HDL particles. Thus, circulating HDL particles *gather cholesterol* from palsma and extrahepatic tissues, change it to cholesteryl esters with the help of LCAT, and carry them in its central lipid core. This changes the 'nscent' HDL particle stepwise into HDL_3 and HDL_2 particles with progressively larger diameters, lower densities, lower triacylglycerol percentage and higher percentage of cholesteryl esters in them. Some C and E apoproteins are transferred from HDL to circulating chylomicrons and VLDL particles. Some cholesteryl esters are also transferred from HDL to VLDL particles with the help of apo-D of HDL, acting as the *cholesteryl ester transfer protein.*

In this way, HDL particles, serve to *mop up cholesterol* from extrahepatic tissues and plasma. From the circulation, HDL particles get bound to specific HDL receptors on the membrane of liver cells and are then taken up by those cells by *adsorptive pinocytosis.* Cholesterol and cholesteryl esters of HDL are then eleminated in the bile either as such or after catabolism into bile acids.

6.5.4 Lipoproteins and atherosclerosis

Because triacyglycerols are mobilized from the liver to the extrahepatic tissues mainly in VLDL particle, a prolonged rise in plasma VLDL produces a sustained rise ain plasma triacylglycerols.

Cholesterol is mobilized from the liver to extrahepatic tissues mainly in LDL particles, and on the contrary, is collected from plasma and extrahepatic tissues mainly by HDL particles and brought back to liver. So, a sustained rise in the plasma LDL, LDL-cholesterol or LDL-cholesterol : HDL-cholesterol ratio, as may happen in habitual high-cholesterol intake, nicotine abuse, hypothyroidism and diabetes mellitus, results in the protracted rise in serum cholesterol level. Repression of synthesis of LDL-receptors (B-100 receptors) resulting from high dietary intake of cholesterol also heightens the plasma LDL-cholesterol by decreasing the removal of LDL from the blood. *Familial hypercholesterolemia* with high plasma LDL-cholesterol is an inborn error of LDL metabolism in humans, resulting from a dominant autosomal genetic defect of LDL-receptors.

High LDL-cholesterol or high LDL : HDL ratio in the plasma leads to the deposition of cholesterol-enriched lipids in vascular walls in the form of fibrous plaques, with consequent thickning and roughening of arterial walls (*altherosclerosis*) and formation of blood clots. Cholesterol plaques and blood clots gradually occlude the vessel, decreasing and finally stopping the blood flow in the affected tissue (*infarction*). Coronary occlusions and myocardial infarctions arise in this way.

High plama level of HDL and HDL-cholesterol would enhance the *scavnging of cholesterol* from extrahepatic tissues and plasma, and tend to lower serum cholesterol. Indeed, high dietary cholesterol has been fond to enhance the formation of HDL; the apo-E protein of the latter wold bind to the E-receptors on hepatocytes, followed by the adsorptive pinocytosis of HDL particle by the liver cells for cholesterol catabolism and excretion. By lowering serum cholesterol HDL prevents the formation of cholesterol plaques in vascular tissues.

6.6 Fatty Liver

Failure in synthesizing required amounts of phospholipids and VLDL in the liver decreases the mobilization of both dietary and endogenous lipids from the liver to extrahepatic tissues including the adipose tissue. The liver consequently gets loaded with accumulated lipids (*fatty liver*) and suffers from fibrosis, cirrhosis and functional failures.

(*a*) Fatty liver may be produced if the liver either receives from food or synthesizes

far more lipids than the amounts of phospholipids and VLDL it can sunthesize for their mobilization to extrahepatic tissues including the adipose tissue. This may happen due to sustainged high-fat diets or heightened adipose tissue lipolysis—free fatty acids rise in the plasma, leading to their enhanced hepatic uptake and conversion to fat which cannot be fully mobilized to adipose tissue in absence of the required enhance amount of VLDL.

(*b*) Fatty liver may result form a failure in hepatic formation of VLDL owing to the dietary deficeincy of essential fatty acids or choline which are constituents of the phospholipids of VLDL particle, or of methionine used as a methyl donor for choline synthesis in the liver. Choline, betaine, methionine, inositol and essential fatty acids are known as *lipostropic factors* as they contribute to the synthesis of phospholipids and VLDL, thus promoting fat mobilization from the liver to the adipose tissue.

(*c*) Protraced high cholesterol intake may also result in fatty liver, because cholesterol is esterified with essential fatty acids, particularly linoleic acid, in the body and therby decreases the availability of essential fatty acids for phospholipid and VLDL synthesis.

(*d*) Toxic effects of chloroform, carbon tetrachloride, puromyci, arsenic, lead and phosphorus prevent the synthesis of apo-B-100 in the liver, which is a constituent of VLDL. This leads to a failure in VLDL formation and a consequent production of fatty liver. Carbon tetrachloride also produces free radicals which peroxidate the lipids of hepatic smooth ER membranes and consequently prevent the formation and release of VLDL.

(*e*) Orotate prevents the release of VLDL from hepatocytes by inhibiting the glycosylation of apo-B-100 of VLDL in the hepatic Golgi cisternae, thus leading to fatty liver.

6.7 Summary

Adipose tissue plays a pivotal role in the metabolism of lipds. Dietary fatty acids, in excess of their immediate requirement for energy production, are used in synthesizing triacylglycerols in adipocytes and stored as adipose tissue. Even carbohydrate metabolites, in excess of their immediate need for energy production, are largely converted to facts and stored in adipose tissue. When needed for energy production, the stored triacylglycerols of adipose tissue undergo lipolysis to release fatty acids for beta-oxidation and energy production. The rate of adipose tissue lipolysis is controlled by regulating the synthesis and activity of its hormone-sensitive rate-limiting enzyme, triacylglycerol lipase.

Nonpolar lipids are transported in aqueous body fluids between the liver, small intestine and other extrahepatic tissues including the adipose tissue by forming phospholipids and lipoproteins. Phospholipids are synthesized in many tissues such as the liver, small intestinal mucosa and adipocytes, and used in forming lipoproteins. Lipoprotein particles are assembled with triacylglycerols, fatty acids, cholesterol and its esters in hepatocytes and enterocytes, with a surface monolayer of amphipathic phospholipids; some intergral and peripheral apolipopreteins are also incorporated into the lipoprotein particles.

According to their densities, lipoproteins are categorised into chylomicrons, very low-density lipoproteins low-density lipopreteins, intermediate density-lipoproteins and high-density lipoproteins. Chylomicrons mainly transport dietary triacylglycerols as well as triacylglycerolglycerols synthesized enterocytes through the lymph and blood from small intestine to adipose tissue and live. VLDL mainly carry triacylglycerols, snthesized inhepatocytes, through the blood to extrahepatic tissues including the adipose tissue. LDL acts as a vehicle for transporting cholesterol and its esters, synthesized in the liver, and also some dietary cholesterol, through the blood to extrahepatic tissues such as connective, lymphoid and vascular tissues. Excess of serum LDL-cholesterol can lead to the formation of cholesterol plaques on vessel walls, and corronary occlusions. HDL functions mainly as cholesterol scavangers, mopping up cholesterol and its esters from peripheral tissues and circulation and delivering them to the liver for catabolism or biliary excertion. HDL thus decreases the tendency of formation of cholesterol plaques in blood vessels. The courses of circulation and metabolism of different plasma lipoproteins have been elaborated in this unit. Fatty liver results from high accumulation of lipids in the liver due to the failure in VLDL formation for their mobilization to extraheptic tissues because of various reasons like dietary deficiencies of essential fatty acids and methionine, high intakes of fats and cholesterol, and exposure to toxic chemical. The uptake of fatty acids from VLDL and chylomicrons by extrahepatic tissues depends upon the hydrolysis of their triacyglycerols by regulated enzyme, lipoprotein lipase, of the endothelial membrane.

6.6 Terminal Questions

1. (a) Describes how stored fats are hydrolyzed into fatty acids in the adipose tissue, mentioning the enzyme actions and using a suitable flow chart.
   (b) Name the rate-limiting enzyme for adipose tissue lipolysis and describe its regulation *in vivo*.
   (c) Name some lipotropic factors and their significance.

2. (a) Summarize the biological roles of plasma lipoproteins.
   (b) Describe the circulation and metabolism of plasma HDL and LDL with suitable flow charts.
   (c) Discuss how LDL-cholesterol and HDL-cholesterol are related to atherosclerpsis.
3. (a) Write how fatty acids are enzymatically thioesterified prior to their use in metabolic roles.
   (b) Describe using suitable flowcharts how triacylglycerols are synthesized by the enzymes of glycerophosphate and dihydroxyacetone phosphate pathways.
   (c) What are the significances of triacylglycerol synthesis?
4. (a) Describe using a suitable flowchart how phospholipids are synthesized, mentioning the actions of the enzymes involved.
   (b) Explain the significances of phospholipid synthesis in fat mobilization.
   (c) Discuss various factors which can give rise to fatty liver.
5. (a) Describe the structure of a typical lipoprotein particle with a suitable sectional sketch and mentioning its principal constituents.
   (b) Give a brief classification and average composition of plasma lipoproteins of different classes.
   (c) How are chylomicrons and VLDL particles circulated in the plasma and metabolized? Describe it with a suitable flow chart.
6. (a) Discuss using suitable flowcharts how different plasma lipoproteins are circulated and metabolized.
   (b) Explain briefly how he LDL : HDL ratio in the plasma influences of deposition in cholesterol in peripheral vessels.
7. Write notes on the following :
   (a) Lipoprotein lipase
   (b) Hormone-sensitive adipose tissue lipase.
   (c) Chylomicrons.
   (d) Lipotropic factors.

6.9 Answers

1. (a) See Section 6.3
   (b) See Section 6.3
   (c) See paragraph (b) of Section 6.6

2. (a) See Section 6.5.2
   (b) See (b) and (c) of Section 6.5.3
   (c) See Section 6.5.4
3. (a) See Section 6.2.1
   (b) See Sections 6.2.2 and 6.2.3.
   (c) See second paragraph of Section 6.2
4. (a) See Sections 6.4.1 and 6.4.2
   (b) See Section 6.4.3
   (c) See Section 6.6
5. (a) See Section 6.5.1
   (b) See Section 6.5.1 and Table 6.1
   (c) See (a) of Section 6.5.3
6. (a) See Section 6.5.3
   (b) See Section 6.5.4
7. (a) See second paragraph of (a) of Section 6.5.3
   (b) See Section 6.3
   (c) See Sections 6.5.1 – 6.5.3
   (b) See third paragraph of Section 6.6

Unit 7 ❐ Genetic Disorders of Phenylalanine, Tyrosine and Glycogen Metabolisms

Structure

7.1 Introduction

You have already read in the preceding units of Group A (2/2) about some important pathways of metabolism of carbohydrates, lipids, proteins and purines. You have also learnt about the kinetics, modulations, inductions and repressions of enzymes.

In this unit, you will read about the discorders of the metabolisms of phenylalanine, tyrosine and glycogen owing to genetic errors.

Genetic errors may result offen from the substitution, omission or addition of one or more mucleotides in the DNA strand of a single gene, and are then called *single-gendisorders* such as albinism. In contrast, *polygenic traits* such as clift lip-and-palate may arise from simultaneous mutations in more than one gene. The genetic disorder may be inherited as either a modinant or a recesive or a codominant trait according as the mutant gene is a Mendelian *dominant, recessive* or codominant one; examples of such traits are respecitively familial hypercholesterolemia, albinism and sickle cell anemia. A mutant gene located in autosome is called and *autosomal trait.* Genetic mutation of a gene affects the protein coded by it and consequently produces disorders in the synthesis, functions and metabolism of such substrates as need that protein for their syntheses and biological roles. Here, in this unit, you will read about such genetic errors of two amino acids and a polysaccharide.

Objectives

Study of this unit should enable you to :

- Understand the natures of inherited defects of two important amino acids, viz., phenylalanine and tyrosine,
- Explain how such genetic errors of these two amino acids affect their metabolism,
- Describe the symptoms of the consequent diseases,
- Know about the causes of important inherited defects of glycogen metabolism.
- Explain how such genetic errors affect the storage of glycogen in animal tissues,
- Describe the principal symptoms of such genetic errors of glycogen metabolism.

7.2 Genetic Errors of Phenylalanine and Tyrosine Metabolisms

You may recall that phenylalanine may be directly incorporated in peptide chains as their phenylalanine residues during translation of tissue proteins. But for all other metabolic functions and fates, phenylalanine has to be converted first to tyrosine which then may be used in synthesizing melanins, catecholamines and iodotyrosines or in being catabolized to fumarate and acetoacetate.

Genetic errors of these two amino acids result from mutations of genes encoding different proteins involved in their metabolism. Such errors are inherited as life-long metabolic defects because of the abnormal or missing proteins whose synthesis and functions have been affected by mutations of their genes. An outline of such genetic errors of Phe and Tyr metabolism is given below.

7.2.1 Hyperphenylalaninemias

Inborn failures to convert phenylalanine to tyrosine many result from genetic defects of several proteins involved in this conversion (Fig. 7.1). In each such case, the inability or deficiency in changing phenylalanine to tyrosine prevents the normal metabolism of phenylalanine through tyrosine in some way or other. Consequently, blood phenylalanineries (hyperphenylalaninemia) and blood tyrosine declines.

(a) ***Hyperphenylalaninemia type I (phenylketonemia) :***

This serius genetic error of phenylalanine metabolism is inherited as an autosomal recessive defect and consists of an inborn deficiency of *phenylalanine hydroxylase,* the enzyme catalyzing the hydroxylation of phenylalanine to tyrosine using molecular O_2 and *tetrahydrobiopterin* as an electron-donor. Failure to change phenylalanine to

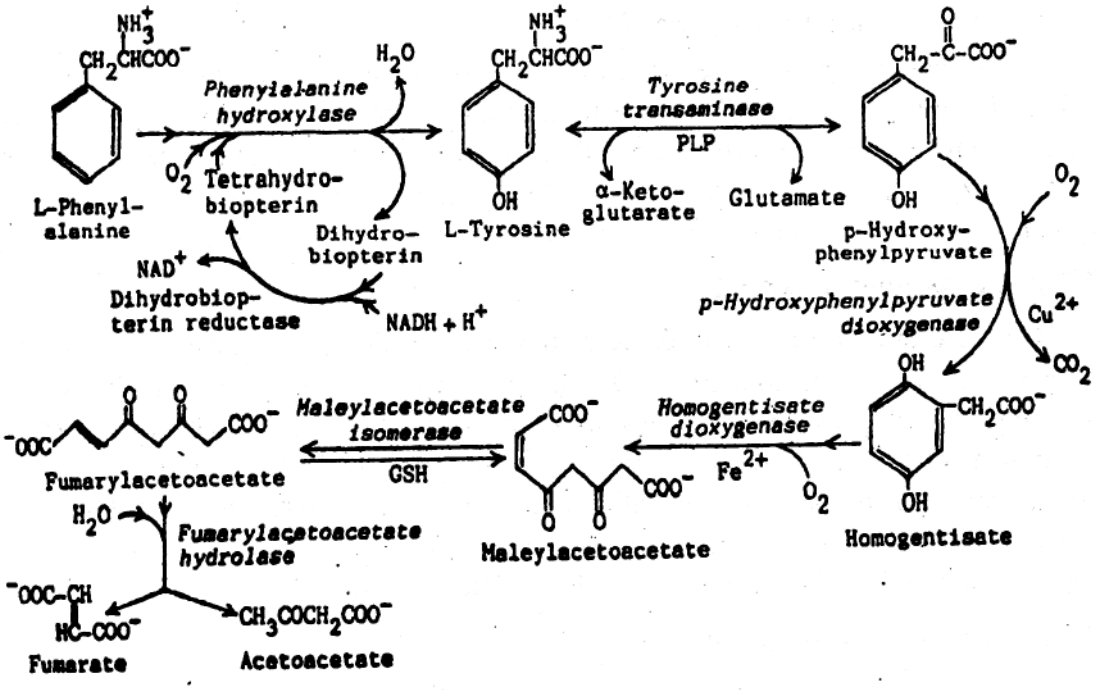


Figure 7.1 : Catabolism of phenylatemine to fumarate and acetoacetate [From D. Das, Biochemistry, Academic Publishers, 2000]

tyrosine leads to high blood levels are urinary eliminations of phenylalanine and its abnormal metabolites such as phenylpyruvate and phenyllactate, severe mental retradation, very low IQ, severe psychosis, seizures and eczematons skin lesions. High blood phenylalanine inhibits *tyrosine hydroxylase* which would have normally hydroxylated tyrosine to dopa for its further metabolism; poor availability of tyrosine for forming melanin pigments leads to decreaed pigmentations of skin and hair.

(b) Hyperphenylalaninemia types II and III :

These rare genetic errors arise from inborn defects of *dihydrobiopterin reductase* which normally would have used NADH to reduce dihydrobiopterin back to tetrahydrobiopterin for continuing phenylalanine hydroxylase activity (Fig. 7.1). Evidently, any genetic defect of this reductase decreases the availability of tetrahydrobiopterin and consequently blocks the hydroxylation of phenylalanine to tyrosine. Symptoms include high blood and urinary levels of phenylalanine, phenylpyruvate and phenyllactate, low blood tyrosine, severe, mental retardation, neurological symptoms, psychosis and seizures. Because tetrahydrobiopterin is also the electron-donor for the synthesis of noradrenaline and serotonin, the decreased synthesis of these neurotransmitters also contribute to neurological symptoms.

(c) Hyperphenylalaminemia types IV and V :

These rare inborn errors results from a genetic deficiencies of *some enzymes* for *dihydrobiopterin synthesis.* This leads to the nonavailability to tetrahydrobiopterin as an electron-donor for converting phenylalanine to tyrosine and consequently blocks phenylalanine metabolism. Symptoms include hyperphenylalaninemia and other symptoms of phenyketonuria, and neurological lesions.

7.2.2 Tyrosinemia type II

This rare inborn error of tyrosine metabolism results from a genetic deficiency of *tyrosine transaminase,* a PLP-dependent enzyme for converting tyrosine to p-hydroxy-phenylpyruvate by transamination (Fig.7.1). This blocks the normal catabolism of tyrosine, and so of phenylalanine also, resulting in high blood and urinary levels of tyrosine, urinary elimination of abnormal tyrosine catabolites like tyramine and N-acetyltyrosine, neuromuscular incoordination, mental retardation, skin lesions, opthalmic nerve lesion and self mutilating tendency.

7.2.3 Neonatal tyrosinemia

This genetic disorder of tyrosine metabolism results from the inborn defect of *p-hydroxpyphenylpyruvate dioxygenase,* the enzyme for oxidizing p-hydroxyphenylpyruvate (see 7.2.2) to homogentisic acid at the next step (Fig. 7.1). Consequently, normal catabolisms of both tyrosine and phenylalanine are blocked, leading to their high blood levels and urinary eliminations, enhancement of their alternative Catbolic pathways, and the resulting urinary eliminations of their abnormal catabolites, such as tyramine, N-acetyltyrosine p-hydroxyphenyllactate and p-hydroxyphenylacetate.

7.2.4 Alkaptonuria

This is a relatively mild autosomal recessive trait. It results from a genetic defect of *homogentisate dioxygenase* the enzyme for oxidizing homogentisate to maleylacetoacetate at the next step of tyrosine catabolism. As this blocks the catabolism of homogentisate, the latter is deposited in connective tissues like cartilages, causing their abnormal pigmentation and leading to rheumatoid arthritis in advanced age. Homogentisate, eliminated in the urine, gets oxidized to a block pigment on expopsure to air, thus turning the urine black.

7.2.5 Tyrosinemia type I

This inborn error of tyrosine catabolism is also called *tyrosinosis.* It produces symptoms like high blood and urinary tyrosine levels, gastrointestinal symptoms

such as nausea, vomiting and diarrhoea, and ultimately death due to hepatic nonfunction. It results from the genetic deficiency either of *maleylacetoacetate isomerase* for isomerizing maleylacetoacetate, a tyrosine catabolite, to fumary lacetoacetate, or of *fumarylacetoacetate hydrolase* which normally hydrolyzes the latter to fumarate and acetoacetate (Fig. 7.1).

7.2.6 Albinisms

These are mild genetic errors of tyrosine metabolism. Melanins, protein-bound polymeric pigments of skin, hair and pigment cells of iris, choroid and retina, are synthesized form tyrosine by a copper enzyme *tyrosinase* in sunlight.

Tyrosine → Dopa → Dopa quinone → Melanins.

Genetic deficiences of malanins may cause two types of albinism with a lack of melanin pigmentation of skin, hair and eye tissues, and the formation of melanomas.

(a) Oculocutaneous albinism :

This autosomal recessive trait results mostly from a genetic dificiency of *tyrosinase.* Skin, hair and eyes have almost no melanin pigmentation in most cases; photophobia is also common and results from the lack of iris pigment. However, some cases may have light cutaneous and iris pigmentation.

(b) Ocular albinism :

This genetic error of melanin formation is inherited as either X-linked or an autosomal trait. Melanin is absent in retina only, but is present in skin and hair. The metabolic defect underlying it is not properly known.

7.3 Genetic Errors of Glycogen Metabolism

Glycogenosis or *glycogen storage diseases* are inherited defects of glycogen metabolism due to the genetic deficiencies of several enzymes for synthesis and breakdown of glycogen.

7.3.1 Type I. von Gierke's disease

You may recall that liver stores glycogen which, when needed, is converted by *glycogenolysis* to glucose for mobilization by blood to extrahepatic tissues. *Type I glycogenosis* is an inborn disorder of glycogenolysis due to a genetic dificiency of *glucose 6-phosphatase,* which catalyzes the hydrolytic dephosphorylation of glucose 6-phosphate to glucose at the final step of that pathway. In this disease, a genetic defect of that enzyme is inherited as an autosomal recessive trait and causes a failure in mobilizing stored glycogen as glucose to extrahepatic tissues. This results in the storage of too much glycogen in hepatic and renal cells, very low blood glucose,

massive hepatic enlargemenmt, no rise in glycogenolysis and consequently in blood sugar even after adrenaline administration, hyperlipidemia and Ketosis—the last two symptoms are due to greater adipose tissue lipolysis and higher fatty acid oxidation because of insufficient glucose supplied to tisues for energy production.

7.3.2 Type II. Pompe's disease

This autosomal recessive trait is characterized by a genetic deficiency of *lysosomal α-1, 4-glucosidases* which normally participate in the lysosomal hydrolysis of polyglycans including glycogen. The defect results in excess glycogen accumulation in tissues and causes infant mortality from respiratory and cardiac failures.

7.3.3 Type III. Cori's disease or limit dextrinosis

In this autosomal recessive trait, *amylo-1, 6-glycosidase* activity of the *glycogen-debranching enzyme* is genetically deficient n the liver. This enzyme activity normally hydrolyzes the α-1, 6-glycosidic bonds at the branching points of glycogen molecules during glycogenolysis. So, a genetic deficiency of this activity causes a failure of hydrolysis of glycosidic bonds at the branching points of glycogen chains. So, *glycogenolysis* remains incomplete, limit-dextrin-type highly-branched polysaccharide molecules accumulate in hepatic cells and muscles, blood sugar remains low, and the liver gets enlarged.

7.3.4 Type IV. Adersen's disease or amylopectinosis

You have read in Unit 3 how glucose is changed to glycogen in the liver, muscles and many other tissues by the processes of *glycogenesis.* In the autosomal recessive trait called amylopectinosis, a genetic deficiency of *amylo-1 4 → 1, 6-transglycosylase* (branching enzyme) of that pathway causes a failure in forming α-1 6-glycosidic bonds between growing oligosaccharide chains and a consequent inability to introduce branchings in glycogen molecules. As a result, abnormal amylose-like unbranched long-chain polysaccharide molecules accumulate in liver, muscles and spleen, leading to infant mortality due to either hepatic cirrhosis or cardiac failure.

7.3.5 Type V. McArdle's disease

You have read about the EMP pathway of *glycolysis* of glycogen in muscles. In type V glycogenosis, an autosomal recessive trait, a genetic deficiency of the *muscle glycogen phosphorylase,* the very first enzyme of theat pathway, birngs about in inability to *glycolyze* muscle glycogen. This results in an increase in muscle glycogen, poor post-exercise blood lactate loevel, severe muscle cramps and inability for heavy exercise.

7.3.6 Type VI. Hers' disease

In this autosomal recessive trait, liver fails to carry out *glycogenolysis* of liver glycogen to glucose because of a genetic error of the *hepatic glycogen phosphorylase* the very first and the rate-limiting enzyme of that pathway. This leads to the accumulation of excessive glycogen in the liver, hepatic enlargement, low blood sugar due to decreased mobilization of glucose from liver into blood, poor hyperglycemic response to epinephrine administration, and Ketosis owing to enhanced fatty acid oxidation because of insufficient availability of glucose for energy production.

7.3.7 Type VII. Taru's disease

Phosphofructokinase I, the rate-limiting enzyme for glycolysis, is genetically deficient in muscles and erythrocytes in this autosomal recessive trait. The consequent failure of glycolysis in those tissus raises the muscles glycogen, decreases the blood lactate, towers the muscular work capacity, and produces muscles camps.

7.3.8 Type VIII. Phosphorylase Kinase deficiency

This disease, inherited as a sec-linked recessive trait, is characterized by a genetic deficiency of the *hepatic phosphorylase kinase* which normally would have phosphorylated and activated the rate-limiting *glycogenphosphorylase* of *glycogenolysis.* This results in poor hepatic glycogenolysis, consequent accumulation of glycogen in the river, hepatic enlargement and mild decline in the blood glucose level.

7.3.9 Type IX. Glycogen synthase deficiency

In this autosomal recessive trait, the rate-limiting *hepatic glycogen synthase* of glycogenesis in the liver bring about a failure in *hepatic glycogenesis* and a consequent decline in the liver glycogen content.

7.4 Summary

Five types of hyperphenylalaninemias result from the genetic errors of different enzymes/cofactors participating in phenylalanine metabolism. They are characterized by high blood and urinary levels of phenylalanine and its abnormal metabolities as well as neurological and mental symptoms.

Out of the genetic disorders of tyrosine metabolism, caused by the genetic errors of different enzymes of the latter, albinisms are characterized by lack of melanin

pigmentation of skin, hair and eye tissues due to genetic failures in melanin synthesis, alkaptonuria is characterized by black urine containing a tyrosine metabolic called homogentisate, and three types of tyrosinemias show symptoms including high blood and uninary levels of tyrosin, and its abnormal metabolites.

Nine types of glycogenoses result from genetic errors of glycogen metabolism and affect the storage of glycogen in liver and extrahepatic tissues. Glycogenoses of types I, III, VI and VIII result from genetic errors of different glycogenolytic enzymes, types IV and IX from genetic deficiencies of glycogenic enzymes, types V and VII from genetic errors of glycolytic enzymes, and type II from genetic defects of lysosomal enzymes for glycogen breakdown.

7.5 Terminal Questions

1. (a) Describe the genetic causes and symptoms of different types of tyrosinemia.
   (b) What is glycogenosis? Describe type IX glycogenosis.
   (c) Describe the inbron errors of glycogen storage owing to genetic defects of glycogenolysis.
2. (a) Discuss the causes and symptoms of different types of albinism.
   (b) Describe the inborn errors of glycogen metabolism due to genetic defects of different enzymes for glycogenesis.
   (c) Describe the genetic cause and symptoms of Pompe's disease.
   (d) What are the symptoms and causes of alkaptonuria.
3. (a) Discuss different inborn deficiencies of metabolism of phenolic amino acids, their genetic causes and effects.
   (b) Give an account of the inborn errors arising from genetic defects in the branching and debranching of glycogen chains.
4. (a) Describe the negetic causes and symptoms of such glycogen storage diseases which result from inherited defects of glycolysis.
   (b) Give an account of the genetic causes and symptoms of different types of hyperphenylalaninemia.

7.6 Answers

1. (a) See Sections 7.2.2, 7.2.3 and 7.2.5.
   (b) See first paragraph of Section 7.3 and 7.3.9.
   (c) See Sections 7.3.1, 7.3.3, 7.3.6 and 7.3.8.
2. (a) See Section 7.2.6.
   (b) See Sections 7.3.4 and 7.3.9.
   (c) See Section 7.3.2.
   (d) See Section 7.2.4.
3. (a) See Sections 7.2.1-7.2.6.
   (b) See Sections 7.3.3 and 7.3.4.
4. (a) See Sections 7.3.5 and 7.3.7.
   (b) See Section 7.2.1.

PREFACE

In the auricular structure introduced by this University for students of Post-Graduate degree programme, the opportunity to pursue Post-Graduate course in Subject introduced by this University is equally available to all learners. Instead of being guided by any presumption about ability level, it would perhaps stand to reason if receptivity of a learner is judged in the course of the learning process. That would be entirely in keeping with the objectives of open education which does not believe in artificial differentiation.

Keeping this in view, study materials of the Post-Graduate level in different subjects are being prepared on the basis of a well laid-out syllabus. The course structure combines the best elements in the approved syllabi of Central and State Universities in respective subjects. It has been so designed as to be upgradable with the addition of new information as well as results of fresh thinking and analysis.

The accepted methodology of distance education has been followed in the preparation of these study materials. Co-operation in every form of experienced scholars is indispensable for a work of this kind. We, therefore, owe an enormous debt of gratitude to everyone whose tireless efforts went into the writing, editing and devising of a proper lay-out of the materials. Practically speaking, their role amounts to an involvement in invisible teaching. For, whoever makes use of these study materials would virtually derive the benefit of learning under their collective care without each being seen by the other.

The more a learner would seriously pursue these study materials the easier it will be for him or her to reach out to larger horizons of a subject. Care has also been taken to make the language lucid and presentation attractive so mat they may be rated as quality self-learning materials. If anything remains still obscure or difficult to follow, arrangements are there to come to terms with them through the counselling sessions regularly available at the network of study centres set up by the University.

Needless to add, a great deal of these efforts is still experimental—in fact, pioneering in certain areas. Naturally, there is every possibility of some lapse or deficiency here and there. However, these do admit of rectification and further improvement in due course. On the whole, therefore, these study materials are expected to evoke wider appreciation the more they receive serious attention of all concerned.

Professor (Dr.) Manimala Das
Vice-Chancellor

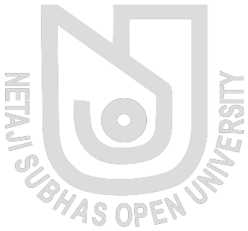
First Reprint : December, 2008

Printed in accordance with the regulations and financial assistance of
the Distance Education Council, Government of India

POST GRADUATE ZOOLOGY
[M.Sc]

PAPER : GROUP
PGZO - 3 : B

	Writer	Editor
Part-I		
Units 1-4	Dr. Kamalesh Mishra	
Units 5-8	Prof. Chandrasekhar Chakraborty	Prof. Chandrasekhar Chakraborty
Units 9-12	Prof. Anil Kumar Saha	
Part-II		
Units 1-8	Dr. Sanjib Kumar Das	Dr. Subir Chandra Dasgupta


NETAJI SUBHAS OPEN UNIVERSITY

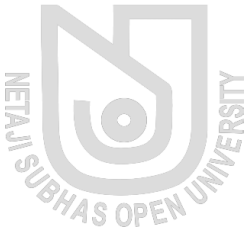
GROUP
B(I)

Part-I : Cytogenetics

GROUP
B(II)
Molecular Biology

GROUP
B (I)

Part-I : Cytogenetics


NETAJI SUBHAS OPEN UNIVERSITY

Unit 1 Biology of Chromosomes

Structure

1.1 Molecular anatomy of eukaryotic chromosomes

The proteins that bind to the DNA to form eukaryotic chromosomes are traditionally divided into two general classes: *histories and the nonhistone chromosomal proteins.* The complex of both classes of protein with the nuclear DNA of eukaryotic cells is known as chromatin. The total mass of histones in chromatin is about equal to that of the DNA. Histones are responsible for the first and most basic level of *chromosome* organization, the nucleosome, which was discovered in 1974. At interphase nuclei most of the chromatin is in the form of a fibre with a diameter of about 30 nm. If this chromatin is subjected to treatments that cause it to unfold partially, it can be seen under the electron microscope as a series of "beads on a string". The string is DNA, and each bead is a "nucleosome core particle" that consists of DNA wound around a protein core formed from histones. The beads on a string represent the first level of chromosomal DNA packing.

The structural organization of nucleosomes was determined after first isolating them from unfolded chromatin by digestion with particular enzymes (called nucleases) that break down DNA by cutting between the nucleosomes. After digestion for a short period, the exposed DNA between the nucleosome core particles, the linker DNA, is degraded. Each individual nucleosome core particle consists of a complex of eight histone proteins—two molecules each of histones H2A, H2B, H3, and H4—and double-stranded DNA that are 146 nucleotide pairs long. The histone octamer forms a protein core around which the double-stranded DNA is wound (Fig. 1.1).

Each nucleosome core particle is separated from the next by a region of linker DNA, which can vary in length from a few nucleotide pairs up to about 80. (The term nucleosome technically refers to a nucleosome core particle plus

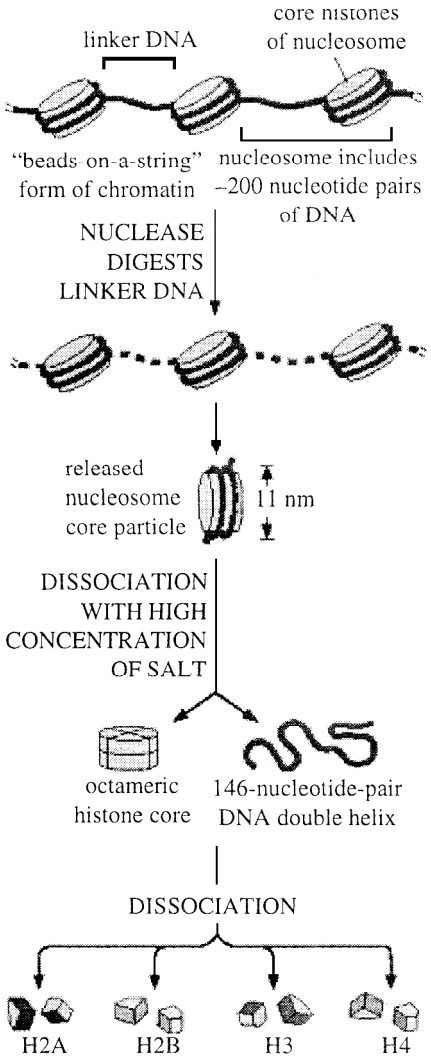


Fig. 1.1 Structural organization of the nucleosome A nucleosome contains a protein core made of eight histone molecules. As indicated, the nucleosome core particle is released from chromatin by digestion of the linker DNA with a nuclease, an enzyme that breaks down DNA. (The nuclease can degrade the exposed linker DNA but cannot attack the DNA wound tightly around the nucleosome core.) After dissociation of the isolated nucleosome into its protein core and DNA, the length of the DNA that was wound around the core can be determined. This length of 146 nucleotide pairs is sufficient to wrap 1.65 times around the histone core

one of its adjacent DNA linkers, but it is often used synonymously with nucleosome core particle.) On average, therefore, nucleosomes repeat at intervals of about 200 nucleotide pairs.

The structure of the nucleosome core particle reveals how DNA is packaged

The high-resolution structure of a nucleosome core particle, solved in 1997, revealed a disc-shaped histone core around which the DNA was tightly wrapped 1.65 turns in a left-handed coil. All four of the histones that make up the core of the nucleosome are relatively small proteins (102-135 amino acids), and they share a structural motif, known as the *histone fold,* formed from three **a** helices connected by two loops (Figure 1.2).

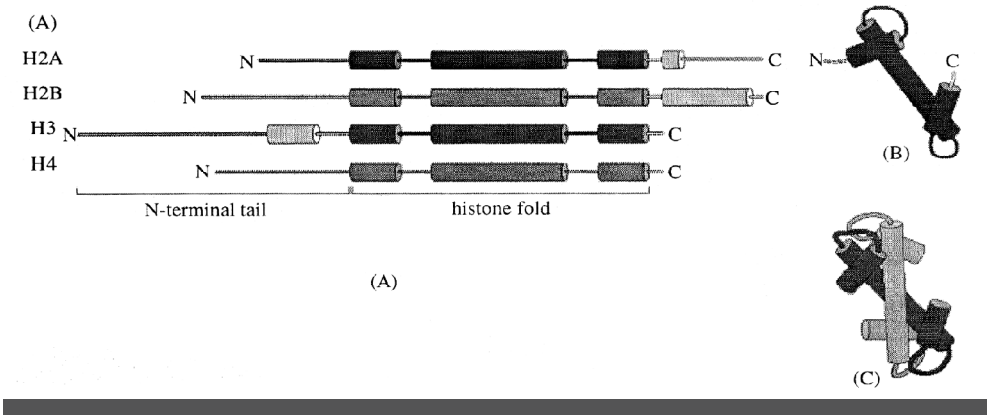


Fig. 1.2 The overall structural organization of the core histones, (A) Each of the core histones contains an N-terminal **tail,** which is subject to several forms of covalent modification, and a histone fold region, as indicated. (B) The structure of the histone fold, which is formed by all four of the core histones. (C) Histones 2A and 2B form a dimer through an interaction known as the "handshake." Histones H3 and H4 form a dimer through the same type of interaction, as illustrated in Figure 1.3

In assembling a nucleosome, the histone folds first bind to each other to form H3-H4 and H2A-H2B dimers, and the H3-H4 dimers combine to form tetramers. An H3-H4 tetramer then further combines with two H2A-H2B dimers to form the compact octamer core, around which the DNA is wound (Fig. 1.3).

The interface between DNA and histone is extensive ; 142 hydrogen bonds are formed between DNA and the histone core in each nucleosome. Nearly half of these bonds form between the amino acid backbone of the histones and the phosphodiester backbone of the DNA. Numerous hydrophobic interactions and salt linkages also hold DNA and protein together in the nucleosome. These numerous interactions explain in part why DNA of virtually any sequence can be bound on a histone octamer core. The path of the DNA around the histone core is not smooth; rather, several kinks are seen in the DNA, as expected from the nonuniform surface of the core. In addition to its histone fold, each of the core histones has a long N-terminal amino acid "tail", which extends out from the DNA-histone core (see Figure 1.3). These histone tails are subject to several different types of covalent modifications, which control many aspects of chromatin structure. The histones are among the most highly conserved eukaryotic proteins. This strong evolutionary conservation suggests that the functions of histones involve nearly all of their amino acids, so that a change in any position is deleterious to the cell. Despite the high conservation of the core histones, many

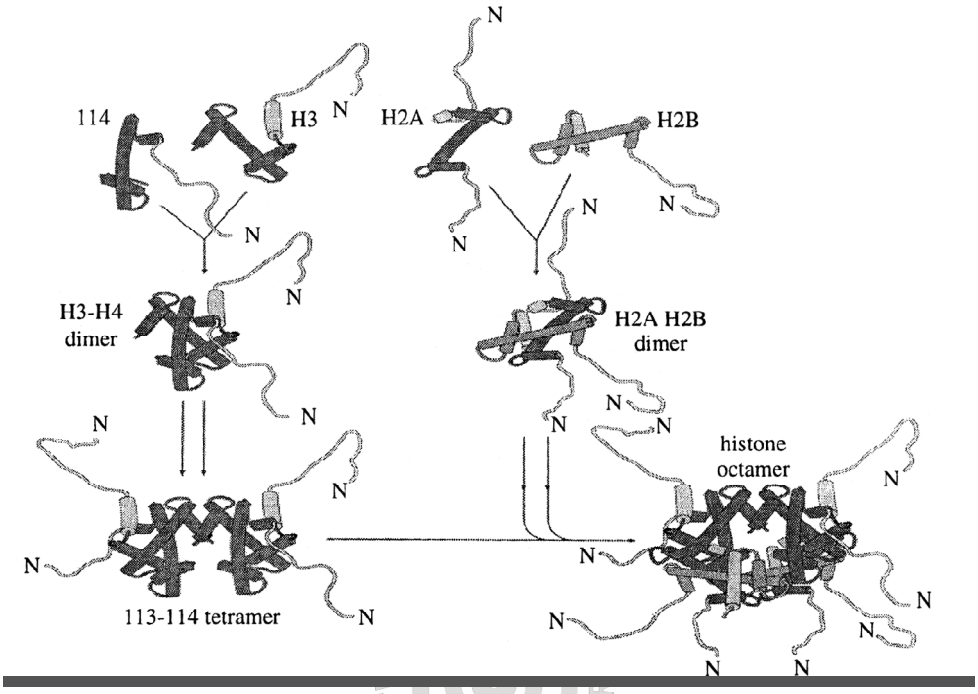


Fig. 1.3 The assembly of a histone octamer. The histone H3-H4 dimer and the H2A-H2B dimer are formed from the handshake interaction. An H3-H4 tetramer forms the scaffold of the octamer onto which two H2A-H2B dimers are added, to complete the assembly. Note that all eight N-terminal tails of the histones protrude from the disc-shaped core structure. In the x-ray crystal, most of the histone tails were unstructured (and therefore not visible in the structure), suggesting that their conformations are highly flexible. (Adapted from figures by J. Waterborg.)

eukaryotic organisms also produce specialized variant core histones that differ in amino acid sequence from the main ones. It is thought that nucleosomes that have incorporated these variant histones differ in stability from regular nucleosomes, and they may be particularly well suited for the high rates of DNA transcription and DNA replication that occur during these early stages of development.

The positioning of nucleosome on DNA is determined by both DNA flexibility and other DNA-bound proteins

Two main influences determine where nucleosomes form in the DNA. One is the difficulty of bending the DNA double helix into two tight turns around the outside of the histone octamer, a process that requires substantial compression of the minor groove of the DNA helix. Because A-T-rich sequences in the minor groove are easier to compress than G-C-rich sequences, each histone octamer tends to position itself on the DNA so as to maximize A-T-rich minor grooves on the inside of the DNA coil. Thus, a segment of DNA that contains short

A-T-rich sequences spaced by an integral number of DNA turns is easier to bend around the nucleosome than a segment of DNA lacking this feature. In addition, because the DNA in a nucleosome is kinked in several places, the ability of a given nucleotide sequence to accommodate this deformation can also influence the position of DNA on the nucleosome.

The second, and probably most important, influence on nucleosome positioning is the presence of other tightly bound proteins on the DNA. Some bound proteins favour the formation of a nucleosome adjacent to them. Others create obstacles that force the nucleosomes to assemble at positions between them. Finally, some proteins can bind tightly to DNA even when their DNA-binding site is part of a nucleosome. The exact positions of nucleosomes along a stretch of DNA therefore depend on factors that include the DNA sequence and the presence and nature of other proteins bound to the DNA.

Nucleosomes are usually packed together into a compact chromatin fibre

The nucleosomes are packed on top of one another, generating regular arrays in which the DNA is even more highly condensed. Thus, when nuclei are very gently lysed onto an electron microscope grid, most of the chromatin is seen to be in the form of a fiber with a diameter of about 30 nm, which is considerably wider than chromatin in the "beads on a string" form. Several models have been proposed to explain how nucleosomes are packed in the 30-nm chromatin fiber; the one most consistent with the available data is a series of structural variations known collectively as the zigzag model. In reality, the 30-nm structure found in chromosomes is probably a fluid mosaic of the different zigzag variations.

Several mechanisms probably act together to form the 30-nm fiber from a linear string of nucleosomes. First, an additional histone, called histone HI, is involved in this process. H1 is larger than the core histones and is considerably less well conserved. A single histone HI molecule binds to each nucleosome, contacting both DNA and protein, and changing the path of the DNA as it exits from the nucleosome. Although it is not understood in detail how H1 pulls nucleosomes together into the 30-nm fiber, a change in the exit path in DNA seems crucial for compacting nucleosomal DNA so that it interlocks to form the 30-nm fibre (Fig. 1.4).

A second mechanism for forming the 30-nm fiber probably involves the tails of the core histones, which, as we saw above, extend from the nucleosome. It is thought that these tails may help attach one nucleosome to another—thereby allowing a string of them, with the aid of histone H1, to condense into the 30-nm fibre (Fig. 1.5).

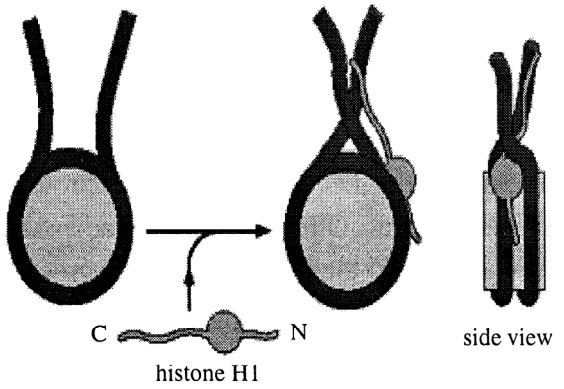


Fig. 1.4 A speculative model for how histone H1 could change the path of DNA as it exits from the nucleosome. Histone H1 (green) consists of a globular core and two extended tails. Part of the effect of H1 on the compaction of nucleosome organization may result from charge neutralization: like the core his-tones, H1 is positively charged (especially its C-terminal tail), and this helps to compact the negatively charged DNA. Unlike the core histones, H1 does not seem to be essential for cell viability; in one ciliated protozoan the nucleus expands nearly two fold in the absence of H1, but the cells otherwise appear normal

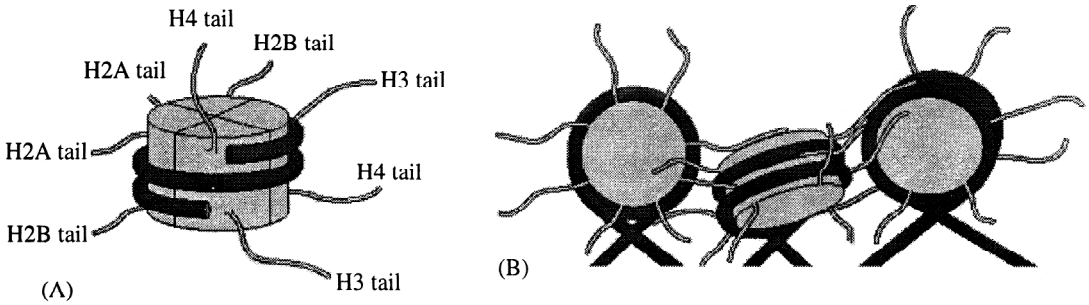


Fig. 1.5 A speculative model for histone tails in the formation of the 30-nm fiber. (A) The approximate exit points of the eight histone tails, four from each histone subunit, that extend from each nucleosome. In the high-resolution structure of the nucleosome, the tails are largely unstructured, suggesting that they are highly flexible. (B) A speculative model showing how the histone tails may help to pack nucieosomes together into the 30-nm fiber. This model is based on (1) experimental evidence that histone tails aid in the formation of the 30-nm fiber, (2) the x-ray crystal structure of the nucleosome, which showed that the tails of one nucleosome contact the histone core of an adjacent nucleosome in the crystal lattice, and (3) evidence that the histone tails interact with DNA

ATP-driven chromatin remodeling machines change nucleosome structure

Eukaryotic cells contain chromatin remodeling complexes, protein machines that use the energy of ATP hydrolysis to change the structure of nucieosomes temporarily so that DNA becomes less tightly bound to the histone core. The remodeled state may result from movement of the H2A-H2B dimers in the nucleosome core; the H3-H4 tetramer is particularly stable and would be difficult to rearrange (Fig. 1.3).

The remodeling of nucleosome structure has two important consequences. First, it permits ready access to nucleosomal DNA by other proteins in the cell, particularly those involved in gene expression, DNA replication, and repair. Even after the remodeling complex has dissociated, the nucleosome can remain

in a "remodeled state" that contains DNA and the full complement of histones—but one in which the DNA-histone contacts have been loosened; only gradually does this remodeled state revert to that of a standard nucleosome. Second, remodeling complexes can catalyze changes in the positions of nucleosomes along DNA (Fig. 1.6); some can even transfer a histone core from one DNA molecule to another.

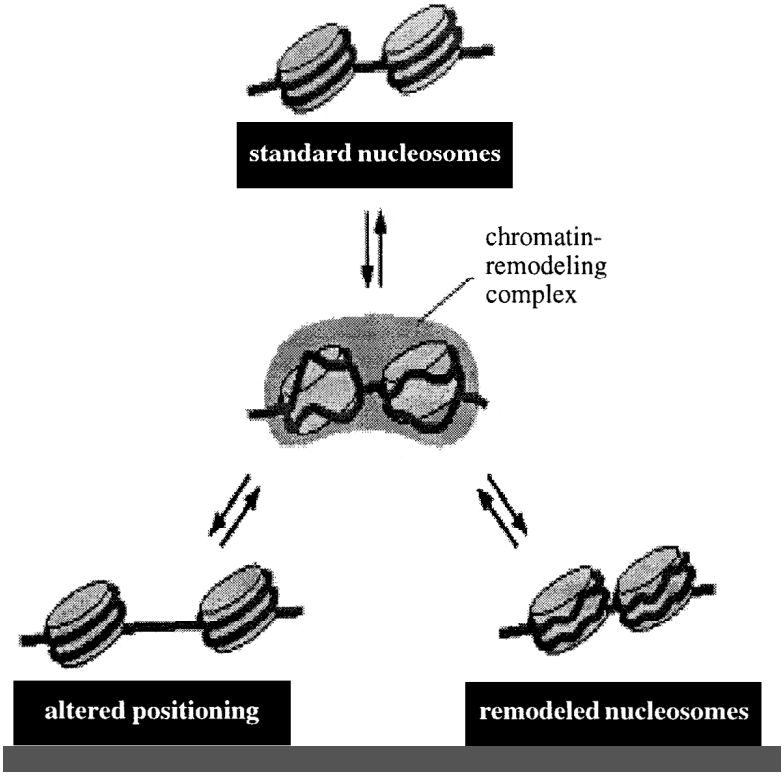


Fig. 1.6 Model for the mechanism of some chromatin remodeling complexes. In the absence of remodeling complexes, the interconversion between the three nucleosomal states shown is very slow because of a high activation energy barrier. Using ATP hydrolysis, chromatin-remodeling complexes (green) create an activated intermediate (shown in the center of the figure) in which the histone-DNA contacts have been partly disrupted. This activated state can then decay to any one of the three nucleosomal configurations shown. In this way, the remodeling complexes greatly increase the rate of interconversion between different nucleosomal states. The remodeled state, in which the histone-DNA contacts have been loosened, has a higher free energy level than that of standard nucleosomes and will slowly revert to the standard nucleosome conformation, even in the absence of a remodeling complex. Cells have many different chromatin remodeling complexes, and they differ in their detailed biochemical properties; for example, not all can change the position of a nucleosome, but all use the energy of ATP hydrolysis to alter nucleosome structure. (Adapted from R.E. Kingston and G.J. Narlikar, Genes Dev. 13:2339-2352, 1999.)

Cells have several different chromatin remodeling complexes that differ subtly in their properties. Most are large protein complexes that can contain more than ten subunits. It is likely that they are used whenever a eucaryotic cell needs direct access to nucleosome DNA for gene expression, DNA replication, or DNA repair. Different remodeling complexes may have features specialized for each of these roles. It is thought that the primary role of some remodeling complexes is to allow access to nucleosomal DNA, whereas that of others is to re-form nucleosomes when access to DNA is no longer required (Fig. 1.7).

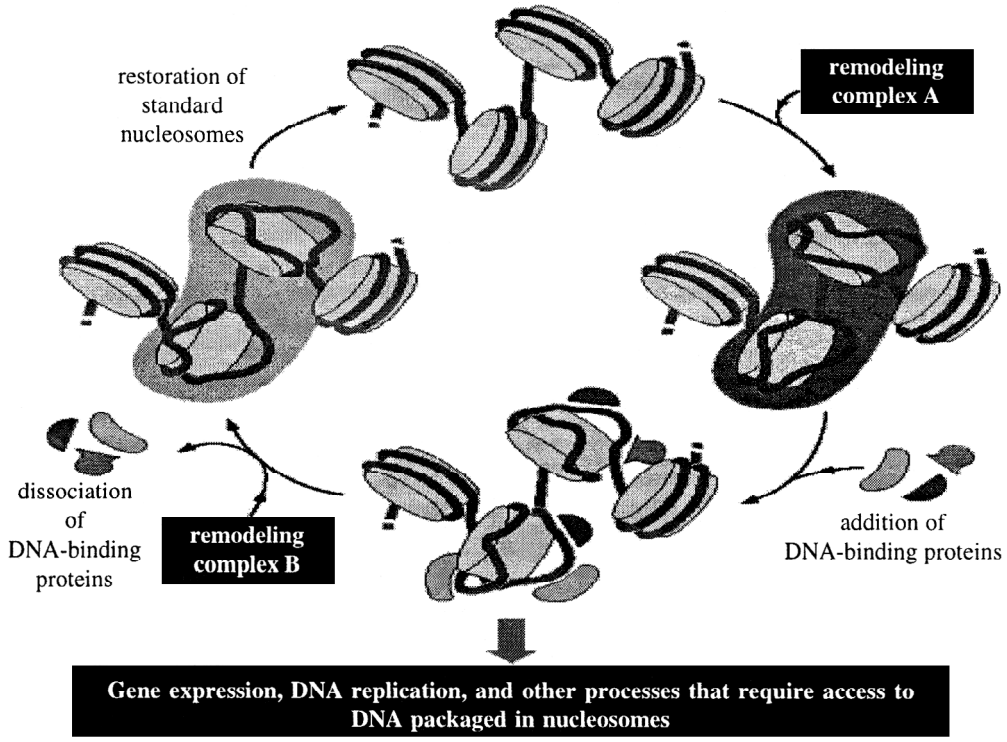
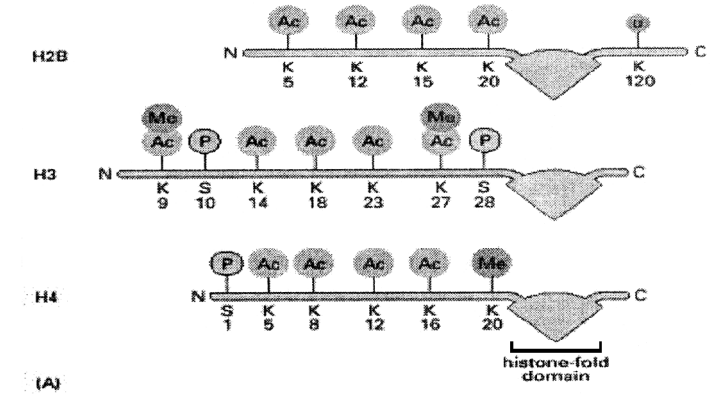


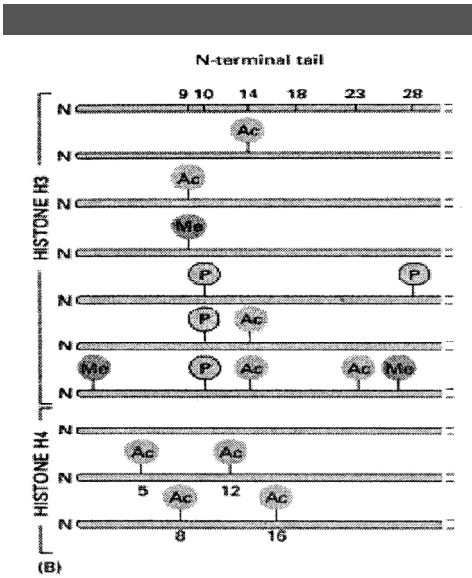
Fig. 1.7 A cyclic mechanism for nucleosome disruption and re-formation. According to this model, different chromatin remodeling complexes disrupt and re-form nucleosomes, although, in principle, the same complex might catalyze both reactions. The DNA-binding proteins could function in gene expression, DNA replication, or DNA repair, and in some cases their binding could lead to the dissociation of the histone core to form nucleosome-free regions of DNA, (Adapted from A. Travers, Cell 96:311-314, 1999.)

Covalent modification of the histone tails can profoundly affect chromatin

The N-terminal tails of each of the four core histones are highly conserved in their sequence, and perform crucial functions in regulating chromatin structure.

Each tail is subject to several types of covalent modifications, including acetyJation of lysine: methylation of lysines, and phosphorylation of serines (Fig. 1.8).





modification state	"meaning"
unmodified	gene silencing?
acetylated	gene expression
acetylated	histone deposition
methylated	gene-silencing/ heterochromatin
phosphorylated	mitosis/meiosis
phosphorylated/ acetylated	gene expression
higher-order combinations	?
unmodified	gene silencing?
acetylated	histone deposition
acetylated	gene expression

Fig. 1.8 Covalent modification of core histone tails

Histones are synthesized in the cytosol and then assembled into nucleosomes. Some of the modifications of histone tails occur just after their synthesis, but before their assembly. The modifications that concern us, however, take place once the nucleosome has been assembled. These nucleosome modifications are added and removed by enzymes that reside in the nucleus; for example, acetyl groups are added to the histone tails by histone acetyl transferases (HATs) and taken off by histone deacetylases (HDACs). The various modifications of the histone tails have several important consequences, histone acetylation tends to destabilize chromatin structure, perhaps in part because adding an

acetyl group removes the positive charge from the Iysine, thereby making it more difficult for histones to neutralize the charges on DNA as chromatin is compacted. However, the most profound effect of modified histone tails is their ability to attract specific proteins to a stretch of chromatin that has been appropriately modified.

The enzymes that modify (and remove modifications from) histone tails are usually multisubunit proteins, and they are tightly regulated. They are brought to a particular region of chromatin by other cues, particularly by sequence-specific DNA-binding proteins. It is likely that histone-modifying enzymes and chromatin remodeling complexes work in concert to condense and recondense stretches of chromatin; for example, evidence suggests that a particular modification of the histone tail attracts a particular type of remodeling complex. Moreover, some chromatin remodeling complexes contain histone modification enzymes as subunits, directly connecting the two processes.

1.2 Metaphase chromosome : Centromere, Kinetochore, Telomere and its maintenance

1.2.1 Centromere

The region of the chromosome that is responsible for its segregation at mitosis and meiosis is called the **centromere.** It is associated with two important features :

It contains the site at which the sister chromatids are held together prior to the separation of the individual chromosomes.

The term "centromere" historically; has been used in both the functional and structural sense to describe the feature of the chromosome responsible for its movement.

The centromere is essential for segregation, as shown by the behavior of chromosomes that have been broken. **Acentric fragment** does not become attached to the mitotic spindle. There can be only one centromere per chromosome. In some species the centromeres are "diffuse", which creates a different situation. Only discrete centromeres have been analyzed at the molecular level,

The regions flanking the centromere often are rich in satellite DNA sequences and contain a considerable amount of constitutive heterochromatin.

1.2.2 Kinetochore

Within the centromeric region, a darkly staining fibrous object of diameter or length ~400 nm can be seen. This object is called as Kinetochore. This

Kinetochore appears to be directly attached to the microtubules. The Kinetochore provides the MTOC on a chromosome.

Genetic engineering has produced plasmids of yeast that are replicated like chromosomal sequences. However, they are unstable at mitosis and meiosis, segregate erratically. Fragments of chromosomal DNA have been isolated by virtue of their ability to confer mitotic stability on these plasmids.

A CEN fragment is defined by its ability to confer stability upon such a plasmid. A *CEN* fragment derived from one chromosome can replace the centromere of another chromosome with no apparent consequence. This suggests that centromeres are interchangeable. *They are used simply td attach the chromosome to the spindle, and play no role in distinguishing one chromosome from another.*

The sequences required for centromeric function fall within a stretch of ~120 bp. The centromeric region is packaged into a nuclease-resistant structure, and it binds a single microtubule. The *S*. cerevisiae centromeric region has three types of sequence element may be distinguished in the *CEN* region.

CDE-I is a sequence of 9 bp that is conserved with minor variations at the left boundary of all centromeres.

CDE-II is a >90% A-T-rich sequence of 80-90 bp found in all centromeres; its function could depend on its length rather than exact sequence. Its base composition may cause some characteristic distortions of the DNA double helical structure.

CDE-III is an 11 bp sequence highly conserved at the right boundary of all centromeres. Sequences on either side of the element are less well conserved, and may also be needed for centromeric function.

A 240 kD complex of three proteins, called Cbf-III, binds to CDE-III. Mutations in the components of the genes coding for Cbf-III block chromosome movement at mitosis. A protein complex with motor activity connects the centromeric region of a chromosome to microtubules and contributes to movement on the mitotic spindle. The yeast S. *pombe* have the centromeres within regions of 40-100 kb that consist largely or entirely of repetitious DNA. The significance of the difference between the short centromeric regions in S. *cerevisiae* and the long regions in *S. pombe* is not clear. The common feature is that the DNA consists of noncoding sequences that are repetitive. Attempts to localize centromeric functions in *Drosophila* chromosomes suggest that they are dispersed in a large region, consisting of 200-600 kb. The large size of this type of centromere suggests that it is likely to contain several separate specialized functions, including sequences required for Kinetochore assembly, sister chromatid pairing, etc.

The primary modification comprising the constitutive heterochromatin of primate centromeres is a satellite DNA, which consists of tandem arrays of a 170 bp repeating unit. There is significant variation between individual repeats, although those at any centromere tend to be better related to one another than to members of the family in other locations. It is not clear whether the a satellite sequences themselves provide this function, or whether other sequences are embedded within the a satellite arrays.

1.2.3 Telomere

Essential feature in all chromosomes is the **telomere.** This "seals" the end. Telomere must be a special structure, because chromosome ends generated by breakage are "sticky" and tend to react with other chromosomes, whereas natural ends are stable. Two criteria in identifying a telomeric sequence :

It must lie at the end of a chromosome (or, at least, at the end of an authentic linear DNA molecule).

It must confer stability on a linear molecule.

Several telomeric sequences have been obtained from genomes of lower eukaryotes. In plant and man the construction of the telomere seems to follow a universal principle. Each telomere consists of a long series of short, tandemly repeated sequences. Table 1.1 lists the repeating units that have been identified at the ends of the linear DNA molecules. All can be written in the general form C_n $(A/T)_m$, where n > 1 and m is 1-4.

Within the telomeric region is a specific array of discontinuities, taking the form of single-strand breaks whose structure prevents them from being sealed by the ligase enzyme. They may be organized in a hairpin so that they are not recognized by nucleases.

Table 1.1 Telomeres have a common type of short tandem repeat. The repeating unit gives the sequence of one strand, in the 5'-3' direction form the telomere toward the centromere

Telomere	**Repeating unit**
Ciliate (*Tetrahymena*) macronucleus	CCCCAA
Ciliate (*Oxytricha*) macronucleus	CCCCAAAA
Trypanosoma minichromosome	CCCTA
Slime molds (*Dictyostelium*) rDNA	CCCTA
Yeast (*Saccharomyces*) chromosome	C_{2-3} A(CA)l-3
Plant (*Arabidopsis*) chromosome	C_3 TA_3
Human chromosome	C_3 TA_2

Addition of telomeric repeats to the end of the chromosome in every replication cycle could solve the problem of end replicating. The addition of repeats by de novo synthesis would counteract the loss of repeats resulting from failure to replicate up to the end of the chromosome. Extension and shortening would be in dynamic equilibrium.

The overall length of the telomere is under genetic control. If telomeres are continually being lengthened (and shortened), their exact sequence may be irrelevant. All that is required is for the end to be recognized as a suitable substrate for addition. This explains how the ciliate *telomere* functions in yeast. Extracts of *Tetrahymena* contain an enzyme, called telomerase, that uses the 3 '-OH of the G+T telomeric strand as a primer for synthesis of tandem TTGGGG repeats. Only dGTP and dTTP are needed for the activity. The telomerase is a large ribonucleoprotein. It contains a short RNA component, 159 bases long in *Tetrahymena,* 192 bases long in *Euplotes.* Each RNA includes a sequence of 15-22 bases that is identical to two repeats of the C-rich repeating sequence given in Table 1.1. This RNA provides the template for synthesizing the G-rich repeating sequence, to which it is complementary. Bases are added individually, in the correct sequence. The enzyme progresses discontinuously. The telomerase is a specialized example of a reverse transcriptase. The protein component provides the catalytic activity of reverse transcriptase, and is (presumably) confined to acting upon the RNA template provided by the nucleic acid component.

The structure of the telomere is organized as single-stranded extension of the G-T-rich strand, usually for 14-16 bases.

A model for the structure of the end proposes the existence of a "quartet" of G residues, formed by an association of one G from each repeating unit. The association between the G residues requires that two of them change the orientation of the base with regard to the sugar (from the usual anti to be usual syn configuration). Since each repeating unit has more than one G, more than one quartet could be formed if other G residues associate, in which case quartets might be stacked upon one another in a helical manner.

It is not known how the complementary (C-A-rich) strand of the telomere is assembled, but we may speculate that it could be synthesized by using the 3'-OH of a terminal G-T hairpin as a primer for DNA synthesis.

1.3 Heterochromatin and Euchromatin

1.3.1 Heterochromatin

Light-microscope studies in the 1930s distinguished between two types of chromatin in the interphase nuclei of many higher eukaryotic cells: a highly condensed form, called ***heterochromatin***, and all the rest, which is less condensed, called ***euchromatin*** which is composed of the types of chromosomal structures—30-nm fibers and looped domains. Heterochromatin, in contrast, includes additional proteins and probably represents more compact levels of organization that are just beginning to be understood. In a typical mammalian cell, approximately 10% of the genome is packaged into heterochromatin. Although present in many locations along chromosomes, it is concentrated in specific regions, including the centromeres and telomeres.

Heterochromatin is classified as :—

(i) Constitutive
(ii) Facultative

(i) Constitutive heterochromatin :

It consists of specific regions that are not expressed. They include satellite DNAs, and could play a structural role in the chromosome. Often these sequences are concentrated in specific regions, typically around the centromere,

(ii) Facultative heterochromatin :

It takes the form of entire chromosomes that are inactive in one cell lineage, although they can be expressed in other lineages. The example *parexcellence* is the mammalian X-chromosome, one copy of which (selected at random) is entirely inactive in a given female. (This compensates for the presence of two X chromosomes, compared with the one present in males.) The inactive X chromosome is perpetuated in a heterochromatic state, while the active X chromosome is part of the euchromatin. Here it is possible to see a correlation between transcriptional activity and structural organization when the *identical DNA sequences are involved in both states.*

Most DNA that is folded into heterochromatin does not contain genes. However, genes that do become packaged into heterochromatin are usually resistant to being expressed, because heterochromatin is unusually compact. Regions of heterochromatin are responsible for the proper functioning of telomeres and centromeres (which lack genes), and its formation may even help protect the genome from being overtaken by "parasitic" mobile elements of DNA. Moreover, a few genes require location in heterochromatin regions if they are to be expressed. In fact, the term *heterochromatin* (which was first defined cytologically) is likely to encompass several distinct types of chromatin structures

whose common feature is an especially high degree of organization. When agene that is normally expressed in euchromatin is experimentally relocated into a region of heterochromatin, it ceases to be expressed, and the gene is said to be *silenced.* These differences in gene expression are examples of position effects, in which the activity of a gene depends on its position along a chromosome. Many ***position effects*** exhibit an additional feature called position effect *variegation,* which result from patches of cells in which a silenced gene has become reactivated; once reactivated, the gene is inherited stably in this form in daughter cells. The study of position effect variegation has revealed two important characteristics of heterochromatin. First, heterochromatin is dynamic; it can "spread" into a region and later "retract" from it at low but observable frequencies. Second, the state of chromatin—whether heterochromatin or euchromatin—tends to be inherited from a cell to its progeny. These two features are responsible for position effect variegation, as explained in Figure 1.9.

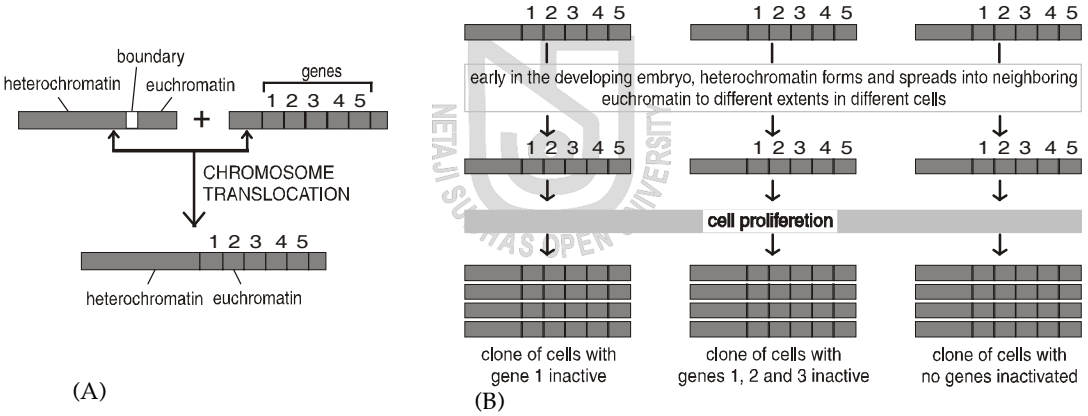


Fig. 1.9 The cause of position effect variegation in *Drosophila.* (A) Heterochromatin is normally prevented from spreading into adjacent regions of euchromatin by special boundary DNA sequences. In flies that inherit certain chromosomal rearrangements, however, this barrier is no longer present. (B) During the early development of such flies, heterochromatin can spread into neighboring chromosomal DNA, proceeding for different distances in different cells. This spreading soon stops, but the established pattern of heterochromatin is inherited, so that large clones of progeny cells are produced that have the same neighboring genes condensed into heterochromatin and thereby inactivated. Although "spreading" is used to describe the formation of new heterochromatin near previously existing heterochromatin, the term may not be wholly accurate. There is evidence that during expansion, heterochromatin can "skip over" some regions of chromatin, sparing the genes that lie within them from repressive effects. One possibility is that heterochromatin can expand across the base of some DNA loops, thus bypassing the chromatin contained in the loop

The ends of chromosomes have a special form of heterochromatin

The molecular nature of heterochromatin is probably best understood in the simple yeast *S. cerevisiae.* Mutations in any one of a set of yeast Silent information regulator (Sir) proteins prevent the silencing of genes located near

telomeres, thereby allowing these genes to be expressed. Analysis of these proteins has led to the discovery of a telomere-bound Sir protein complex that recognizes underacetylated N-terminal tails of selected histones (Fig. 1.10A). One of the proteins in this complex is a highly conserved histone deacetylase known as Sir2, which has homologs in diverse organisms, including humans, and presumably has a major role in creating a pattern of histone under acetylation unique to heterochromatin.

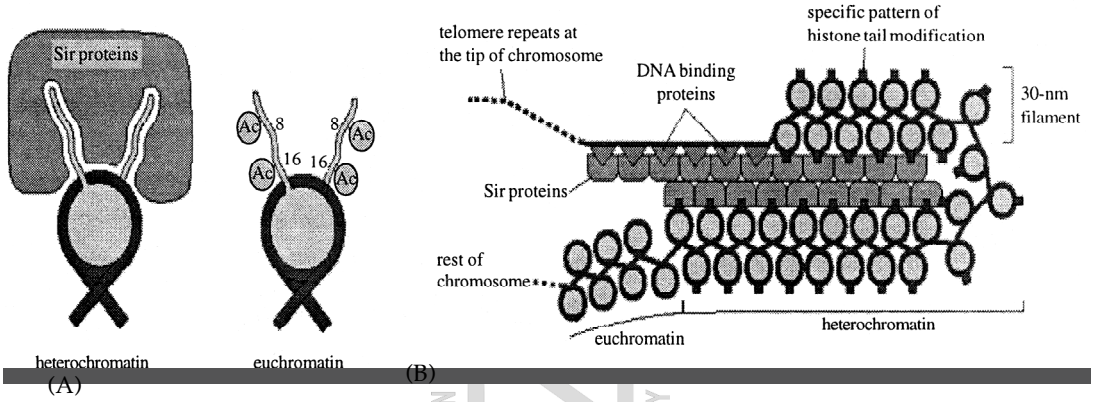


Fig. 1.10 Speculative model for the heterochromatin at the ends of yeast chromosomes. (A) Heterochromatin is generally underacetylated, and underacetylated tails of histone H4 are proposed to interact with a complex of Sir proteins, thus stabilizing the association of these proteins with nucleosomes. Although shown as fully unacetylated, the exact pattern of histone H4 tail modification required to bind to the Sir complex is not known with certainty. In some organisms, the methylation of lysine 9 of histone H3 is also a critical signal for heterochromatin formation. In euchromatin, histone tails are typically highly acctylated. Those of H4 are shown as partially acetylated but, in reality, the acetylation state varies across euchromatin. (B) Specialized DNA-binding proteins (*blue triangles*) recognize DNA sequences near the ends of chromosomes and attract the Sir proteins, one of which (Sir2) is a NAD^+ dependent histone deacetylase. This then leads to the cooperative spreading of the Sir protein complex down the chromosome. As this complex spreads, the deacetylation catalyzed by Sir2 helps create new binding sites on nucleosomes for more Sir protein complexes. A "fold back" structure of the type shown may also form

But how is the Sir2 protein delivered to the ends of chromosomes in the first place ? A DNA-binding protein that recognizes specific DNA sequences in yeast teiomeres also binds to one of the Sir proteins, causing the entire Sir protein complex to assemble on the telomeric DNA. The Sir complex then spreads along the chromosome from this site, modifying the N-terminal tails of adjacent histones to create the nucleosome-binding sites that the complex prefers. This "spreading effect" is thought to be driven by the cooperative binding of adjacent Sir protein complexes, as well as by the folding back of the chromosome on itself to promote Sir binding in nearby regions (see Fig. 1.10B). In addition, the formation of heterochromatin probably requires the action of chromatin remodeling complexes to readjust the positions of nucleosomes as they are packed together.

Whatever the precise mechanism of heterochromatin formation, it has become clear that covalent modifications of the nucleosome core histones have a critical role in this process. Of special importance in many organisms are the *histone methyl transferases,* enzymes that methylate specific lysines on histones including lysine 9 of histone H3 . This modification is "read" by heterochromatin components (including HPl in *Drosophila*) that specifically bind this modified form of histone H3 to induce the assembly of heterochromatin. It is likely that a spectrum of different histone modifications is used by the cell to distinguish heterochromatin from euchromatin.

Centromeres are also packaged into heterochromatin

In many complex organisms, including humans, each centromere seems to be embedded in a very large stretch of heterochromatin that persists throughout interphase. The stmcture and biochemical properties of this so-called *centric heterochromatin* are not well understood, but, like other forms of heterochromatin, it silences the expression of genes that are experimentally placed into it. It contains, in addition to histones (which are typically underacetylated and methylated in heterochromatin), several additional structural proteins that compact the nucleosomes into particularly dense arrangements.

Heterochromatin may provide a defense mechanism against mobile DNA elements

DNA packaged in heterochromatin often consists of large tandem arrays of short, repeated sequences that do not code for protein, as we saw above for the heterochromatin of mammalian centromeres. This suggests that some types of repeated DNA may be a signal for heterochromatin formation. This feature, called *repeat-induced gene silencing,* may be a mechanism that cells have for protecting their genomes from being overtaken by mobile genetic elements. These elements can multiply and insert themselves throughout the genome.

Once a cluster of such mobile elements has formed, the DNA that contains them would be packaged into heterochromatin to prevent their further proliferation. The same mechanism could be responsible for forming the large regions of heterochromatin that contain large numbers of tandem repeats of a simple sequence, as occurs around centromeres.

1.3.2 Euchromatin

Euchromatin is a lightly packed form of **chromatin** that is rich in gene concentration, and is often (but not always) under active **transcription.** Unlike **heterochromatin,** it is found in both **eukaryotes** and **prokaryotes.**

Structure

The structure of euchromatin is reminiscient of an unfolded set of beads along a string, where those beads represent nucleosomes. Nucleosomes consist of eight proteins known as histones, with apprxomiately 145bp of DNA wound around them; in euchromatin this wrapping is loose so that the raw DNA may be accessed. Each core histone possesses a 'tail' structure which can vary in several ways; it is thought that these variations act as "master control switches" which determine the overall arrangement of the chromatin. In particular, it is believed that the presence of-methylated lysine 4 on the histone tails acts as a general marker for euchromatin. The exact organization of the DNA within euchromatin is not known, but with the electron microscope it is possible to see loops of DNA within the euchromatin regions, each loop between 40 and 100 kb in length and predominantly in the form of the 30 nm chromatin fiber. The loops are attached to the nuclear matrix via AT-rich DNA segments called **matrix-associated regions (MARs)** or **scaffold attachment regions (SARs).**

Appearance

Euchromatin generally appears as light-colored bands when stained in **GTG banding** and observed under an **optical microscope;** in contrast to **heterochromatin,**
which stains darkly. This lighter staining is due to the less compact structure of euchromatin. It should be noted that in **prokaryotes,** euchromatin is the only form of chromatin present; this indicates that the heterochromatin structure evolved later along with the **nucleus,** possibly as a mechanism to handle increasing genome size and therefore a decrease in safety/managability.

Function

Euchromatin participates in the active transcription of DNA to mRNA products. The unfolded structure allows gene regulatory proteins and RNA polymerase complexes to bind to the DNA sequence, which can then initiate the transcription process. Not all euchromatin is necessarily transcribed, but in general that which is not is transformed into **heterochromatin** to protect the genes while they are not in use. There is therefore a direct link to how actively productive a cell is and the amount of euchromatin that can be found in its nucleus. It is thought that the cell uses transformation from euchromatin into heterochromatin as a method of controlling gene expression and replication, since such processes behave differently on densely compacted chromatin- this is known as the 'accessibility hypothesis'.

1.4 Giant chromosomes : Polytene and Lampbrush chromosomes

1.4.1 Polytene chromosomes

In dividing diploid cells the DNA synthetic phase (S phase) is regularly followed by mitosis (M phase). The alternation of Gl, S, G2, M and Gl phases is called the **cell cycle.** The process of recurrent duplication cycle without consequent mitosis is called endoreduplication. Many of the cells of certain fly larvae grow to an enormous size through multiple cycles of DNA synthesis without cell division. The resulting giant cells contain as much as several thousand times the normal DNA complement. Cells with more than the normal DNA complement are said to be **polyploid** when they contain increased numbers of standard chromosomes. In several types of secretory cells of fly larvae, however, all the homologous chromosome copies are held side by side, creating a single polytene chromosome. The fact that, in some large insect cells, polytene chromosomes can disperse to form a conventional polyploid cell demonstrates that these two chromosomal states are closely related, and that the basic structure of a polytene chromosome must be similar to that of a normal chromosome. Instances of polyploid chromosomes in Drosophila include ovary nurse cells, follicle cells surrounding oocytes, abdominal histoblasts (see **Escargot),** fat body cells, gut cells, and cells of the late prepupal **salivary gland.** During the process of polyploidization, chromosomes become multistranded.

Polytene chromosomes are large and precisely aligned side-by-side adherence of individual chromatin strands greatly elongates the chromosome axis and prevents tangling. Polyteny has been most studied in the salivary gland cells of *Drosophila* larvae, in which the DNA in each of the four *Drosophila* chromosomes has been replicated through 10 cycles without separation of the daughter chromosomes, so that **1024** (2^{10}) identical strands of chromatin are lined up side by side.

When viewed in the light microscope, distinct alternating dark **bands** and light **interbands** are visible. Each band and interband represents a set of **1024** identical DNA sequences arranged in register. About **95 %** of the DNA in polytene chromosomes is in bands, and **5%** is in interbands. The chromatin in each band appears dark, either because it is much more condensed than the chromatin in the interbands, or because it contains a higher proportion of proteins, or both. Depending on their size, individual bands are estimated to contain 3000-300,000 nucleotide pairs in a chromatin strand. The bands of *Drosophila* polytene chromosomes can be recognized by their different thickness and spacings, and each one has been given a number to generate a chromosome "map." There are

approximately 5000 bands and 5000 interbands in the complete set of *Drosophila* polytene chromosomes.

Polyploid chromosomes exhibit a banded structure that is reproducible from individual to individual. In Drosophila there are thousands of recognizable bands. In situ hybridization of cloned complementary DNA of identified genes to banded polyploid chromosomes allows the localization of genes to individual chromosome bands. Chromosomal rearrangements are easily documented by comparing the order of bands between individuals, lines or even species. The degree of rearrangement observed between species is indicative of their evolutionary distance. *Drosophila melanogaster* has four pairs of chromosomes, three pairs of autosomes and a pair of sex chromosome.

The reference system proposed by Bridges divides the limbs of salivary gland chromosomes into 102 sections called "divisions" designated by number from 1 to 102. Each of the five main limbs (X, 2L, 2R, 3L, and 3R) contains 20 divisions; the short chromosome 4 contains only two divisions. The divisions are started with a prominent band and divided further into 6 subdivisions, each designated with capital letters from A to F. Each subdivision starts with a sharp band. Thus each individual band of salivary gland chromosomes can be identified by giving the division number, subdivision, and the number of the band starting from the beginning of the subdivision. Bridges presents the following minimum numbers of bands for the salivary gland chromosomes of *Drosophila melanogaster*: 537 bands for the X chromosome, 1032 bands for the second chromosome, 1047 bands for the third chromosome, and 34 bands for the fourth chromosome, totalling a minimum of 2650 bands for the whole genome. In this initial count doublets were listed as single bands; more recent interpretations give the total number of bands as 3286 (Sorsa, 1988).

In late prepupal salivary gland chromosomes, not all DNA in each of the chromosomes is polyploid. Approximately a third of the Drosophila genome is represented by heterochromatin, and heterochromatic regions are underrepresented in polytene chromosomes as these regions do not undergo endoreduplication. For example, the *rolled* locus is found in a heterochromatic region of chromosome 2 that is considered to remain condensed (and for the most part transcnptionally inactive) throughout all or most of the cell cycle, *rolled* lies in what is considered to be alpha heterochromatin, a chromosome region that makes up the chromocenter of polytene salivary gland chromosomes. The chromocenter is thought to be made up of DNA and protein in a dense, tightly knit structure that is transcnptionally inactive. Such heterochomatic regions, which make up 30% of the *Drosophila* genome, have a much lower density of genes as compared to euchromatin. *Rolled* gene activity is unusual in that it requires the surrounding heterochromatin for gene function. Rolled gene

activity is severly impaired by bringing *rolled* close to any euchromatic position. However, these position effects can be reversed by chromosomal rearrangements that bring the *rolled* gene closer to any block of autosomal or X chromosome heterochromatin (Eberl, 1993).

Both bands and interbands in Polytene chromosomes contain genes

Since the number of bands in *Drosophila* chromosomes was once thought to be roughly equal to the number of genes in the genome, it was initially thought that each band might correspond to a single gene; however, we now know this simple idea is incorrect. There are nearly three times more genes in *Drosophila* than chromosome bands, and genes are found in both band and interband regions. Moreover, some bands contain multiple genes, and some bands seem to lack genes altogether.

It seems likely that the band-interband pattern reflects different levels of gene expression and chromatin structure along the chromosome, with genes in the less compact interbands being expressed more highly than those in the more compact bands. The remarkable appearance of fly polytene chromosomes is thought to reflect the heterogeneous nature of the chromatin compaction found along all interphase chromosomes. The remarkable appearance of fly polytene chromosomes is thought to reflect the heterogeneous nature of the chromatin compaction found along all interphase chromosomes.

Individual Polytene chromosome bands can unfold and refold as a unit

A major factor controlling gene expression in the polytene chromosomes of *Drosophila* is the insect steroid hormone ***ecdysone***, the levels of which rises and falls periodically during larval development. When ecdysone concentrations rise, they induce the expression of genes coding for proteins that the larva requires for each molt and for pupation. As the organism progresses from one developmental stage to another, distinctive ***chromosome puffs*** arise and old puffs recede as new genes become expressed and old ones are. turned off. Most puffs arise from the decondensation of a single chromosome band.

Puffing is the term that describes structural changes in polytene chromosomes. If one observes polytene chromosomes during the late prepupal stage, different bands appear to be puffed up. Puffs, then, afford **a** view of the temporal sequence of gene activation. A temporal pattern to puffing in the salivary glands of late prepupal flies is inducible by ecdysone injection and is therefore under control of the **ecdysone receptor.** A small number of genes react by puffing within minutes of exposure to ecdysone, and a much larger number (>100) react within hours. It is hypothesized that the time sequence of puffing represents a genetic hierarchy of gene activation. Early puffs are independent of protein

synthesis while late puffs require prior protein synthesis (Ashburner, 1990).

In recent years, transcription factors and chromosomal proteins have been localized to various bands. Binding of these proteins is thought to have functional significance and to reflect the activity of these proteins in gene regulation. For more information on the binding of various proteins and RNA species to bands, see **HP1/ Su(var)205, polycomb, male sex lethal 2,** and **suppressor of hairy wing.**

An example of binding of specific proteins to polytene chromosomes is found in a study of the protein CHDl (chromo-ATPase/helicase-DNA-binding domain). Proteins related to CHDl via the helicase domain have been shown to exist in large multiprotein complexes. For example SNF2/SWI2/Brm proteins are thought to participate in ATP-dependent remodeling of chromatin. Antibodies to CHDl localize this protein to extended chromatin (interbands) and regions associated with high transcriptional activity (puffs) on polytene chromosomes from salivary glands. These observations support the idea that CHDl functions to alter chromatin structure in a way that facilitates gene expression (Stokes, 1996).

Polyploidization by endoreduplication requires regulation of the **cell cycle.** What makes one region of the chromosome become polyploid while another remains underreplicated. Information about the roles of cell cycle genes in the regulation of polyploidization can be found in **cyclin E, Escargot,** and **origin recognition complex 2.**

Electron micrographs of certain puffs, called Balbiani rings, of *Chiwnomus* salivary gland polytene chromosomes show the chromatin arranged in loops, much like those observed in the amphibian lampbrush chromosomes. Each loop contains a single gene. When not expressed, the loop of DNA assumes a thickened structure, possibly a folded 30-nm fiber, but when gene expression is occurring, the loop becomes more extended. Both types of loops contain the four core histones and histone HI.

It seems likely that the default loop structure is a folded 30-nm fiber and that the histone modifying enzymes, chromatin remodeling complexes, and other proteins required for gene expression all help to convert it to a more extended form whenever a gene is expressed.

1.4.2 Lampbrush chromosomes

In 1882 Fleming first observed these chromosomes in urodele amphibian ovary. Riickert (1892) first described in great detail in shark oocytes. He coined the name “Lampbrush Chromosome” because of their brush-like appearance. These chromosomes’ occur at diplotene stage of meiotic prophase in oocytes of all animal species, in spermatocytes of several species and even in giant nucleus

of unicellular algae *Acetabularia*. These chromosomes are characterized by several lateral projections called "Lateral Loops". They are very large and best seen in salamander oocytes because of their high DNA content.

Morphology : Lampbrush chromosomes are extensible and elastic. These chromosomes with well-developed lateral loops can be stretched to about 2½ times of original length. Since, these chromosomes are found in meiotic prophase they are present in the form of bivalents in which maternal and paternal chromosomes are held together by chiasmata. The axis of each chromosome consists of a row of granules or chromomeres and from which lateral loops extend. :

(i) Centromeres : These are round, smooth, and Fuelgen positive and bear no lateral loop. In many species of urodele, centromeres are identifiable chromosome landmarks as "axial bars", formed by the amalgamation of neighbouring chromomeres, whereas in certain species of urodeles centromeres are not flanked by axial bars and are difficult to be identified, flank them. In such urodele species partner centromere do not fuse whereas urodele having axial bars centric fusion occurs.

(ii) Telomeres : Ends of Lampbrush chromosomes are occupied by distinctive telomeres consisting of a small fuelgen positive part closely applied to the surface of a smooth round Fuelgen negative part. Fuelgen negative material can be digested by proteolytic enzyme. Like centromere, Telomere do not possess lateral loop. They are of different sizes, large in T. c. *cristatus* and small in T. c. *karelini.* In some urodele fusion b/w telomeres are common in T. c. *aristatus* while in T. c. *karetini* fusions are rare.

(iii) Lateral loops : Loops are always symmetrical. Each chromosome having two of them, one for each chromatid. Loops can be distinguished by size, thickness and by several other morphological characteristics. Each loop appears at a constant position in the chromosome and there are about 10,000 loops /chromosome set. Each loop has axis formed by a single DNA molecule. About 5-10% DNA is present in the lateral loops the rest is condensed in chromomeres of chromosome axis, which is transcriptionally inactive.

Types of loops :

1. **Normal loop :** Most loops can form one pattern termed as normal loop while other loops are distinguished by their matrix deposition.
2. **Granulous loop :** They are so called because they accumulate granule at the distal end of the fine fibers projection from other classes of lateral loop differ from granulous loop in material accumulation.

(a) In some cases matrix plastering the axis but leaving the tips of the projecting fibers visible.

(b) In other cases, matrix fusion is irregular over the loop length. Such loops have uneven outlines.

3. Lumpy Loop : Situated on either side of centromere much degree of matrix fusion, usually so great that the loop pattern is wholly observed. Sometimes sister lumpy loop may fuse together so that instead of a pair of loops a single amorphous body is present.

4. Giant Loops : They are much larger than lumpy are matrix that they accumulate is exceedingly heterogeneous in texture. Each loop has its own developmental sequence of extension and regression. For example, giant granular loops are already of full size in very young oocytes and remain the same throughout oocyte development. The giant fusing loops are small in small oocytes and regress only just before ovulation. The granulous loops are largest in young oocytes is not regress early.

Unineme theory and C. value paradox :

The results of most of the earlier studied have revealed that each loop has just 1 DNA molecule as a major finding because it showed a single thread of DNA runs through each chromatid and lead to the elaboration of unineme model concept of chromosome structure.

A matrix covers each loop that consists of RNA transcripts with RNA binding proteins attached to them. In general ribonucleoprotein matrix is asymmetrical being thickness at one end of the loop than at the others. RNA synthesis starts at the thinner end and progresses to the thicker end. Many of the loops correspond to a single transcriptional unit while the other loop contains several units of transcription. Some t unite on lampbrush chromosome are extremely long i.e. over lOOum in length i.e. (lu-m of DNA = 3000 bases) why they are so enormous? Even more puzzling is that the length of loop increases with C. value as a result Salamander has 10 times longer transcription units than those of a frog although both code for a similar gene product. We have no answer to this paradox but it may be possibly connected with inefficient termination of transcription in oocytes.

The results of *in situ* hybridization studies have revealed that long Lampbrush transcripts are due to failure of termination i.e. transcription eventually stops where the next t. unit is reached. However the function of these long transcripts remain unexplainable, though the majority of them are degraded in the nucleus but presumably some RNA has some role to play in preparing an oocyte for the journey that an egg undertakes after fertilization i.e. development of a new organism.

Anyway, at the cellular level of analysis both Lampbrush chromosome and

p.c. provide remarkably favorable opportunities to study the mechanisms responsible for gene ordered synthesis. Most of the recent exciting advances in our understanding of the nature and mode of action of the genetic material have come from the genetic studies on microorganisms. Cytologists have contributed rather little to this advances.

I.5 Suggested questions

1. Describe the structure of nucleosome along with diagrams.
2. State and explain the ways of chrornatin remodeling.
3. Sate the different types of covalent modifications in histone tails and its significance.
4. Explain the significance of centormeric sequence in chromosomal segregation.
5. How is the length of telomere maintained in eukaryotic systems?
6. Explain "position effect variegation" with example.
7. How is heterochromatinization brought about?
8. Elucidate polytene chromosome structural organization.
9. Explain chromosomal puffs.
10. State the morphology of lampbrush chromosomes.
11. Validate unineme theory by lampbrush chromosome structure.

Unit 2 Sex Chromosomes, Sex Determination and Dosage Compensation

Structure

2.1 Introduction

In multicellular organisms sex is determined by many different mechanisms, which vary greatly. Of the various mechanisms of sex determination known till date, sex-chromosomal method of determination is perhaps the best understood and intriguing. Here, sex of an individual is determined by the presence or absence of its species-specific sex chromosomes. In this system of sex determination there are defined set of autosomes and well-defined pair of *allosomes* (sex chromosomes). The allosomes may be of one kind (e.g. in C. *elegans,* Grasshopper etc. has only X chromosome; thus two sexes are determined by either XX or XO) or of two different kinds (e.g. *Drosophila* has both X and Y chromosomes and sexes are determined by XX or XY).

The paradox for such mechanism of sex determination is in the fact that either of the two sexes have different sex chromosomal constitution, leading to differential allosomal gene dosages. In many organisms there are two X chromosomes in female and one X in male. Therefore, it is essential to make a balance between the products of the genes of two X chromosomes and the products of one X chromosome. The mechanism by which the balance between two dosages and one dose is maintained is known as dosage compensation. This is done either by suppressing the activities of the genes of one of the two X chromosomes of the female (inactivation of one of the female X chromosome) or by hyperactivation of the male X chromosome. This would thus require dosage compensation to negate the genie imbalance for the sex chromosomes. Although sex determination pathway and dosage compensation are different pathways they may have few steps in common but must not be considered to be same under any circumstances.

2.2 Sex determination and dosage compensation in *Caenorhabditis elegans*

Caenorhabditis elegans has two sexes: hermaphrodites and males. Hermaphrodites are essentially female animals that produce sperm during larval development and oocytes during adulthood. Hence, hermaphrodites are capable of self-fertilization, as well as cross-fertilization by males. Although some adult structures such as the pharynx are similar in males and hermaphrodites, most tissues and many aspects of behavior are different.

The pathway is not as linear and that several loops and branches in the pathway play important roles in specifying sexual development.

2.2.1 Control of xol-1 by the X:A Ratio

The primary signal for sex determination is the ratio of X chromosomes to sets of autosomes, which causes XX animals to become hermaphrodites and XO animals to become males. Early in development this ratio regulates the activity of *xol-l* (Fig.2.1), a key developmental switch gene that controls both sex determination and dosage compensation, *xol-1* encodes a novel protein, and during early embryogenesis, high levels of XOL-1 protein activity promote male development and low levels promote hermaphrodite development. The male specifying *xol-l* transcript is not needed after the end of gastrulation.

The early time at which *xol-1* acts strongly suggests that it is a direct target of the X : A signal. This signal must involve elements on the X and elements autosomes that are compared.

The X chromosome signal is polygenic, and that the combined action of these X signal element is required to inhibit *xol-1* activity in hermaphrodites. At least four different regions, regions 1-4, of the X contain signal elements, and two of these elements have been identified molecularly: *sex-land fox-1.* Increasing

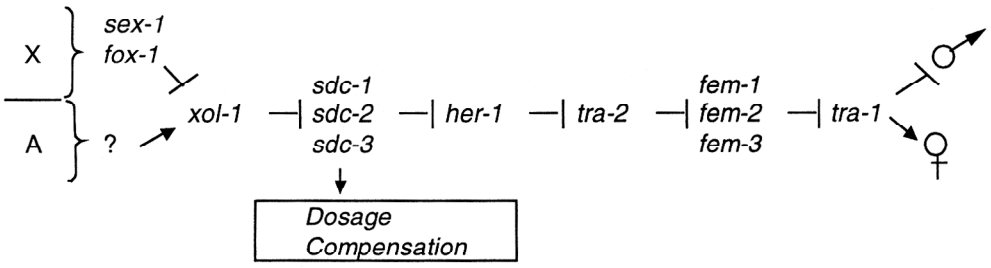


Fig. 2.1 The basic sex determination and dosage compensation pathways in *C. elegans*

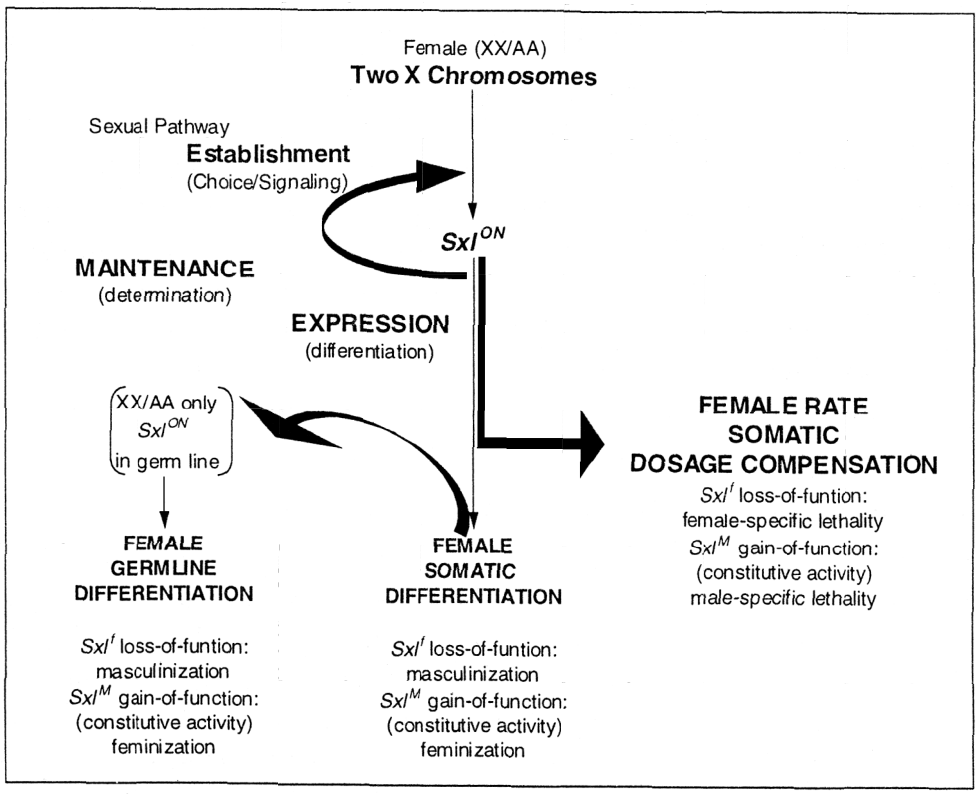


Fig. 2.2 The female-specific developmental switch gene, *Sex-lethal*, counts X chromosomes early in development to establish the choice between male and female alternative pathways of development at the cellular level. Through an autoregulatory feedback loop, *Sxl* subsequently maintains this choice throughout development, and ultimate, directs sexually dimorphic aspects of differentiation through its effects on different sets of subordinate genes downstream. Because one of these sets controls the vital process of X chromosome dosage compensation, misregulation of *Sxl* caused by upsets in the X chromosome counting process is lethal to one sex or the other. This lethality obscures potential effects on sexual phenotype. Loss-of-function {f-type) mutations in this gene are deleterious to chromosomal females (XX), while gain-of-function mutations (M-type) lead to constitutive expression and arc deleterious to chromosomal males (XY). Control ofS.t/in the germ line requires sex-specific input from the soma

the dose of these elements in XO animals represses *xol-1,* promotes hermaphrodite development and causes death because dosage compensation is activated. Decreasing their dose in XX animals activates *xol-1,* promotes male development, and causes death due to failure to initiate dosage compensation. To date, no autosomai signal elements have been identified. Evidence indicates that *sex-1* regulates the transcription *of xol-1.*

In contrast to *sex-1, fox-1* and region 2 act postranscriptionally to regulate *xol-1* expression. The *fox-1* gene encodes a protein with ribonuclear protein (RNP)

motifs, suggesting that it might bind the *xol-1* RNA. It is possible that the OF-1 protein regulates *xol-1* alternative splicing, or it might govern another aspect of *xol-1* mRNA metabolism.

Combinatorial effect of these regulatory mechanisms allows the worm to discriminate accurately between small differences in the X : A.

2.2.2 Control of the *sdc* genes by *xol-1*

Three genes are required in XX animals to promote both hermaphroditic development and dosage compensation—*sdc-], sdc-2* and *sdc-3.* The primary means by which XOL-1 transmits the X:A signal appears to be by negative regulation of *sdc-2,* as *sdc-2* is not expressed in wild-type XO embryos, but is expressed in *xol-1* XO embryos.

Null mutations in *sdc-2* and *sdc-3* have no effect on XO animals but cause complete reversal of sexual fate n XX animals; null mutations in *sdc-1* cause only a partial reversal of sexual fate. The *sdc* genes control XX hermaphrodite development by regulating the expression of the downstream sex-determining gene, *her-1*, a gene required for male development.

SDC-2 and SDC-3 might act in a complex to directly repress *her-1* transcription. This is supported by the finding that SDC-2, is highly charged and contains coiled-coil motifs, is targeted to transgenic copies of the *her-1* promoter. Moreover, this localization is blocked by specific *sdc-3* mutations, called *sdc-3 (Tra* alleles. SDC-3 is a novel protein that contains two functional domains. The first is a zinc finger motif that is required for dosage compensation but not for sex determination. The second resembles a **myosin** ATPase domain, and is necessary for sex determination but not for dosage compensation. The *sdc-3 (Tra)* mutations after this letter domain. The role of SDC-1 in regulating sexual development is less clear. SDC-1 has seven zinc fingers and resembles TFIIIA.

2.2.1 Somatic sex determination

Regulation of TRA-2A by HER-1

her-1 is required for male development, as mutations in *her-1* cause XO animals to develop as hermaphrodites but do not affect dosage compensation. The *her-1* gene is predicted to encode a novel protein with an ami no-terminal signal sequence and protein cleavage and glycosylation sites, suggesting that HER-1 is a secreted protein.

HER-1 promotes male development by repressing the activity of *tra-2.* Major transcript of this gene, *tra-2,* encodes a transmembrane protein, a direct interaction between secreted HER-1 and TRA-2A.

Regulation of sexual fate by *tra-2*

TRA-2A might be processed to release its intracellular domain, TRA-2ic. Production of TRA-2ic might occur by the action of TRA-3, a member of the calpain protease family. The *tra-3* gene is necessary for hermaphroditic development. TRA-3 can proteolytically cleave TRA-2A to release TRA-2ic in insect cells. How HER-1 inhibits TRA-2A activity is unclear. One possibility is that HER-1 inhibits production of TRA-2ic by TRA-3. *tra-2* activity is controlled not only at the level of protein processing or protein interaction, but also at the translational level by two elements called *tra Gli* elements (TGEs) which are located in the 3' untranslated region (3'UTR) of the *tra-2* message. Epistasis test show that *tra-2* promotes hermaphrodite development by inhibiting the activity of three genes, *fem-1, fem-2* and *fem-3.* TRA-2A does not transcription ally regulate these genes, as they are all expressed at high levels in both sexes. Instead it appears that TRA-2A or TRA-2ic inhibits FEM activities by protein-protein interaction. FEM-2 can bind FEM-3, so they might interact to promote male development. In addition, TRA-2A and TRA-2ic can bind FEM-3, which they might inactivate to allow hermaphrodite development.

Regulation of TRA-1A activity by the FEM proteins

The final gene in the sex-determination pathway is *tra-1,* which acts cell autonomously to promote hermaphrodite development. Although genetic experiments indicate that the FEM proteins promote male development by inhibiting TRA-1 A activity, how they do so is a mystery. They are unlikely to act transcriptionally, as *ra-1* mRNA levels do not differ between males and hermaphrodites. The phosphatase activity of Fem-2 is necessary for its activity, so it is possible that FEM-2 control the activity of TRA-1 A by altering its phosphorylation state. Alternatively the FEM proteins might control sexual

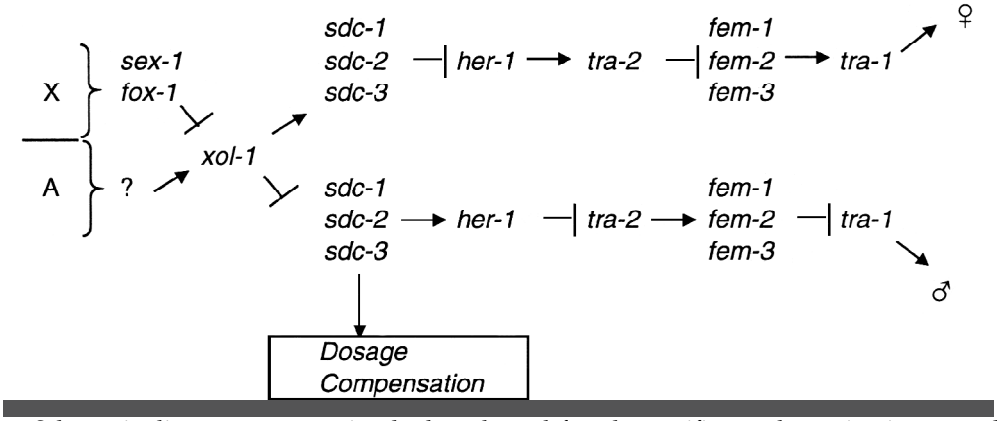


Fig. 2.3 Schematic diagram representing both male and female specific sex determination cascade in C. *elegans*

development by regulating the nuclear levels of TRA-1A. Recent analyses have revealed that hermaphrodite tissues have higher TRA-1A nuclear levels that male tissue. Thus TRA-1A transcriptional regulatory activity might be specified by nuclear versus cytoplasmic distribution of the protein. Furthermore, the FEM proteins might regulate nuclear import or export of TRA-1A, as TRA-1A is almost completely nuclear in loss-of-function *fem-1* animals. The entire sex-determination cascade can be summarized as Figure 2.2.

2.3 Sex determination and dosage compensation in *Drosophila*

In 1921 Bridges proposed that '..Sex in D. *melanogaster* is determined by a balance between the genes contained in the X chromosomes and those contained in the autosomes. It is not the simple possession of two X chromosomes that makes a female, and of one that makes a male'.

The discovery of numerator and denominator elements appears to validate the concept of sex determination of; as a ratio-measuring process. Proposed regulation cascade for *Drosophila* sex determination is schematically given below (Fig. 2.4)

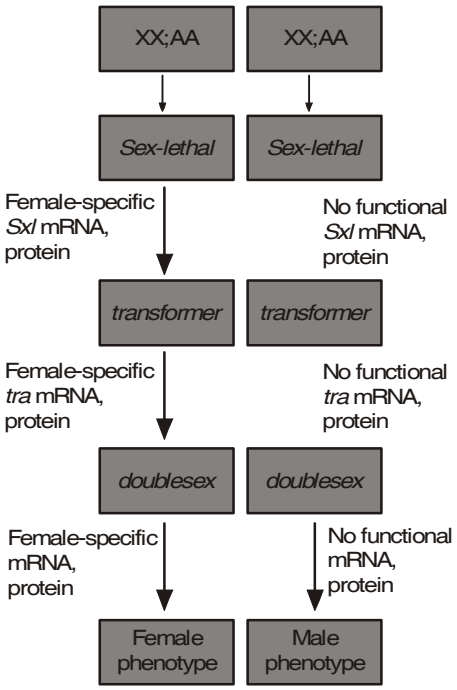


Fig. 2.4 Sex determination in *Drosophila.* This simplified scheme shows that the X-to-autosome ratio is monitored by the Sex-*lethal* gene. If this gene is active, it processes the *transformer* nRNA into a functional female-specific message. In the presence of the female-specific Transformer protein, the *doublesex* gene transcript is processed in a fern ale-specific fashion. The female-specific Doublesex protein is a transcription factor that leads to the production of the female phenotype. If the *transformer* gene does not make a female-specific product (i.e., if the *Sex-lethal* gene is not activated), the *doublesex* transcript is spliced in the male-specific manner, leading to the formation of a male-specific Doublesex protein. This is a transcription factor that generates the male phenotype

Sxl, the target of the somatic sex determination signal

Sexual differentiation and dosage compensation are a consequence of *Sxl* being turned on in diplo-X individuals and remaining off in haplo-X individuals. In somatic cells, the active (female) state is maintained by a positive feedback loop (Fig.2.2), rather than continued input from the X-linked genes that trigger the initial activation.

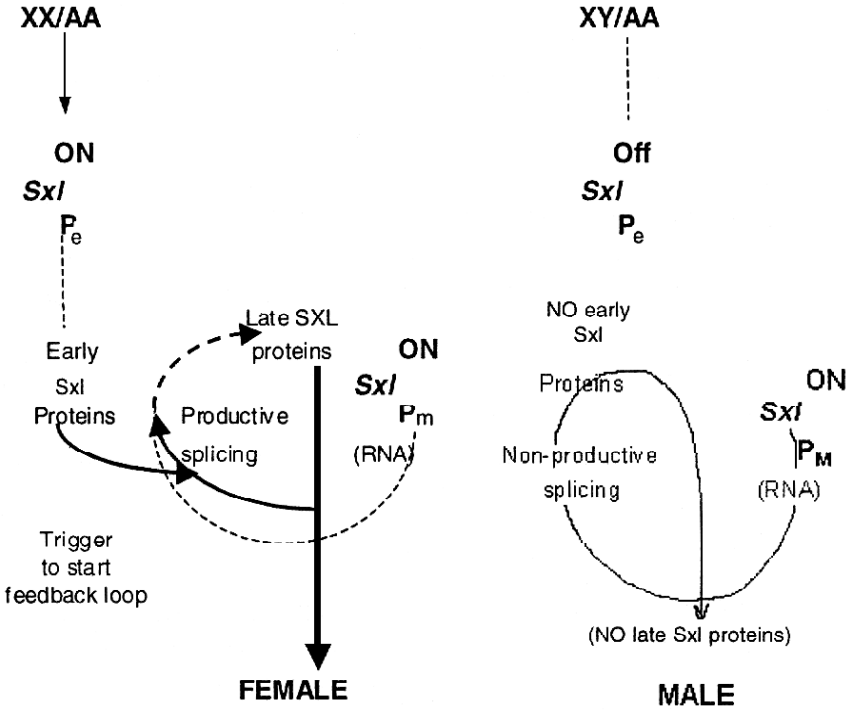


Fig. 2.5 Molecular steps in the operation of Sxl to establish and maintain a sexual pathway choice in somatic cells that is appropriate for their X chromosome dose

Figure 2.5 summarizes the current model for how, in the soma, a brief, very early effect of X chromosome dosage has a long-lasting effect on the activity of *Sxl*. The double dose of X chromosomes in females activates an *Sxl* 'establishment' promoter, $P_{e'}$ before the blastoderm stage, resulting in production of the 'early' *Sxl* proteins. In contrast, the single X chromosome dose in males leaves P_e inactive; hence males fail to produce early *Sxl* proteins.

Transition from the sexual pathway establishment (signaling) level of *Sxl* regulation to the pathway maintenance (determination) level reflects a switch in *Sxl* promoters and an attendant shift from transcriptional regulation to regulation at the level of RNA splicing. At the blastoderm stage, P_e shuts down and a 'maintenance' promoter, P_m, located 5 kb upstream, becomes active in both sexes and remains active throughout the rest of development. In contrast to

transcripts from P_e, transcripts from Pm can be spliced into mRNA encoding full-length Sxl proteins only if Sxl proteins (early or late) are already present.

The early burst of P_e expression in diplo-X individuals generates a pulse of Sxl early protein that directs the splicing of P_m-derived transcripts to eliminate a male-specific, translation-terminating exon that would otherwise block synthesis of active Sxl proteins. The Sxl proteins generated from Pm-derived transcripts then maintain productive RNA splicing through a positive auto regulatory feed back loop, and act on downstream genes to elicit female differentiation and suppress X chromosomes hyper activation.

In the absence of Sxl protein, downstream gene targets involved in both sexual differentiation and dosage compensation are expressed in a male-specific manner.

Numerator and denominator elements

The first candidate was *sisterless-a.* A zygotically acting positive regulator of Sxl. The gene was subsequently defined as 'numerator in experiments at that also revealed a second numerator elements, sis-b. the behavior of these two genes is characterized by reciprocal, zygotic dose-dependent *Sxl*-based, sex specific lethality.

One additional numerator element, the apir-rule segmentation gene runt, the first known denominator element, the semi lethal behavioral mutant *deadpan*. As a denominator element, *dpn* shows reciprocal, zygotic dose-dependent, *Sxl*-based, sex specific lethality that is the inverse of that for numerator elements; duplications, rather than deletions of this autosomal gene kill females.

Maternal involvement in progeny sex determination

Mother plays an important role in, sex determination by building into the egg the biochemical machinery the embryo needs to count its X chromosomes and decide whether to activates *Sxl*. Maternal daughterless da is a positive regulator that is necessary but not sufficient for Sxl activation. Without da activity cannot activate *Sxl* after fertilization, regardless of their X chromosome dosage. All progeny develop as males.

Analysis of the cDNA from *Sxl* mRNA shows that the *Sxl* mRNA of males differs from the *Sxl* mRNA of females. This is the result of differential RNA processing. Moreover, the Sxl protein appears to bind to its own mRNA precursor to splice it in the female manner. Since males do not have any available Sxl protein, their new 5x7 transcripts are processed in the male manner. The male *Sxl* mRNA is non functional. While the female-specific *Sxl* transcript contains a translation termination codon (UGA) after amino acid 48. The differential RNA processing that puts this termination codon into the male-specific mRNA. In

males, the nuclear transcript is spliced in a manner that yields three exons, and the termination codon is within the central exon. In females, RNA processing yields only two exons, and the male-specific central exon is now spliced out as a large intron. Thus, the female-specific mRNA lacks the termination codon. The protein made by the female-specific *Sxl* transcript can be predicted from its nucleotide sequence. This protein would contain two regions that are important for binding to RNA. Bell and colleagues (1988) have proposed that there are tow targets for the RNA-binding protein encoded by *Sxl.* One of these targets is the pre-mRNA of *Sxl* itself. This would be the mechanism that would maintain the female state of the pathway after the initial activating event had passed. The second target of the female-specific Sxl protein would be the pre-mRNA of the next gene on the pathway, *transformer.*

The transformer genes

The *Sxl* gene regulates somatic sex determination by controlling the processing of the *transformer* gene transcript. The *transformer* gene (*tra*) is alternatively spliced in males and females. There is a female-specific mRNA and also a nonspecific mRNA that is found in both females and males. The non-specific tra mRNA contains a termination codon early in the message, making the protein non-functional. The second exon of the non-specific mRNA has the termination codon. This exon is not utilized in the female-specific message. The female-specific protein form the *Sxl* gene activates a female-specific 3' splice site in the transformer pre-mRNA, causing it to be processed in a way that splices out the second exon. To do this, the *Sxl* protein blocks the binding of splicing factor U2AF to the nonspecific splice site by specifically binding to the polypyrimidine tract adjacent to it. This causes U2AF to bind to the lower-affinity (female-specific) 3' splice site and generate a female-specific mRNA. The protein encoded by this message is critical in female sex determination.

The female-specific tra product acts in concert with the *transformer-2 (tra2)* gene to help genrate the female phenotype. The *tra2* gene is eonstitutively active and makes the same protein product in both males and females. This Tra2 protein, like that of the female-specific *Sxl* protein, contains and RNA-binding domain. It is proposed tha the *tra2* gene can bind to the transcript of the *doublesex* gene, but only in the presence of the female-specific Tra protein.

Doublesex : the switch gene of sex determination

The *doubelsex* gene is active in both males and females, but its primary transcript is processed in a sex-specific manner. Male and female transcripts are identical through the first three exons. The 3'exons differ markedly. What is an exon for the female-specific transcripts is part of the untraslated 3'end of the male-specific message.

The alternative RNA processing appears to be the result of the *transformer* genes. The Tra2 and female-specific Tral proteins bind specifically to a DNA sequence adjacent to the female-specific 3′splice site of the *dsx* pre-mRNA, and they recruit nonspecific splicing factors to this site. If *tra* is not produced, the *doublesex* transcript is spliced in the male-specific manner. The downstream 3' splice site is used, and a male-specific transcript is made. This encodes an active protein that inhabits female traits and promotes male traits. The Transformer proteins bid to sequences within the female-specific exon and activate the female-specific 3′ splice site. This activation of an otherwise unused female-specific 3′ splice site produces an mRNA encoding a female-specific protein that activates female-specific genes (such as those of the yolk proteins) and inhibits male development.

The functions of the Doublesex proteins can be seen in the formation of the *Drosophila* genitalia.

Target genes for the sex determination cascade

Numerous proteins in *Drosophila* are present in one sex and not the other. In females, these include yolk proteins and eggshell (chorion) proteins. In males,

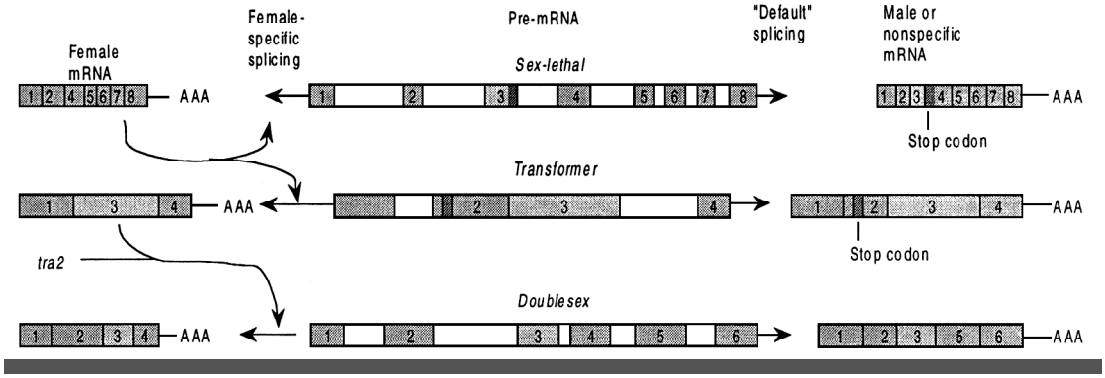


Fig. 2.6 The pattern of sex-specific RNA splicing in three major *Drosophila* sex-determining genes. The pre-mRNAs are located in the center of the diagram and are identical in both male and female nuclei. In each case, the female-specific transcript is shown at the left, while the default transcript (whether male or non-specific) is shown to the right. Exons are numbered, and the positions of the termination codons and poly(A) sites are marked. (After Baker, 1989)

the sex combs of the legs are sex -specific structure. Both the male and female *doublesex* transcripts bind to three sites within the 127-base-pair enhancer of the *yolk* protein genes. Their binding and mutagenesis studies demonstrate that the male-specific Doublesex product inhibits transcription by its binding to these sites, whereas the female-specific Doublesex protein activates gene transcription from the same sites.

2.4 Genetic regulation of sex determination and gonadal differentiation in humans

It is clear that only a small region of the Y-chromosome is endowed with the gene(s) for sex determination. The mystery of the sex reversed (Sxr) male mice having XX-chromosome constitution was resolved unambiguously by Singh and Jones (1982) who showed that one of the 2 X-chromosomes in the Sxr males carried a min]ute segment of the short arm of the Y-chromosome.

Testis determining gene on the Y-chromosome (TDY)

Subsequent studies led to the discovery of SRY gene (sex determining region of the Y) that coded for an HMG domain DNA binding protein. SRY is present on the short arm of human Y, and is conserved. The expression of Sry occurred from the onset of differentiation of medulla (primordial testis) in the genital ridge. Clinching evidence in favour of Sry as the male determining gene came from the sex reversal of the XX embryos to male through insertion of only a 14kb fragment of Y-chromosomal DNA having Sry. Transgenic mouse males are sterile the role of Sry as the male determining factor is confirmed. This single-exon gene codes for a protein that has a 79 amino acid long HMG (High mobility group proteins)-domain. The proteins harbouring this domain constitute a SOX (Sry-box) family of transcription factors that bind to a heptamer of nucleotides A/TAACAAT. Mutations in the HMG domain or in the upstream promoter region of Sry have been shown to result in XY pseudohermaphrodites, gonadal dysgenesis and other gonadal pathologies. The mouse Sry carries a polyglutamine stretch, which is absent in the human Sry.

Autosomal genes

XY individuals have however been reported with anomalies of gonad and urogenital system, in spite of having completely normal SRY (Sry) gene. More than 75% of the XY individuals suffering from acute dwarfism due to Campomelic dysplasia (a rate skeletal disorder) are hermaphrodite with ambiguous genitalia. This disorder is caused by mutation in a gene SOX9. Importance of SOX9 (Sox9 in mouse) in testis differentiation as well as subsequent organogenesis of the male genital system has since been established. An X-linked transcription factor, DAX1 (Xp21.2-22.2), has repressive effect on male development. The XY, SRY-positive pseudohermaphrodite individuals were found to have a duplication of DAX1 on the single X-chromosome, showing dosage-dependent effect of DAX1. Individuals suffering from WAGR (Wilms aniridia, genitourinary malformation, menl&l retardation) and Denys Drash syndromes both fail to have genital as

well as the renal development and are caused due to mutation in the WT-1 gene. Loss of function mutation in the autosomalWT-1, the Wilms1 tumour gene, also leads to gonadal dysgenesis. Similarly, mutation in SF-1 (orphan steroid factor-1 gene), has a broader effect on the mesonephric gonadal complex.

Interaction of genes in sexual differentiation

SF-1 and Sox9 bring about the activation of MIS gene, whose product leads to the regression of MD. These two transcription factors bind to the MIS promoter to induce its activity. WT-1 and GATA-4 act as cofactors of SF-1 and facilitate its binding to the MIS promoter. The X-linked, Dax-1, acts as a repressor of SF-1 in the female. Thus in the absence (or low level) of Sox9 and Sf-1, MIS (Mullerian Inhibiting Substances) is not produced and there is no regression of the Mullerian duct. Dax-1 is expressed also in males coincidentally with Sry, Sox9 and Sf-1 but its level is much lower than that in the female, the dosage of Dax-1 vis-a-vis Sry, Sox9, perhaps makes the difference between the differentiation of the male and female sex.

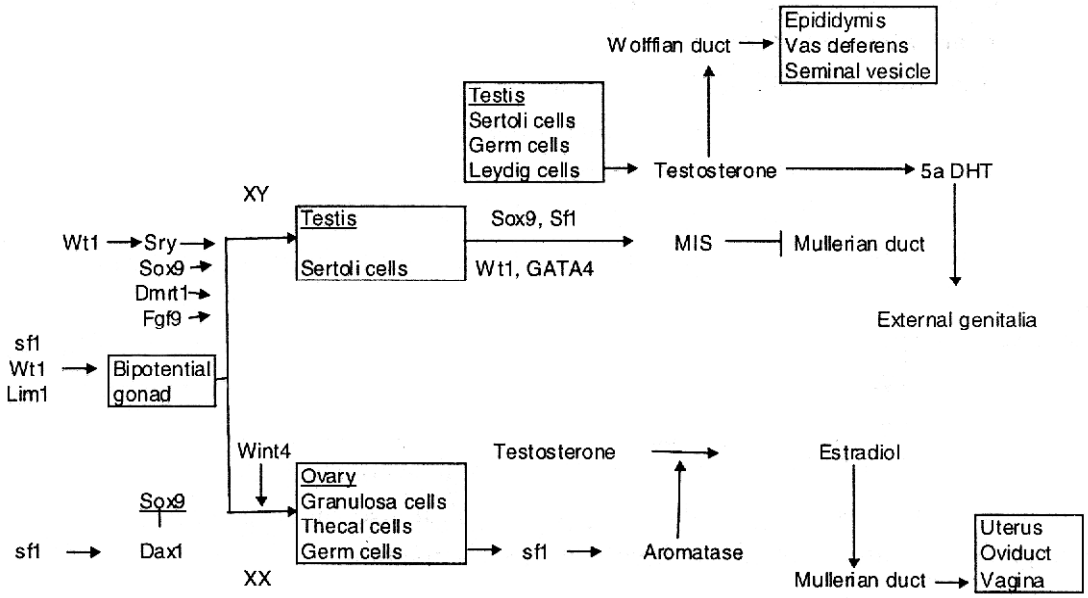


Fig. 2.7 Cartoon giving a simplified scheme of the genes participating in sex determination, and differentiation of the urogenital system in mammals

Obviously, Daxl is an "anti-male" rather than "female" gene. Its elevated level in female must be blocking not omy SF-1 induced activation of MIH but also testicular development.

More genes in the gonadal pathway

In addition a number of cases of sex reversals and ambiguous genitalia have been observed through mutations in other parts of the genome. Wnt4, coding for a cell-signaling molecule, is likely to have a positive role in female determination, unlike the "anti-testis" DAXl.

Like SF-1 artd WT-1, genes, LIM1 and LIM9, also act in the mesonephros to trigger the differentiation of GR (GR = Gonadal Ridge). Fgf9 mediates the migration of mesonephric cells in the testicular sex cords in GR of XY fetus. Therefore its absence is not detrimental to ovary in females. The XY-male specific expression of Vanin-1, a cell membrane asso

ciated G-protein, during testis differentiation may also be playing an important role in directing the mesonephric cells towards the testicular sex cords in the genital ridge.

It is obvious from the foregoing that though autosomal and X-linked genes play a role in the differentiation of the urogenital system, SRY is the key gene that switches the bipotential genital ridge towards testis determination. DNA as well as RNA binding functions are assigned to Sry (and other Sox proteins). It is intriguing that it occurs only in mammals. Functional SRY initiates a chemotactic action that leads to the migration of mesonephric cells into GR. It has also been demonstrated that the Sry expression in GR is not only temporally but also spatially restricted. The most important, role of Sry in gonadogenesis is to decisively direct the bipotential precursor cells in GR to for m the Sertoli cells instead of the follicular cells that form ovary. Centrally of SRY in testis determination is established. Nevertheless, which genes are the Sry-specific target gene(s) still remains a mystery. In a recent experiment, introduction of WT1 - promoter-driven Sox9 fusion gene succeeded in imparting male phenotype to the XX mouse, suggesting that if the Sox9 level could be raised then Sry was redundant for testis formation. The immediate conclusion from this evidence is that SOX9 may be the immediate target of SRY. It also suggests that WT-1 which is expressed upstream of Sry is involved in its activation.

Although many sex-chromosomal as well as autosomal genes play a role in mammalian (human) sex determination pathway, only few of them have' been tabulated below (Table 2.1).

Table 2.1 Genes Commonly Involved in Sex Determination in Man, Other Mammals and Lower Vertebrates

Genes	Gene product	Mutant Phenotype in man/mouse	Mutational mechanisms	Orthologue in other vertebrates
SRY	protein with an HMG domain	XY gonadal dysgenesis	loss of function	None (C. versicolor?)
WT1	Zn finger	WAGR syndrome Denys-Drash syndrome, Agenesis of gonads, kidney	Haplo-insufficiency, Dominant negative, loss of function	TSD-reptiles, birds
SF-1	orphan steroid receptor	Agenesis of urogential system	loss of function	TSD-reptiles, birds
SOX9	SRY-like HMG domain	XY-sex reversal, Campometic dysplasia,	loss of function	fish (Sox9a, b), reptiles, birds
DAX1/DSS	nuclear receptor	AHC, HGG XY SEX reversal	loss of function duplication	reptiles, birds
MIS	TGF-β gene family	persistance of Mullerian duct	loss of function	birds, TSD-reptiles
DMRT1	protein with dm-vdomain	XY-sex reversal	loss of function	fish, amophibia, reptiles, birds
Aromafase	enzyme in steroid biosynthesis	female to male sex reveral in birds & reptiles	loss of function	reptiles, birds
Wint4	signal transduction	XX, Leydig like cells	loss of function	—
Liml, Lirn9	homebox	Lim-1-loss of gonad, kidney Lim9-gonadal agenesis	loss of function	—
Fgf9	growth factor	XY abnormal gen italia	loss of function	—
Vanin2	G-protein	XY-abnormal genitalia	loss of function	—

2.5 Suggested questions

1. Describe the molecular mechanisms that regulate *Sxl-* promoter activity and its maintenance.
2. With the help of diagrams, describe the events of alternative splicing that lead to sex determination in *Drosophila.*
3. Briefly describe the cascade of events that lead to sex determination in C. *elegans.*
4. Explain the role of xol-J in C. *elegans sex* determination.
5. State and explain the effect on sex determination and dosage compensation due to mutation in the following genes:
   (a) Loss of function of *fox-l* in XX
   (b) Gain of function of *Xol* in XO
   (c) Gain of function of *Sxl* in XY
   (d) Loss of function of *da* in XX and XY
   (e) Gain of function of *Sdc* in XX
6. With the help of diagram briefly describe the initial pathway that regulates sex determination in Human.

Unit 3 Imprinting of Genes, Chromosomes and Genomes

Structure

3.1 Introduction

In humans and other mammals, several biallelic genes are known where the expression of one parental allele, either the paternal or the maternal allelebut not both, is *normally* repressed in some cells (**allelic exclusion**). In such cells the relevant gene is said to exhibit functional hemizygosity; *even although the sequences of both parental alleles are perfectly consistent with normal gene expression and may even be identical*. In some cases the allelic exclusion may be a property of select cells or tissues while in other cells of the same individual both alleles may be expressed normally. A variety of different expression mechanisms can be involved and two broad classes of mechanism are involved :

- *Allelic exclusion according to parent of origin* (*imprinting*). The choice of which of the two inherited copies is expressed is not random. This means that for some genes the allele whose expression is repressed is always the paternally inherited allele; in others it is always the maternally inherited allele.
- *Allelic exclusion independent of parent of origin*. Here the decision as to which of the two alleles is repressed is initially made randomly, but afterwards that pattern of allelic exclusion is transmitted stably to daughter cells following cell division. A variety of different mechanisms may be involved. A unique form of control is the programmed DNA rearrangements.

3.2 Genomic imprinting

Genomic imprinting is an epigenetic phenomenon, which, in most cases, is believed to occur during gametogenesis. Genomic imprinting occurs when both maternal and paternal alleles are present, but one allele will be expressed while the other remains inactive. The most prominent assumption is that this process is necessary for development and may somehow regulate growth in the embryo and neonate. Some

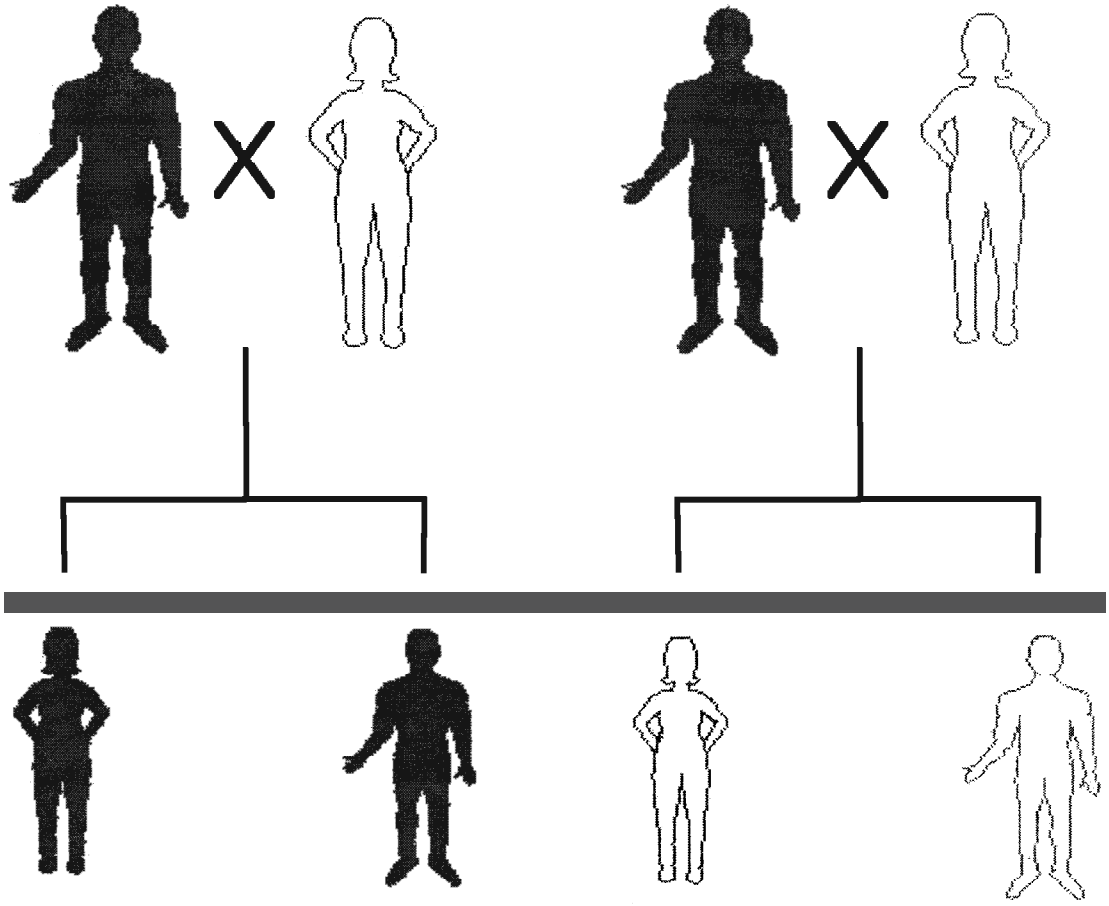


Fig. 3.1 Two examples of a hypothetical imprinted gene responsible for body color. (LEFT) In this example the pigment gene is maternally imprinted (maternal allele is inactivated). Matings between a male who possesses the allele for pigment and a female who possesses the allele for no pigment produces offspring that show only the pigmented phenotype. In this example, the mother's allele is imprinted and inactivated in the offspring. Therefore, the only actively-expressing allele is the father's pigment allele, which is not imprinted in the offspring. (RIGHT) In this example the pigment gene is paternally imprinted (paternal allele is inactivated). Matings between a male who possesses the allele for pigment and a female who possesses the allele for no pigment produces offspring that show only the pigmented phenotype. In this example, the father's allele is imprinted and inactivated in the offspring. Therefore, the only actively expressing allele is the mother's no pigment allele, which is not imprinted in the offspring. (Figure courtesy of Ross McGowan, Dept. Zoology, University of Manitoba)

mechanism must be able to distinguish between maternally and paternally inherited alleles: as chromosomes pass through the male and female germlines they must acquire some imprint to signal a difference between paternal and maternal alleles in the developing organism (Fig. 3.1).

An optimal method for gene imprinting, at least in maintaining the imprinted status, is allele-specific DNA methylation. The imprinting of several

imprinted genes has been shown to be disrupted in mutant mice that are deficient in the ***Dnmt 1*** cytosine methyltransferase gene and all imprinted genes are characterized by CG-rich regions of differential methylation. This process is carried out with the enzyme DNA methyltransf erase (DNA MTase) in mammals. DNA MTase acts on the DNA sequence 5'-CpG-3'. Some organisms (primarily higher eukaryotes) have aggregates of CpG (known as CpG islands) in their genomes. These islands are rarely methylated in animal cells. This may be due to the bound transcription factors that block DNA MTase. De novo methylation and maintenance of methylation are two distinct processes that are required for establishment and mitotic inheritance of tissue specific methylation patterns. *Dnmt1* is the major maintenance methyltransferase. ***Dnmt3a*** and ***Dnmt3b*** are essential for de novo methylation. And Sequences that are methylated are usually not active (Gold and Pedersen, 1994). Recent investigations, however, have shown that this is not always the case (Li *et al*., 1993).

It has been postulated that if a mutation was introduced to the DNA MTase gene in the embryonic stem cells of mice, the methylation of CpG would be abnormal, and gene expression would be affected. The mutation of the DNA MTase gene was caused by homologous recombination. The three genes used in this experiment were HI9, Igf2 (insulin-like growth factor) and Igf2r (Igf2 receptor).

For the H19 gene, it is the maternal allele that is expressed, while the paternal allele is silent. It should be noted that the inactive paternal allele is methylated while the maternal allele is not. It was shown that typical DNA methylation is a requirement to keep the paternal allele inactive for the H19 gene, a result that is consistent with the hypothesis.

In contrast to the H19 gene, the Igf2 gene is expressed only from a methylated paternal allele. It has now been concluded that a normal level of DNA methylation is needed for expression of the paternal Igf2 allele.

The gene Igf2r (insulin like growth factor receptor) is expressed from a methylated maternal allele. DNA methylation by DNA MTase is a requirement for the expression of the Igf2r gene.

Disruption of primary imprinting during oocyte growth leads to the modified expression of imprinted genes during embryogenesisThus far, experiments have not demonstrated how the imprinting process is regulated.lt has been proposed that further research should be attempted to discover and isolate imprinting genes . Already, some progress is being made in these areas. For instance, genomic imprinting has been implicated in cancer and shown to be involved in chromosomal deletion syndromes, such as Prader-Willi and Angelman Syndromes (Peterson and Sapienza, 1993).

3.3 Uniparental disomy and genomic imprinting

Uniparental disomy refers to the presence of two copies of a chromosome (or part of a chromosome) from one parent and none from the other. Several additional disorders resulting from uniparental disomy of single genes or multiple genes (including whole chromosomes) have been reported. A readily detectable adverse outcome of uniparental disomy is the consequence of a newly recognized phenomenon *called genetic or genomic imprinting.*

The first recognized example of such human abnormality resulting from the presence of uniparental disomy of an imprinted part of the genome was in Prader-Willi syndrome (PWS). Uniparental maternal disomy for chromosome 15

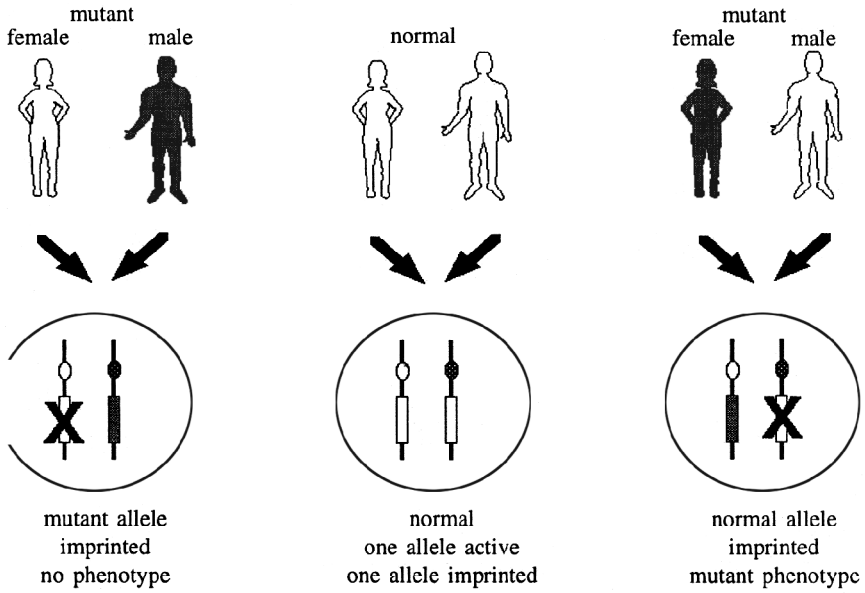


Figure 3.2 Schematic representation of the phenotypic effects of maternal imprinting of a mutant allele. Darkened body indicates individual that is mutant for the hypothetical imprinted locus. A cross is used to indicate the imprinted/inactive allele. (CENTER) Both parents are homozygous for the normal allele at the imprinted locus. Although only one allele is active (the paternal copy) in the offspring produced from these parents, it must be a normal allele and therefore all offspring will have a normal phenotype. (LEFT) The mother is homozygous mutant at the imprinted locus, and the father is normal. Since this hypothetical locus is maternally imprinted, the maternal mutant copy will be inactivated in their offspring and the paternal normal copy will be the only active allele. The offspring will be phenotypically normal, and the mutant allele will appear to be a recessive mutation. (RIGHT) The mother is homozygous normal at the imprinted locus, and the father is homozygous mutant. The maternal normal allele is imprinted and inactivated in the offspring of these parents. The only allele that is active is the mutant paternal copy. Therefore, all offspring produced from these parents will display the mutant phenotype, and the mutant allele will appear to be a dominant mutation. (Figure courtesy of Ross McGowan, Dept. Zoology, University of Manitoba)

is thought to cause Prader-Willi syndrome because there is absence of needed paternally contributed genes in the critical PWS region (del 15q 11 -q 13). The paternal contribution is hypothesized to be necessary because the homologous maternally derived genes are inactivated or imprinted (perhaps by methylation).

Interestingly, a very different disorder called Angelman syndrome also involves imprinting of the same chromosome region - only in Angelrnan syndrome the maternal contribution of the critical region is missing. The terminology used to describe the role of imprinting in these two disorders is somewhat confusing but goes as follows.

It is hypothesized that the critical genetic region which determines Prader-Willi syndrome is *maternally* imprinted (i.e. inactivated when inherited from the mother), whereas the critical region which determines Angelman syndrome is *paternally* imprinted (i.e. inactivated when inherited from the father). Both disorders result when the expected active genetic contribution from one parent is missing, either by deletion or uniparental disomy (Fig.3.2).

Interestingly, a number of human congenital tumors show evidence of genomic imprinting. For example, in cells from Wilms' tumor, loss of the maternal chromosome 11 is common. This suggests that the maternal chromosome 11 has a tumor suppresser role not present on the paternal 11. This phenomenon in relation to cancer is referred to as "loss of heterozygosity".

3.4 Suggested questions

1. How is genomic imprinting established and maintained?
2. Explain the phenomena and importance of uniparental disomy.
3. Differentiate paternal and maternal imprinting.

Unit 4 Somatic Cell Genetics

Structure

4.1 Cell fusion and hybrids-agents and mechanisms of fusion

Cultured animal cells infrequently undergo cell fusion spontaneously. The fusion rate, increases greatly in the presence of certain viruses that have a lipoprotein envelope similar to the plasma membrane of animal cells. Cell fusion is also promoted by polyethylene glycol, which causes the plasma membranes of adjacent cells to adhere to each other and to fuse. As most fused animal cells undergo cell division, the nuclei eventually fuse, producing viable cells with a single nucleus that contains chromosomes from both "parents." The fusion of two cells that are genetically different yields a hybrid cell called a **heterokaryon** (Fig. 4.1).

Because some **somatic cell** from animals can be cultured from single cells in a well-defined medium, it is possible to select for genetically distinct cultured animal cells. Genetic studies of cultured animal cells are called *somatic-cell genetics* to distinguish them from *classical genetics,* which deals with whole organisms derived from **germ cells** (sperm and eggs).

Assigning genes to chromosomes

The technique of somatic cell hybridization is extensively used in human genome mapping, but it can in principle be used in many different animal systems. The procedure uses cells growing in culture. A virus called the Sendai virus has a useful property that makes the mapping technique possible.

If suspensions of human and mouse ceils are mixed together in the presence of Sendai virus that has been inactivated by ultraviolet light, the virus can mediate fusion of the cells from the different species. The initial fusion products are described as heterokaryons because the cells contain both a human and a rodent nucleus. Eventually, heterokaryons proceed to mitosis, and the two nuclear

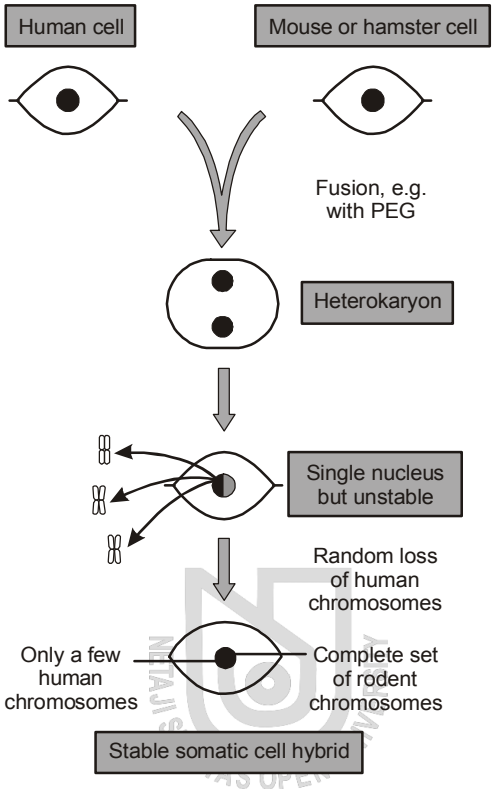


Fig. 4.1 Fusion of cells from different species can result in stable somatic cell hybrids. The example shows how stable human-rodent somatic cell hybrids can be generated following initial fusion using polyethylene glycol (PEG). For reasons that are not understood, human chromosomes are selectively lost from the initial fusion products. The loss occurs essentially at random so that eventually the stable products of a single fusion experiment will include a variety of cells with different complements of human chromosomes. They can be cloned to establish individual cell lines with a specific complement of human chromosomes. The identity of the human chromosomes can be established by PCR-based typing for chromosome-specific markers

envelopes dissolve. Because the mouse and human chromosomes are recognizably different in number and shape, the two sets in the hybrid cells can be readily distinguished. However, in the course of subsequent cell divisions, for unknown reasons the human chromosomes are gradually eliminated from the hybrid at random.

The loss of human chromosomes can be arrested in the following way to encourage the formation of a stable partial hybrid. The cells used are mutant for some biochemical function; so, if the cells are to grow, the missing function must be supplied by the other genome. This selective technique results in the maintenance of hybrid cells that have a complete set of mouse chromosomes

and a small number of human chromosomes, which vary in number and type from hybrid to hybrid but which always include the human chromosome carrying the wild-type allele defective in the mouse genome.

4.2 Heterokaryon-selecting hybrids and chromosome segregation

In cells, DNA can be made either de novo ("from scratch") or through a salvage pathway that uses molecular skeletons already available. The selective technique involves the application of a chemical, aminopterin that blocks the de novo synthetic pathway, confining DNA synthesis to the salvage pathway. Two essential salvage enzymes, thymidine kinase (TK) and hypoxanthine-guanine phosphoribosyl transferase (HGPRT), are relevant to the system, as shown in the following two reactions :

$$\text{thymine} \xrightarrow{\text{TK}} \text{thymidilic acid (a DNA precursor)}$$

$$\text{hypoxanthine} \xrightarrow{\text{HGPRT}} \text{inosinic acid (a DNA precursor)}$$

The mouse cell line to be fused is genetically unable to make TK because it is homozygous for the allele tkr, whereas the human cell line is genetically unable to make HGPRT because it is homozygous at another locus for the allele hgprT. So the genotypes of the two fusing cell lines are :

Mouse : tk^- / tk^-; $hgprt^+ / hgprt^+$

Human : tk^+ / tk^+; $hgprt^- / hgprt^-$

Because each is deficient for one enzyme, neither the mouse nor the human cells are able to make DNA individually. In. the hybrid cells, however, the tk^+ alleje complements the $hgprt^+$ allele, so the cells can make both enzymes. Therefore, DNA is synthesized and the cells can proliferate. Most human chromosomes are eliminated from the hybrid cell cultures because their loss has no effect on the cultures' ability to grow. But, to continue to grow in medium containing hypoxanthine, aminopterin, and thymidine (*HAT medium*), a hybrid culture must retain at least one of the human chromosomes that carry the tk^+ allele.

4.2.1 Selecting for the chromosome contents of hybrids

Hybrids can be selected for retention of a given human chromosome or chromosome fragments if it corrects an otherwise lethal abnormality in the rodent cell. Frequently used systems include :

- **Hybrid cells often are selected on HAT medium :** The medium most often used to select hybrid cells is called HAT *medium,* because it contains hypoxanthine (a purine), aminopterin, and thymidine. Normal cells can grow in HAT medium because even though aminopterin blocks de novo synthesis of purines and TMP, the thymidine in the media is transported into the cell and converted to TMP by TK and the hypoxanthine is transported and converted into usable purines by HGPRT. On the other hand, neither TK nor HGPRK cells can grow in HAT medium because each lacks an enzyme of the salvage pathway. However, hybrids formed by fusion of these two mutants will carry a normal *TK* gene from the HGPRK parent and a normal *HGPRT* gene from the TK parent. The hybrids thus will produce both functional salvage-pathway enzymes and grow on HAT medium. Likewise, hybrids formed by fusion of mutant cells and normal cells can grow in HAT medium. Somatic cell hybrids can be forced to retain human chromosome 17 by using thymidine kinase deficient (TK) rodent cells and growing the hybrids in *HAT* (hypoxanthine-aminopterin-thymidine) medium.
- **G418 selection** : Hybrids can be selected for the presence of a particular human chromosome segment if it has been tagged by incorporation of a neomycin resistance (neo^R) gene. The neomycin analog G418 kills nonresistant cells. NeoR is a typical example of a dominant selectable marker.

Selecting for the chromosome contents of hybrids

Hybrids can be selected for retention of a given human chromosome or chromosome fragments if it corrects an otherwise lethal abnormality in the rodent cell. Frequently used systems include :

- **HAT selection :** Somatic cell hybrids can be forced to retain human chromosome 17 by using thymidine kinase deficient (TK) rodent cells and growing the hybrids in *HAT* (hypoxanthine-aminopterin-thymidine) medium. TK cells are killed in HAT medium, but are rescued by the human TK gene on chromosome 17.
- **G418 selection :** Hybrids can be selected for the presence of a particular human chromosome segment if it has been tagged by incorporation of a neomycin resistance (neo^R) gene. The neomycin analog G418 kills nonresistant cells. Neo^R is a typical example of a dominant selectable marker.

4.2.2 Somatic cell hybrid panels can permit chromosomal localization of any human DNA sequence

The human chromosomes in somatic cell hybrids can conveniently be identified by PCR screening with sets of chromosome-specific primers. By

collecting hybrid cell lines with different human chromosome contents it is possible to generate a hybrid cell panel that can be used to map any human DNA sequence to a specific chromosome. To do this, each of the hybrid cell lines is tested for the presence of the human sequence of interest. A PCR assay can be used with primers specific for that sequence or the relevant DNA sequence can be labeled and used as a hybridization probe.

4.3 Radiation hybrids, hybrid panels and gene mapping

Subchromosomal mapping is possible using hybrid cells containing defined portions of a human chromosome

Conventional somatic cell hybrids are a relatively crude tool for physical mapping. More refined mapping is possible using hybrids that contain only part of a particular human chromosome. Translocation hybrids and deletion hybrids are made using donor human cells that have a chromosomal translocation or deletion. To be useful, the hybrids must lack the normal homolog of the chromosome of interest. Such hybrids can be used for subchromosomal mapping of a human sequence-tagged site or biochemical marker. (Fig. 4.2). They are especially useful for defining the sequences removed by microdeletions, by segregating the deletion-carrying chromosome away from its normal homolog.

Chromosome-mediated gene transfer

One of the first techniques to use this approach was **chromosome-mediated gene transfer (CMGT)**. Fragments of purified mitotic chromosomes from a donor, such as a human fibroblast, are coprecipitated with calcium phosphate on to the surface of a recipient rodent cell line in monolayer culture. Human chromosome fragments enter the recipient cells, such as mouse fibroblasts, and integrate into the chromosomes, resulting in stable transformation. As a result, hybrids can be established that retain segments of human DNA (**transgenomes**) of a size that is useful for mapping (usually in the range of 1-50 Mb). However, the transgenomes are prone to frequent rearrangements, so CMGT is more suited to functional assays of complex loci than as a mapping tool.

Irradiation fusion gene transfer

The most valuable hybrids for gene mapping are radiation hybrids. Donor cells are subjected to a lethal dose of radiation, which fragments their chromosomes. The average size of a fragment is a function of the dose of radiation. After irradiation the donor cells are fused with recipient cells of a different species. A selection system is used to pick out recipient cells that have taken

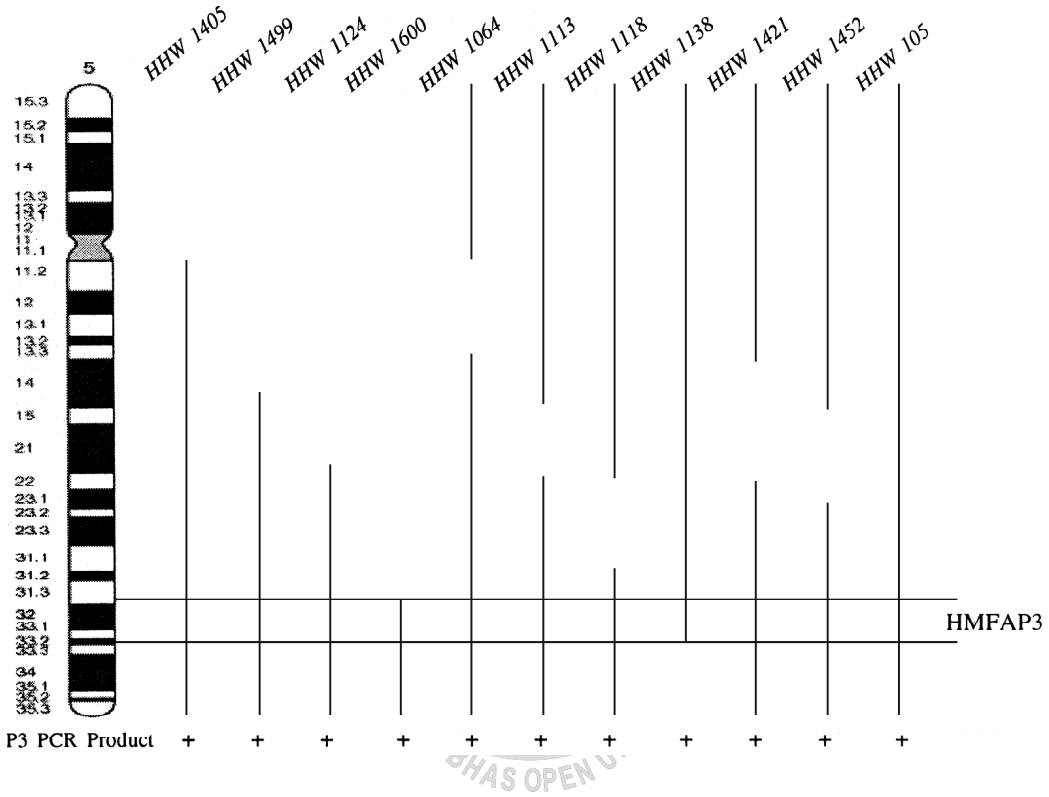


Fig. 4.2 Subchromosomal localization can be achieved by mapping against a panel of hybrid cells containing translocation or deletion chromosomes. The figure illustrates PCR-based mapping of the human microfibrillar protein MFAP3 using a panel of 5q translocation and deletion hybrids. Vertical black bars to the right indicate the extent of human chromosome 5 sequences, which are retained in the hybrids. Hybrids HHW1405, 1499, 1124 and 1600 contain translocation chromosomes with 5q breakpoints and retention of the segment distal to the breakpoint. By contrast, translocation hybrid HHW1138 retains material proximal to the 5q breakpoint. Hybrids HHW1064, 1113, 1118, 1421 and 1452 have different interstitial deletions of 5q. The solid blue vertical bar to the left indicates the inferred subchromosomal location as defined by breakpoints in hybrids HHW1600 and HHW1138 (blue horizontal lines near bottom). Reproduced from **Abrams et al. (1995)** Genomics, **26**, pp. 47-54, with permission from Academic Press, Inc

up some of the donor chromosome fragments. These cells are useful for mapping insofar as they have taken up a random set of other chromosome fragments from the donor, as well as the selected fragment. Stably incorporated donor fragments are either integrated into rodent chromosomes or are assembled into novel human minichromosomes formed around fragments containing a functional centromere. When a set of DNA markers from the human chromosome is assayed in a panel of such radiation hybrids, the patterns of cross-reactivity can be used to construct a map. (Fig. 4.3).

The principle is very similar to meiotic linkage analysis : the nearer together two DNA sequences are on a chromosome, the lower the probability that they

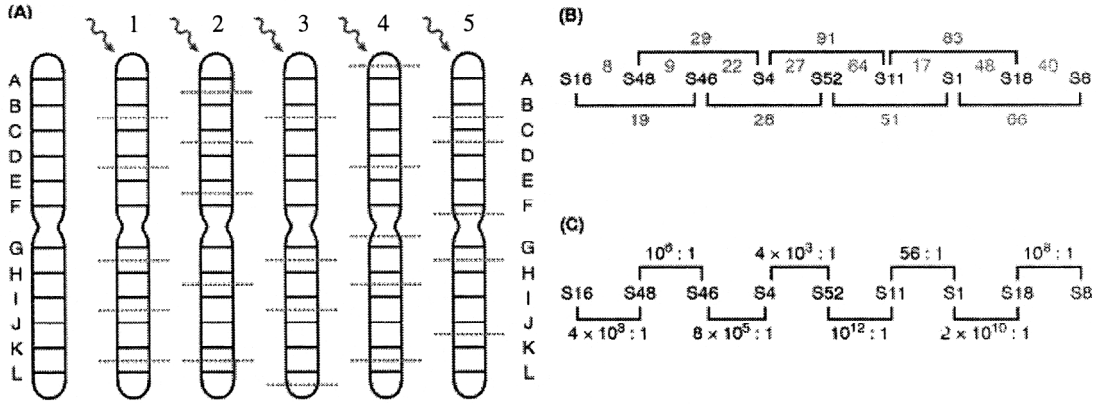


Fig. 4.3 Constructing radiation hybrid maps. (**A**) Breakpoints occur randomly. Five possible examples of breakpoints (dashed blue lines) on the same type of chromosome are shown. Markers close together will tend to occur on the same fragment, e.g. A and B in all cases other than example 2. Thus, if a radiation hybrid contains marker A it will frequently also contain marker B, but rarely a distant marker such as L. (B) Ordering of markers on human 21q. The order of markers *D21S16-D21S8* as inferred by *Cox et ai* (1990) from radiation hybrid mapping is shown. Figures on the top panel refer to distances between markers in centiRays8000. For example, the SO 6-S48 interval is 8 cR8000: at a radiation dose of 8000 rad, there is 8% frequency of breakage between them, and so a 92% chance they will occur together on one fragment. (C) Odds ratios refer to the likelihood of the indicated order for pairs of markers compared with that with the markers inverted. For example, the calculated likelihood for the order S16-S48-S46-S4 is 106 times greater than for the order S16-S46-S48-S4

will be separated by the chance occurrence of a breakpoint between them. The frequency of breakage between two markers can be defined by a value, analogous to the recombination frequency in meiotic mapping. The value,varies from 0 (the two markers are never separated) to 1.0 (the two markers are always broken apart). As in meiotic mapping, the value underestimates the distance between markers that are far apart on the same chromosome, in this case because a cell can take up two markers on separate fragments. A more accurate estimate is provided by a mapping function, D = -In (1 - value), which is analogous to the Haldane mapping function used in meiotic linkage analysis. D is measured in centiRays (**cR**). *D* is dependent on the dosage of radiation, so it is referenced against the number of rads. For example, a distance of 1 cR_{8ooo} between two markers represents a 1% frequency of breakage between them after exposure to 8000 rad of X-rays.

Radiation hybrids derived from monochromosomal hybrid donor cells have been superseded by whole-genome radiation hybrids where the donor is an irradiated normal human diploid cell. The first such panel consisted of 199 hybrids made by fusing an irradiated 46,XY human fibroblast cell line to TK

hamster cells (Walter *et al.*,>1994), Gyapay et al. (1996) used 404 microsatellite markers of known location to show that this hybrid panel could generate accurate maps, and then used it to map 374 unmapped ESTs. A subset of 93 of the hybrids has been made widely available as the Genebridge 4 panel. The 93 hybrids average 32% retention of any particular human sequence, with an average fragment size of 25 Mb. Laboratories can map any unknown STS by scoring the 93 Genebridge hybrids and comparing the pattern with patterns of previously mapped markers held on a central server.

This has turned into an extremely powerful and convenient tool for physically mapping any STS or EST. A second human-hamster panel, Stanford G3, was made using a higher dose of radiation, so that the average human fragment size is smaller. The 83 hybrids in G3 average 16% retention of the human genome, with an average fragment size of 2.4 Mb. Thus G3 can be used for finer mapping. The impressive results of large-scale use of these panels can be accessed at http://www.ncbi.nlm.nih.gov/genemap98/.

4.4 Suggested questions

1. What is heterokaryon? What are the methods of heterokaryon selection?
2. Explain the principle of mapping by radiation hybrids.
3. How can genes be assigned to specific regions on chromosomes by low-resolution mapping (radiation hybrids)?

4.5 Suggested books

1. Albert', B. et at. (2003) *Molecular Biology of the Cell,* Fourth edition; Garland Sciences, New York.
2. Lewin, B. (2004) *Genes* VIII, John Willey, New York.
3. Klug, W, and Cummings, M. (2003) *Concepts of Genetics;* Seventh Edition; Pearson Education, Singapore.
4. Karp, G. (2005) *Cell and Molecular Biology,* Fourth Edition; John Willey, New York.
5. Gilbert, S. (1997) *Developmental Biology,* Fifth Edition, Sinauer Associates Publishers, Massachusetts.
6. Russell, P. (1998) Genetics, Fifth Edition; Addison Wesley Longman, New York.
7. Strachan, T. and Read, A. P. (2004) *Human Molecular Genetics,* Third Edition; Oxford University Press, Oxford.

Unit 5 Human Cytogenetics

Structure

5.1 Techniques in human chromosome analysis

5.1.1 Introduction

The correct chromosome number for man was established only after the *application of tissue culture methods* to cytogenetics. Before 1956, the chromosome number of man was considered to be 48. Tjio and Levan (1956), analyzing cultures of human embryonic lung fibroblasts, found a consistent chromosomal number of 46. At the same time and independently Ford and Hamerton (1956), using meiotic cells obtained from testicular biopsy material, found only 23 pairs of chromosomes. Mitotic cells from the same specimens contained 46 chromosomes. These two papers mark the beginning of modern human cytogenetics. The 15-year period from 1956 to 1971 saw the development of a standardized system of nomenclature which became more refined as the identification of human chromosomes became more precise. The delineation of most of the syndromes associated with chromosomal abnormalities occurred during syndromes associated with chromosomal abnormalities occurred during this period.

5.1.2 Terminologies used in the identification of human chromosomes karyotype & idogram

Human chromosomes can be arranged in an orderly fashion to produce a *karyotype* (Rowley, 1969). A karyotype is comoposed of individual chromosomes from a particular cell; the chromosomes are aligned in pairs and identified according to the standard nomenclature accepted by cytogeneticists. Karyotypes of different cells will reflect the variations in chromosomal morphology present in these cells. An *idiogram* is the schematized drawing of a composite of many karyotypes and is not directly related to any particular cell. Idiograms generally are not prepared for clinical purposes. But are used for comparing chromosomal patterns of different species.

Metaphase chromosome

Metaphase chromosomes differ from one another in size and shape (Fig. 5.1) Each metaphase chromosome is identified by its size, shape, and specific banding pattern. The absolute size of any chromosome varies with the stage of mitosis. Chromosomes are longer and less coiled in prophase and shorter and more compact at the end of metaphase. The duration of treatment with mitotic blocking agents and the type of hypotonic solution also influence the absolute size of the chromosomes. In general, the longest human metaphase chromosome is about 7-8 urn in length, whereas the shortest is about 2μm long. Each metaphase chromosome is composed of two *chromatids* joined at the *centromere* (the site of attachment of the spindle fibre). The position of the centromere is specific for

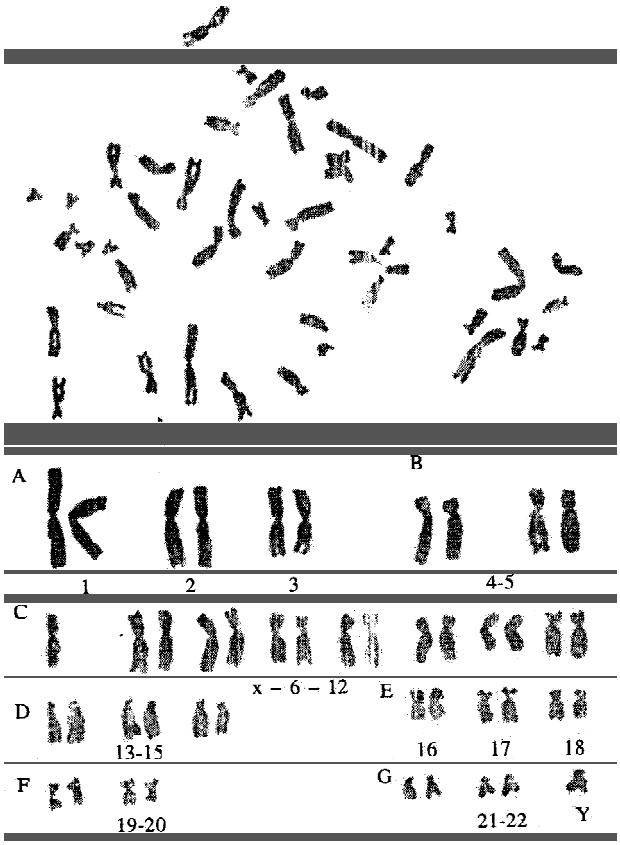


Fig. 5.1 Intact metaphase plate from a normal male with 46 chromosomes from bone marrow specimen (top) and karyotype of cell (bottom). Chromosomes are arranged in seven groups of morphologically similar chromosomes. The X chromosome is included with group C from which it cannot be distinguished. The Y chromosome in this cell is about the size of G-group chromosomes, but it can be differentiated from them by the size of the short arm

each chromosome and divides it into a long and short arm. The relative length of the two arms (arm ratio) is important for the identification of chromosomes.

A chromosome with a centreomere in the middle that divides it into two equal arms is called *metacentric.* When the centromere is somewhat nearer to one end of the chromosome, so that there is a distinct long and short arm, the chromosome is said to be submetacentric. If the centromere is very near to one end of the chromosome, which thus only a very short small arm and a relatively longer long arm, the chromosome is called *acrocentric* (Fig. 5.2). *A telocentric* chromosome, not normally found in human cells, has the centromere at the end.

5.1.3 Standardization of nomenclature of chromosome

Four conferences were held between 1960 and 1971 to revise the nomenclature used to identify human chromosomes along with the improvement of technologies. The last two conferences were held in Chicago and Paris. When the first reports on human chromosomal abnormalities appeared in 1956, each group of investigators used its own system of arranging and numbering the chromosomes. The necessity for adopting a uniform system was generally recognized, and a standard nomenclature of human mitotic chomosomes was adopted at the Denver Conference in 1960. Numbers were assigned to each pair of autosomes as nearly as possible in decending order of length. The sex chromosomes, XX pairs of autosomes (44) plus two sex chromosomes, making a total of 46 chromosomes. Seven groups of morphologically similar chromosomes could be distinguished (Table 5.1). It soon became evident, however, that not all pairs of chromosomes could be identified with certainty, even in preparations of the highest technical quality.

Patau (1960) proposed that the seven groups of morphologically similar chromosomes be indentified by capital letters A through G, an arabic number being added only when the individual chromosome could be identified with certainty. This system had the great advantage of flexibility, since it permitted general recognition of the group to which a chromosome belonged without implying identification of the specific chromosome involved in an abnormality. Patau's recommendations were accepted at the London Conference on the normal human karyotype (1963).

Whereas individual investigators were using the same systems (Denver and London) for identifyng individual chromosomes, they were using different systems for describing chromosomal abnormalities. This led to confusion when data from different laboratories were collated. A major achievement of the Chicago Conference in 1966 was the development of a uniform system of notation designed to facilitate coding for data retrieval. It was agreed that the analysis of the karyotype would be recorded with the total chromosomal number first, followed

by the sex chromosomes, and finally by any additional abnormalities. Thus the karyotype of a normal male was written 46, XY; a normal female was 46, XX. The recommended nomenclature symbols used in describing normal or abnormal chromosomes are summarized in Table 5.2.

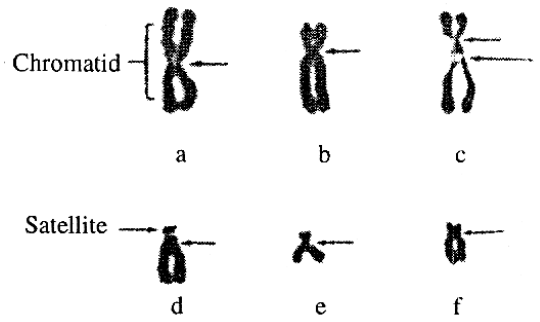


Fig. 5.2 Centromere position in typical chromosomes, a, Two chromatids joined at centromere or primary constriction. Centromere (arrow) is median and divides chromosomes in two equal arms. Chromosome is metacentric. b, Submedian centromere (arrow) divides the chromosome into short arm and long arm Chromosome is submetacentric. c, Submetacentric chromosome with secondary constriction (long arrow) in long arm. d, Large acrocentric chromosome with subteminal centromere (short arrow). Satellites (long arrow) are separated from short arm by secondary constriction, e and f, G-group chromosome and Y from same cell. Long arm chromatids of Y are close together and short arm is larger than G-group chromosome

Table 5.1 Systems of identification of human metaphase chromosomes*

Denver			Chromosome description	
Patau		Chromosome	Centromere	
group	Group	number	position	Morphology
A	1-3	1,3	Median	Metacentric
		2	Submedian	Less metacentric than above
B	4-5	4,5	Submedian	Submetacentric
C	X 6-12	X,6,7,9,ll	Submedian	Submetacentric but more metacentric than remainder
		8,10,12	Submedian	Less metacentric than above
D	13-15	13,14,15	Subtermina	Acrocentric—all may have satellites
E	16-18	16	Median	Metacentric
		17,18	Submedian	Submetacentric
F	19-20	19,20	Median	Metacentric
G	21-22	21,22	Subterminal	Acrocentric—all may have satellites
		Y	Subterminal	Acrocentric—similar in size to group G but usually morphologi cally distinct; long arms tend to be close together: satellites are not present

*Rowley (1969)

Table 5.1 Nomenclature symbols

Chicago Conference

A—G	the chromosome groups
1—22	the autosome numbers
X,Y	the sex chromosomes
diagonal(/)	separates cell lines in describing mosaicism
?	questionable identification of chromosome or chromosome structure
*	chromosome explained in text or footnote
ace	acentric
cen	centromere
dic	dicentric
end	endoreduplication
h	secondary constriction or negatively staining region
i	isochromosome
inv	inversion
mar	marker chromosome
mat	maternal origin
p	short arm chromosome
pat	paternal origin
q	long arm of chromosome
r	ring chromosome
s	satellite
t	translocadon
s t repeated symbols	duplication of chromosome structure

Paris Conference

A. Recommended changes in Chicago Conference nomenclature

\+ 1. The + and — signs should be placed *before* the appropriate symbol where they mean additional or missing whole chromosomes. They should be placed *after* a symbol where an increase or decrease in length is meant. Increases or decreases in the length of secondary constriction, or negatively staining regions, should be distinguished from increases or decreases in length owing to other structural alterations by placing the symbol h hetween the symbol for the arm and the + or — sign (e.g., 16qth +).

2. All symbols for rearrangements are to be placed before the designation of the chromosome (s) involved in the rearrangement, and the rearranged chromosome (s) always should be placed in parentheses, e.g., r(18), i (Xq), die (Y).

B. Recommended aditional nomenclature symbols

del	deletion
der	derivative chromosome
dup	duplication
ins	insertion
inv ins	inverted insertion
rep	reiprocal translocation*
rec	recombinant chromosome
rob	Robertsonian translocation* (“centric fusion”)

tan	tandem translocation*
ter	terminal or end ("p ter" for end of short arm; "q ter" for end of long arm)
:	break (no reunion, as in a terminal deletion)
: :	break and join
→	from-to

* Optional, where greater precision is desired than that provided by the use of t as recommended by the Chicago Conference.

5.1.4 Recommendations of the Paris Conference

The fluorescent karyotype, published by Caspersson et ah (1971), was accepted as the basis for the assigning of numbers to each chromosome.

A. Definitions

The bands seen with the fluorescent dyes (quinacrine) were called Q-bands and were accepted as the reference bands. Those bands of chromatin stained by methods that demonstrate "constitutive heterochromatin" were called C-bands and they are mainly confined to the centromeric region. The bands stained with basic dyes such as Giemsa were called G-bands and, except in one techique, they correspond quite well with Q-bands. One of the techniques using Giemsa, the exception just noted, gives patterns that are opposite in intensity to the G-bands; these were called R-bands.

A band was defined as a part of a chromosome which is clearly distinguishable from its adjacent segments by appearing darker or lighter with the Q, G, R, or C staining methods. By definition there were no "interbands." In the construction of the chromosome map, each band was refered to by its midline and not by its margins. A chromosome *landmark* was defined as a consistent and distinct morphological feature that is an important diagnostic aid in identifying a chromoosome. A region was defined as any area of a chromomsome lying between two adjacent landmarks. A chromosome arm lacking a prominent landmark consists of only one region.

B. Band numbering

Regions and bands are numbered consecutively from the centromere outward along each chromosome arm (Fig. 5.3). A band used as a landmark is considered as belonging entirely to the region distal to the landmark and is accorded the band number of "1" in the region. A band bisected by the centromere is considered as two bands, each being labeled as band 1, in region 1, of the appropriate chromosome arm.

For the designation of a particular band, four items are required: the chromosome number, the arm symbol, the region number and the band number within that region. These items are given in order without spacing or punctuation.

For example, Ip33 indicates chromosome No. 1, short arm, region 3, band 3. If a band defined in the present chromosome map has to be subdivided, the original band designation will be followed by a decimal point and the sub-bands will then be numbered sequentially from the centromere outward, e.g., Ip33.1; Ip33.2; Ip33.3, indicating that the orginal band 33 in the short arm of chromosome No. 1 has been divided into three sub-bands, 33.1 being proximal and 33.3 distal to the centromere. This system is thus relatively simple and yet sufficiently flexible to accommodate further refinements as the banding techniques are improved.

C. Characterization of chromosomes by the various banding techniques

The technical quality of the chromosomes is of the utmost importance for characterizing chromosomes by their banding patterns; in fact, it is much more important than previously adopted methods when morphology was the sole criterion for preparing the karyotype. The most distinct banding patterns are obtained when the chromosomes are relatively long and free of overlaps. It is important to note that the morphological characteristics of size and centromere position remain critical parameters used in the identification of chromosomes. Thus, the cytogencticist uses banding patterns as well as overall morphology to distinguish individual chromosomes. In general, the number of distinct bands increases with increasing length of the chromosome. It is thus, meaningless to mention the absolute number of bands in a chromosome arm, since the number varies with the state of contraction, the quality of the preparation and the type of treatment and stain. Once identification of the individual chromosomes by means of the major landmarks is mastered, careful observation of very good preparations can reveal a number of fine bands which can be used, among other things, for defining the site of chromosomal breakage and rejoining.

The intensity of fluorescence is influenced by the position of the chromosome in the metaphase. Chromosomes in the center of the cell frequently fluoresce more brightly than homologous chromosomes that are on the periphery. It is therefore necessary to consider the *pattern* of the bands in a particular chromosome and to appreaciate the fact that the overall intensity of fluorescence of homologs may be different. There may also be "spreading" of fluorescence from brighter to duller chromosomes, such that a 19 or a 22 adjacent to an X or the long arm of the Y may appear to be much brighter than normal.

The terms "distal" and "proximal" refer to the position of a band with respect to the centromere. The following terms were used in the Paris report to indicate the approximate intensity of fluorescence:

negative	no or almost no fluorescence
pale	as on distal lp
medium	as the two broad bands on 9q
intense	as the distal half of 13q
brilliant	as on distal Yq

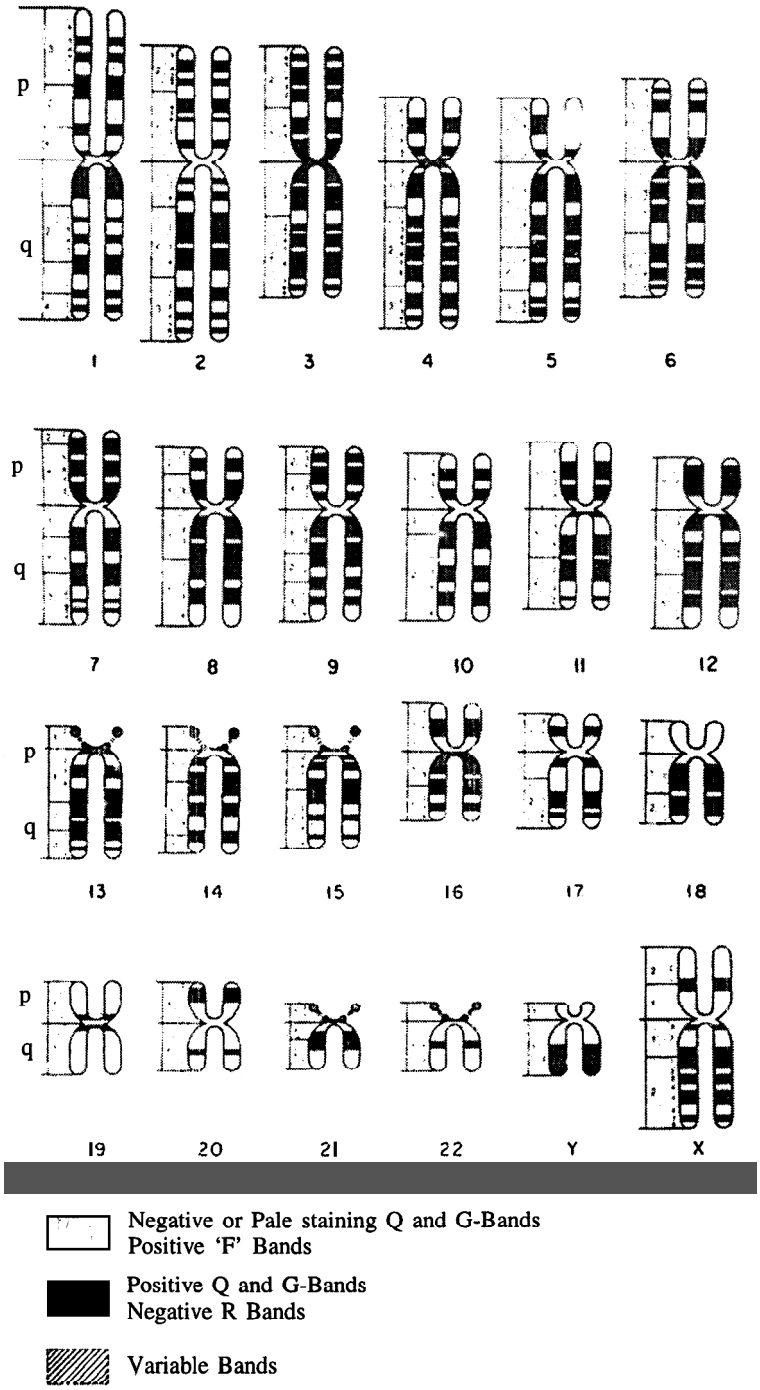


Fig. 5.3 Diagrammatic representation of chromosome bands as observed with the Q-, G-, and R-staining methods; centromere representative of Q-staining method only. Reproduced from the report of The Paris Conference (1972)

D. Characterization of chromosomes on the basis of different boarding techniques :

Chromosome 1 is the largest chromosome and is usually metacentric. The distal 40% of the short arm shows pale fluorescent bands (32 to 36), and the proximal segment shows two bands of medium fluorescence; the more proximal band divides region 1 from 2, and the more distal band divides region 2 from 3 (Fig. 5.4). The area of the secondary constriction, a poorly staining gap, which by definition is in the long arm, is adjacent to the centromere and shows negative fluorescence; it constitutes region 1. The long arm also contains a central intense band which divides region 2 from 3, with a less intense band distal to it which divides region 3 from 4.

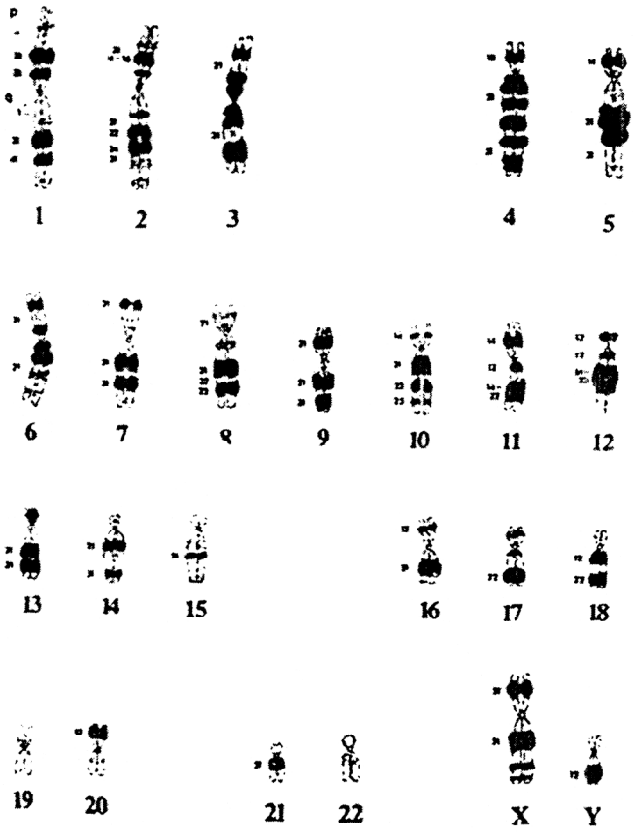


Fig. 5.4 Idiogram of fluorescent bands in human chromosomes (see text)

The large block of densely stained material in the long arm adjacent to theL centromere is the most prominent feature of No. 1 in cells treated to produce G-bands (Fig. 5.5). This area corresponds to the negatively quinacrine-stained secondary constriction. The two proximal bands in the short arm and the very

darkly stained central band and less darkly stained distal band in the long arm are also present. The end of the short arm is faintly stained.

The technique for staining C-bands revals the same densely staining region of the secondary constriction, generally called "constitutive heterchromatin." as do the G-band techniques. The R-banding technique (Dutrillaux and Lejeune, 1971) demostrates moderately staining material in this region; elsewhere along the chromosome, however, dark bands appear pale and vice versa. This reversal of staining intensity as compared with G-bands is particularly noticeable at the end of the short arm.

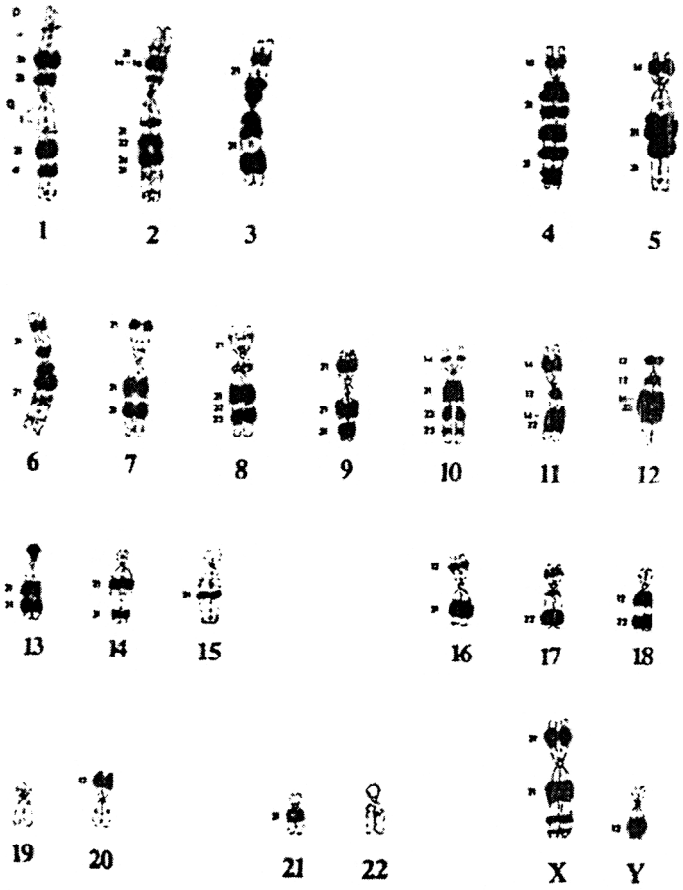


Fig. 5.5 Idiogram of Giemsa bands in human chromosomes (see text)

Chromosome 2 is a large chromosome, less metacentric than No. 1, and lacking conspicuous landmarks. It shows a number of bands of medium fluorescence along the entire length; the central band (14 to 16) in the short arm and the two cenral bands (22 to 24) of equal intensity in the long arm are most prominent. In the short arm, the dull band distal to the central band divides

region 1 from 2. In the long arm, the dull band proximal to the proximal central band divides region 1 from 2, whereas the dull band distal to distal central band divides region 2 from 3.

Chromosome 3 is a large metacentric chromosome which is smaller than No. 1. It has a nearly symmetrical banding pattern. There is a distinct band of pale fluorescence in the center of each arm which separates two broad bands of medium intensity. This pale band separates region 1 from 2 in each arm. The distal medium band in the short arm is narrower, but frequently appears more intrensely fluorescent than that in the long arm, whereas the terminal pale band is longer in the short than in the long arm.

Similar morpological features are observed in cells treated to produce G-bands. The centromeric regions stains darkly, but the variation in intensity of stain, seen with fluorescence, is not obsrved in G-bands.

Chromosome 4 is a long submetacentric chromosome that, similar to No. 2, lacks prominent landmarks. It has one band (15) of medium fluorescence in the short arm and four or five relatively evenly spaced bands of medium intensity in the long arm. The long arm is divided into three regions by a proximal dull band separating region 1 from 2 and a distal dull band separating 2 from 3.

Chromosome 5 has more distinctive characteristics than No. 4 and is frequently the first pair of B-group chromosomes indentified in the cell. The central band (14) of medium fluorescence in the short arm is brighter and frequently wider thanthat in No. 4. There is a broad central band of medium fluorescence which separates region 1 from 2 in the long arm, with a prominent distal pale band which separates region 2 from 3. Frequently, the terminal portion of the long arm of No. 5 is paler than No. 4.

Chromosome 6 is the largest and one of the last submetacentric C-group chromosomes. The most prominent feature is a distinct band of pale fluorescence in the middle of the short arm, which separates two bands of medium intensity and which also separates region 1 from 2. The long arm contains a number of bands of medium intensity; those bands near the centromere are frequently brighter and more idstinct. A dull band in the middle of the long arm separates region 1 form 2.

Chromosome 7 is slightly smaller than No. 6, but their centromere positions are similar. The two bands of intense fluorescence in the center of the long arm are the most prominent feature of this chromosome. The proximal band divides regions 1 and 2,whereas the distal band divides regions 2 and 3. There is a distinct band of medium fluorescence at the end of the short arm which divides region 1 from 2.

Chromosome 8 is one of the most submetacentric C group chromosomes, and it lacks distinctive landmarks. The short arm shows less intense fluorescence

than the long arm, and a cenrtral pale band which divides region 1 from 2 may be seen in good preparations. A distal pale band (22) separates two medium bands in the long arm, The more proximal of these medium bands separates region 1 from 2.

Chromosome 9 is the middle-sized C chromosome and is more metacentric than Nos. 8 or 10. The long arm shows a negatively staining centromeric region (12) with two distal evenly spaced bands of medium intensity. The proximal band of the pair may appear wider and separates region 1 from 2; the distal band separates 2 from 3. The short arm has a characteristic heart -shaped appearance with a central band of medium intensity which divides region 1 from 2, Both G and Q techniques show a prominent, faintly staining region near the centromere. The R-bands are the reverse, except for the centromeric region which is pale; the R-bands are similar in staining intensity to Q-and G-bands. The negatively fiuorescing and pale-staining region in the long arm near the centromere shows a large block of material that stains intensely with Giemsa after treatment to produce C-bands. This C-band material presumably represents one of the types of constitutive heterochromatin, the size of which may vary with individuals.

Chromosome 10 is one of the smaller, less matacentric C-group chromosomes which can be identified by the three evenly spaced bands in the long arm. The proximal band is most intense (and divides region 1 from 2) and the distal is the least intense. The short arm shows medium fluorescence.

Chromosome 11 is one of the least submetacentric C group chromosomes; and it may be very slightly larger than No. 10. It is somewhat similar to No. 9, but can be distinguished on the basis of the following features: There is a narrow medium band (12) in the middle of the negatively staining centromeric portion of the long arm. In poor preparations, this narrow band may be very faint or not apparent. A broad band of medium fluorescence is present in the middle of the long arm. A narrow dull band in the middle of this broad band separates region 1 from 2. This broad band usually appears as a *single* band that distinguishes it from the definite double band in the long arm of No. 9. The short arm shows medium fluorescence (14) and has a rather squarish appearance. In contrast, the short arm of No. 9 tends to taper near the centromere.

The C-banding technique reveals that the amount of centromeric stining material in No. LI, and the X is second only to that in No. 9.

Chromosome 12 is the most submetacentric of the C-group chromosomes; it is similar in size to No. 10. The short arm shows a band of mediumn fluorecence (12) which is smaller than that in No. 11, because the short arm is smaller..The band (12) of medium intensity in the long arm near the centromere is wider than that found in No. 11. A short band (13) of negative fluorescence separates band

ql2 from a distal long segment of medium intensity which divides region 1 from 2. This distal segment and the terminal dull band are both longer than the corresponding bands of No. 11.

The X *chromosome* is the third largest chromosome in the C group; it is, with 6 and 11, among the least submetacentric chromosomes in this group. The X chromosome is frequently the most fluorescent of the C-group chromosomes. Both Xs in the female show identical fluorescence patterns. In both arms, the region proximal to these medium bands shows pale fluorescence. Two other evenly spaced, less distinct and less intense bands of fluorescence are found in the distal long arm.

Chromosome 13 is the largest of the D-group chromosomes and shows intense fluorescence of the distal half of the long arm. In good preparations this segment is seen as two bands; the more proximal band divides regions 1 and 2. and the more distal divides 2 and 3. Flourescence of the satellites and short-arm region shows variable intensity and is an inherited variant; a proximal intense band in the long arm has also been observed as an inherited variant.

Chromosome 14 shows a broad medium to intense band in the proximal half of the long arm which divides regions 1 and 2 and a narrow medium and, dividing regions 2 and 3. close to the distal end of the long arm. Variable fluorescence of the satellites and short-arm region has been noted.

Chromosome 15 is the smallest D chromosome and shows the last intense fluorescence of its group. The proximal half of the long arm shows medium florescence which divides region 1 from 2, whereas the distal half is pale. This chromosome is distinguished from No. 14 by the absence of a distal *medium* band, although there may be a distal band of faint fluorescence in No. 15. In poor preparations, it is frequently difficult or impossible to differentiate No. 14 from 15. The satellites and short-arm region show variable fluorescence.

Chromosome 16 is the largest and most metacentric of the E-group chromosomes and is one of the few chromosomes that could be identified solely by morphology. The short arm shows a band (12) of medium fluorescence which is less intense than the central band which separates region 1 from 2 in the long arm. The long arm contains a proximal segment of negative fluorecence (11) which corresponds to the secondary constriction.

The G-band technique reveals the medium bands in the long and short arm and the densely staining secondary constriction. The R-bands are the reverse of Q-bands, except that the region of the secondary constriction is pale. Obanding reveals a large block of constitutive heterochromatin in the region of the secondary constriction which varies in length in different individuals.

Chromosome 17 is the palest staining of the E-group chromosomes. It has a

single distal band (22) of medium fluorescence in the long arm. The dull band proximal to this band divides region 1 from 2.

Chromosome 18 is the smallest E-group chromosome. The long arm contains two bands of medium intensity (12and 22), the proximal band being the brighter and wider of the two. The dull band between these two bands separates regions 1 and 2.

Chromosome 19 is the most weakly fluorescent chromosome in the karyotype and it is difficult to distinguish the long and short arms except in very good preparations. There is a fluorecent spot (12) in each arm adjustment to the centromere; the spot in the short arm is longer and brighter than that in the long arm.

G-bands are similar to Q-bands, and the chromosome shows the same very pale staining except for the centromeric region which is well stained.

Chromosome 20 also show weak fluorescence, but more than No. 19. The short arm is brighter than the long arm. The C-band is medium sized and smaller than in No. 19.

Chromosome 21 is the smaller of the G-group chromosomes and is much brighter than No. 22. The long arm shows a proximal intense segment. Which divides regions 1 and 2, and distal pale segment. There is variable fluorescence of the short-arm region and the satellites; this is an inherited variant. This variability has been useful in determining the source of the meiotic error that results in Down's syndrome (Robinson, 1973).

Chromosome 22 is the larger of the G-group chromosomes and shows very dull fluorescence, similar to No. 19. In fact, in some preparations it may be difficult to distinguish Nos. 19 and 22. A narrow pale band may be observed in the middle of the long arm. A bright fluorescent band in the short arm and variable fluorecence of the satellites are inherited variants.

G-bands are similar to Q-bands, although the centromere region stains darkly with Giemsa. The C-band is medium sized and larger than in No.21.

The *Y chromosome* was the first human chromosome identified with quinacrine fluorescence (Zech, 1969) because of the brilliance of the distal long arm in cells from all but a few males. The variability in the length of the Y chromosome is well established; it is now evident that this polymorphism is correlated with variation in the length of the brilliant segment.

The Y has a variable appearance with G-banding techniques, but may show two distal bands; it is relatively pale when stained with the R-banding technique. The size of the distal C-band in the long arm is directly related to the length of the brilliant segment on fluorescence.

5.2 Human karyotype—banding—nomenclature

5.2.1 Introduction

Metaphase chromosomes show little morphological differentiation in conventional preparations. The size, the position of the centromere, and occasional secondary constriction(s) are the only criteria that can be employed for recognizing chromosomes. In species with high diploid numbers, chromosome pairs with similar morphology become increasingly common, thus making the identification of individual pairs extremely difficult. The human karyotype is comparatively favorable because at least the chromosomes can be classified, according to morphology, into seven groups, and a few pairs can be identified unequivocally. In the karyotype of the laboratory mouse, all chromosomes are acrocentric and do not even allow grouping.

Cytologists have attempted a variety of ways, such as distribution of chemical-i induced breaks, unstained chromosome regions induced by low temperature, and diferential DNA replication time revealed by autoradiography, to further differentiate the chromosomes longitudinally, but all these methods are laborious and the results are ambiguous. The first break through was recorded not long ago when Caspersson I and his collaborators (1969a, b) found that certain fiuorochromes, e.g., quinacrine mustard, when applied to cytological preparations and observed with ultraviolet optics, produced characteristically bright and dark bands. Later Caspersson *et al.* (1970a.b) applied the technique to human chromosome prepsrations and found that the fluorescent banding pattern is likewise specific for each chromosome pair.

Credit to the second major advance must go to Joseph G. Gall and Mary Lou Pardue who perfected the in situ DNA/RNA hybridization technique. In their studies on the cytological locations of the satelite DNA of the laboratory mouse (Pardue and Gall. 1970), they treated the cytological preparations with a series of chemicals in order to achieve molecular hypridization. In these preparations, the centromeric areas of the mouse chromosomes stained more deeply with Giemsa than the Chromosome arms. They regarded the densely stained centromeric areas as heterochromatin. The discovery of a simple staining procedure led to an explosive activity in inventing new procedures, particularly regarding the chromosomes of man and other mammals.

5.2.2 Classification of banding patterns :

The Q-bands. Florescent banding with quinacrine mustrard or quinacrine dihydrochloride.

The C-bands. Constitutive heterochromatin revealed by the PardueGall *in situ* hybridization procedure or its modifications.

The G-bands. Crossbands of chromosomes revealed by a variety of procedures. These bands coincide well with Q-bands, i.e., deeply stained G-bands are brightly fluorescent in Q-band preparations.

The R-bands. The "reverse" banding pattern following the procedure of Dutrillaux and Lejeune (1971).

5.2.3 Cell harvest and slide preparations

Arresting

Harvesting cells is strictly conventional. Bone marrow cell culture or any cell population containing a high incidence of mitosis are suitable. Agents such as colchicine, colcemid, and vinca alkaloid can be used to accumulate mitosis. However, over condensed chromosomes yield very poor banding, so that prolonged mitotic arrest should not be done.

Fixation

The cell polulations should be treated with a hypotonic solution prior to fixation. It matters little which kind of hypotonic solution is used. The cells, after hypotonic solution treatment, are fixed according to the type of slide preparations to be made, viz,, "Carnoy" fixative (1 glacial acetic acid: 2 methanol) for air-dired slides. For cell populations which require squash technique (e.g., many solid tissues), the fixative to use is 45—50% acetic acid. However, the preparations are not suitable for G-bands, though they are excellent or C-bands.

It is suggested that the air-dried slides be incubated at 37°C for 1 hour (without covering) and thereafter be kept in air-tight slide boxes containing a drying agent such as silica gel. The slides may be used immediately after this incubation period or may be stored for a few weeks.

5.2.4 C-Banding

In good C-band preparations, the constitutive heterochromatin should stain deeply and the euchromatin should show only a faint outline of the chromosome. However, flame-dried preparations, when improperly treated, will show G-bands as well as C-bands, which is confusing.

The original procedure (Pardue and Gall, 1970) and the modifications thereof (Arrighi and Hsu, 1971; Yunis et al., 1971) are useful for distinguishing constitutive heterochromatin and euchromatin in mammalian chromosomes. The procedure to be described is a simplified version designed to reveal C-bands using air-dried and flame-dried preparations, although some comments will also be made for squash preparations.

A. Reagents

1. HCL : prepare 0.2 N solution.

2. NaOH : Prepare a 0.07 N solution.
3. SSC : Prepare a lOx concentrate (a solution of 0.15 M sodium citrate and 1.5 M NaCL) and dilute with distilled water to the desired concentrration.
4. Giemsa staining solution.
5. Phosphate buffer solution (0.01 M Sorensen's phosphate buffer, pH 7.0).

B.Procedure

1. Treat the slides with HCL at room temperature for 15 minutes. Rinse with distilled water three times.
2. Treate the slides with NaOH for 2 minutes. Rinse with 70%, then 95% ethanol three times for a period of 5 minutes each. Air dry the slides.
3. Place slide horizontally, with cell-side up, in a moist chamber, and add either 2x or 6x SSC to the cell area of the slide. Place a coverglass over the SSC solution.
4. Incubate the moist chambers containing the slides at 60°—65°C for 16-20 hours.
5. Rinse in either 2x or 6x SSC (three times, 5 minutes each), 70% ethanol (three times, 5 minutes each), 95% ethanol (three times, 5 minutes each), and air dry.
6. Stain in Giemsa solution.

C. Comments

1. After a considerable amount of experimentation, it is generally opined that the HCL treatment is an important step in eliminating the G-bands in C-band preparation, particularly when air-dried and flame-dried slides are used. In sqush preparations, HCL treatment is not a vital step and may be omitted for C-band preparations.

2. The concentration of NaOH and the duration of the NaOH treatment are important. As a standard, one may start with 0.07 N for 2 minutes. This combination may be too strong and the resulting euchromatic chromosomes may appear bloated and show an empty appearance. If such a result is obtained, one must experiment with reduced concentration of NaOH solution (0.02 N, 0.01 TV) and time of treatment (1 minute, 30 seconds, or even 15 seconds). If the chromosomes still appear distorted, one should then try the solution suggested by Stefos and Arrighi (1970): a 2 x SSC solution with pH adjusted to 12 by NaOH. This solution is particularly useful for small chromosomes such as the microchromosomes of the birds and the chromosomes of *Drosophila* (Hsu, 1971). Conversely, some C-bands require a prolonged NaOH treatment.

3. Many laboratories use Coplin jars filled with 2 x or 6 x SSC for the over night incubation. This is undersirable because the glass slides will stain heavily with Giemsa, thus interfering with the observations on the chromosome banding.

Incubating slides in moist chambers eliminates this defect. If, however, such defect is not observed, Coplin Jars are of course convenient.

A simple moist chamber can be constructed as follows. Use a Petridish of suitable size. If square (120 mm each side) Petri dishes are used, 10 ml of either 2 x or 6 x SSC is placed in the bottom (15 ml for overnight treatment at 65°C). Next a stand for the slides is placed in the bottom half. The stand should be as small as posible and of sufficient height so that the slide is above the salt solution in the bottom half of the Petri dish. The slide(s) is placed on the stand. A few drops of the solution are placed on the slide to cover the cells. A coverglass is then placed over the solution, the Petri dish is covered, and the entire chamber is placed in an oven set at the desired temperature.

4. The stock Giemsa-staining solution is diluted with phosphate buffer, and the concentration varies with each new lot of stain. Usually concentrations varying from 2 to 10% have been used and stained the slides from 5 to 30 minutes.

5. Slides should be of good quality and should be cleaned in some manner. Slides can be cleaned in 95% ethanol, soap, and dilute HCL. AH seemed to be acceptable. Coverglasses should also be cleaned.

D. Squash preparation

This section is for squash preparations only. If squash preparations are used for C-bands, the slides should be dipped into a solution of 0.1 % gelatin and 0.01 % chrome alum and dried prior to squashing. This thin coat prevents cellular loss during the treatments. However, if the HCL treatment is omitted, the slides should be treated with RNase (100 pg/ml. diluted in 2 x SSC) at 37°C for 1 hour using the moist chamber method. Rinse the slides for 5 minutes each in three changes of 2 x SSC, 70% ethanol, and 95% ethanol and air dry. Treat the slides with NaOH solution (0.07 N NaOH or 2 x SSC, pH 12). Try several treatment times, e.g., 1 minute, 2 minutes, 4 minutes, etc. if 0.07 //NaOH is used, rinse in 70% ethanol, 95% ethanol, dry,and incubate in 6 x SSC as usual. If 2 x SSC at pH 12 is used, rinse slides in three changes of 2 x SSC for 10 minutes each. Do not dry but drain and immediately place the slides at 65°C in a moist chamber, as suggested earlier.

5.2.5 G-Banding

G-bands are the crossbands of various width and shades stained with Giemsa, Leishman's, Writht's, or similar stains. They usually correspond to the Q-bands but do not always correespond to the C-bands. In some cases, the C-bands and the G-bands may be opposite in staining behavior. For example, the C-band of human chromosome 9 is relatively unstained with G-band techniques.

The Y chromosome of man shows a distinct C-band in the distal portion of the long arm, but the same chromosome is somewhat variable in G-band staining, usually deeply stained throughout. Thus, G-banding does nor replace C-banding in assessing information.

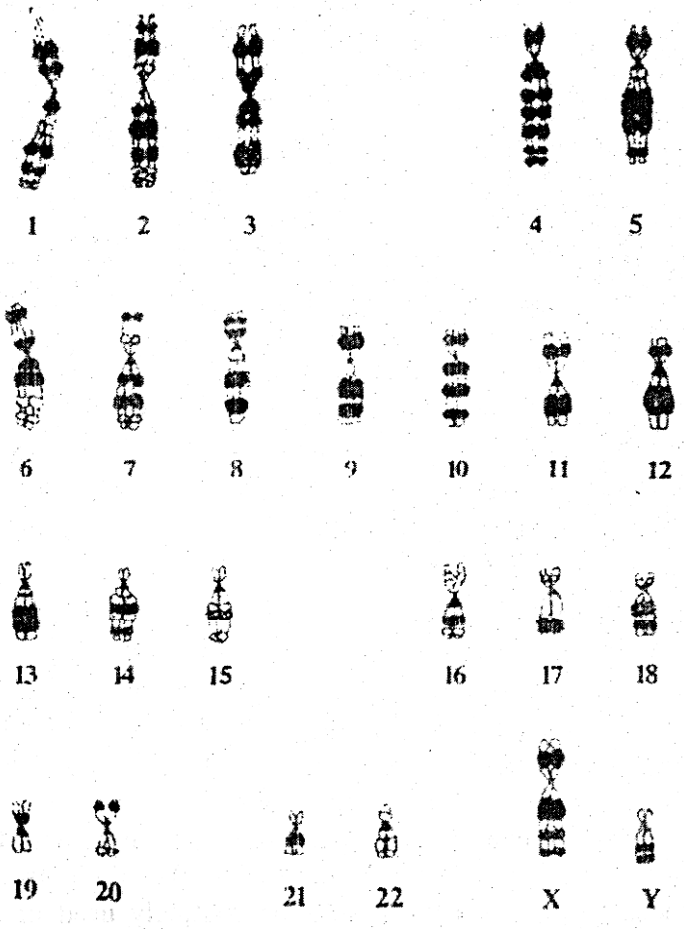


Fig. 5.6 Kayotype of human male showing C-banding pattern from a flame-dried preparation. Note the large amount of heterochromatin at the centromere areas of Al, C9, E16, and the distal portion of the Y. Variations in amounts exist in the two homologs of Al, Courtesy of Mrs.Ann Craig-Holmes and Dr. M. W. Shaw

There were many papers published in 1971 and 1972, each proposing a certain procedure to reveal crossbands in mammalian chromosomes (Summer et al., 1971; Drets and Shaw, 1971; Patil et al, 1971; Schned, 1971a,b; Seabright, 1971; Wang and Fedoroff, 1972; Kato and Yosida, 1972; Utakoji, 1972). Since the results of the various procedures are similar to one another, it is unnecessary to present the procedures for all of them. The trypsin procedure described here more or

less follows the one devised by Seabright (1971) with recommendations for individual laboratory modifications.

A. Reagents

1. Trypsin solution : Seabright uses Bacto trypsin (Difco Catalogue No. 0153) prepared by adding 10 ml of sterile distilled water or isotonic saline to each vial as the stock solution. This stock solution is diluted 1 : 10 with saline before use.

It is really not necessary to use the particular brand of trypsin recommended by Seabright. Most laboratories carry monolayer cell cultures which require trypsin to dislodge the cells for harveest or for subculturing. Usually it is a crude trypsin solution (0.20—0.25% dissolved in a balanced salt solution without Ca^{2+} and Mg^{2+}). In some laboratories, purified trypsin solution (0.01—-0.02%) is used. Whatever the kind and the concentration, the trypsin solution routinely used in the cell culture laboratory can be considered as the "stock solution". The trypsin solution used for G-bands is prepared by diluting the stock trypsin solution with saline, balanced salt solution, or "rinsing solution" (balanced salt solution without Ca^{2+} and Mg^{2+}). In our laboratory we use the rinsing solution.

For laboratories using trypsin solution for G-banding only, it is advisable to dispense the stock trypsin solution in small containers and store them in a freezer. Keep only a small amount in the refrigerator for immediate use.

2. Rinsing solution : Physiological saline or balanced salt solution without Ca2~ and Mg2'. This solution is used diluent of the stock trypsin solution as well as for rinsing the slides after the trypsin treatment.

3. **95% Ethanol**
4. **Giemsa staining solution :** See C-bands above.
5. **Phosphate buffer :** Like that of C-bans.

B. General principles

Many factors may influence the success of the G-band staining by trypsin treatment. It is, therefore, pointless to follow a set recipe without knowing these factors because the preparations may give excellent results if one knows how to modify the procedure. The success of G-banding depends primarily on the combination of the concentration of the trypsin solution and the duration of treatment, but the following factors dictate the correct combination :

1. **The method for preparing the slides :** The flame-dried preparations are more resistant to the trypsin treatment than air-dired preparations.
2. **The age of the slides :** The longer the slides are stored, the more resistant the cells are to the treatment. Cells of very old slides often give spotty, instead

of banded, chromosomes. Refixing of the slides in the Carnoy fixative sometimes helps.

3. **Heating the slides :** Air-dried slides be heated at 37°C for 1 hour (without covering). This procedure seems to give more consistent results.
4. **The salt composition of the trypsin solution (including the diluent) :** The presence of divalent cations in the solution slow the reaction but do not prevent it.
5. **The temperature of the trypsin solution (the higher the temperature, the faster the reaction) :** In laboratories with air conditioning, room temperature is suitable. The trypsin solution should be stabized at room temperature for approximately 30 minutes prior to use. For laboratories without room temperature control, it is probably a good practice to stabilize the trypsin solution at 4°C (refrigerator) or even in an ice bucket (Deaven and Petersen, 1973).
6. **Trypsin concentration :** We suggest a dilution of 1:5 or 1:10. If these are too strong, dilute further. If too weak (as in the case of flamelried preparations), the concentration may be raised to 1:2 or the undiluted trypsin used.
7. **Time of treatment :** The time of treatment of course depends on all the factors mentioned above. As a general principle, it should be adjusted to give good results in not more than 2 minutes but not less than 30 seconds.

C. Procedure

1. Prepare trypsin solution in a Coplin jar. Using a Coplin jar is somewhat more convenient than flooding the slides with the solution but either way is acceptable. Use two or three slides and vary the duration of trypsin treatments as the initial monitor. Since for best results monitoring is necessary, it is advisable to prepare at least 10 to 12 slides of good quality.
2. Rinse with physiological saline or rinsing solution. Seabright suggested, at this stage, inspection of the wet slides by phase-contrast microscopy to determine the effect of trypsin treatment. The chromosomes should appear slightly swollen. The preparations can be treated again with trypsin if necessary.
3. Rinse with 95% ethanol and let dry.
4. Stain with diluted Giemsa (2% Giemsa solution in phosphate buffer) for 4— 10 minutes. Overstaining may obliterate some of the lighter bands.
5. The slides can be pulled out of Giemsa, rinsed quickly with deionized water, and air dried. It is not necessary to mount the slides.

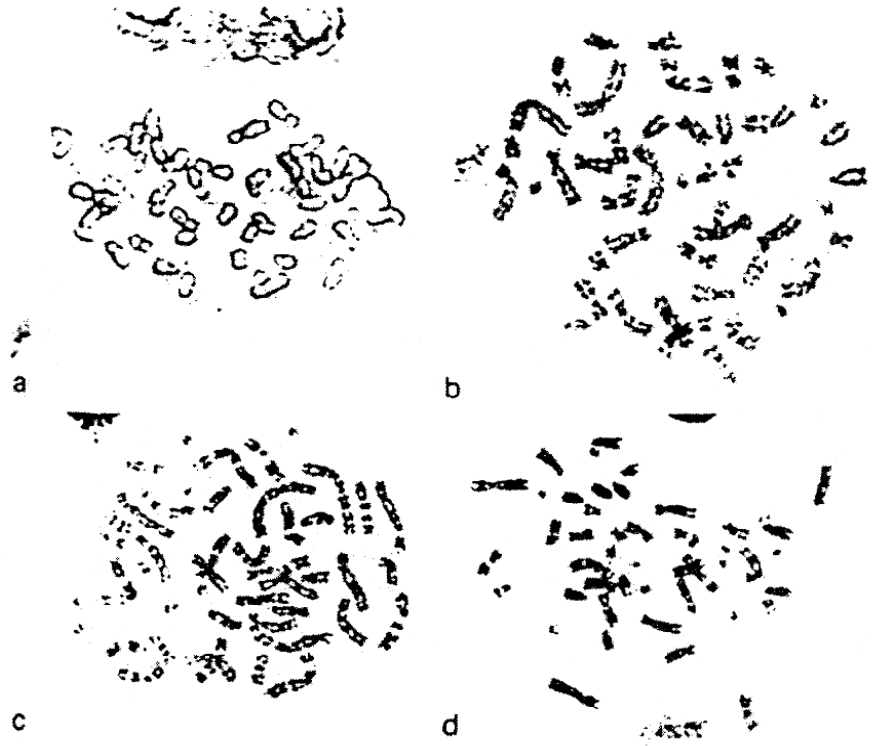


Fig. 5.7 Metaphase plates of human cells showing results of various durations of trypsin treatment from extreme over treatment to proper timing, (a) Extreme overtreatment; (b) overtreatment, highly unsatisfactory; (c) slight overtreatment; chromosomes are fuzzy but discreate bands can be seen; (d) proper treatment

Examine the stained preparations to determine the proper duration of the trypsin treatment. The chromosomes in undertreated preparations, in overtreated preparations, the chromosomes will show a series of appearances ranging from completely "ghost" chromosomes (Fig 5.7a) to those with poorly differentiated crossbands and fuzzy outlines (Fig. 5.7b). The appearance of these cells indicates that the treatment time or the concentration of trypsin solution should be reduced. Figure 5.7c shows a metaphase with reasonably good but slightly overtreated chromosomes, and Fig 5.7d shows proper G-bands.

6. Once the proper combination of trypsin concentration and the duration of treatment is determined by the preliminary monitor, treat the rest of the slides according to the best combination in the same day using the same solutions. The solutions in the Coplin jars should be discarded each day.

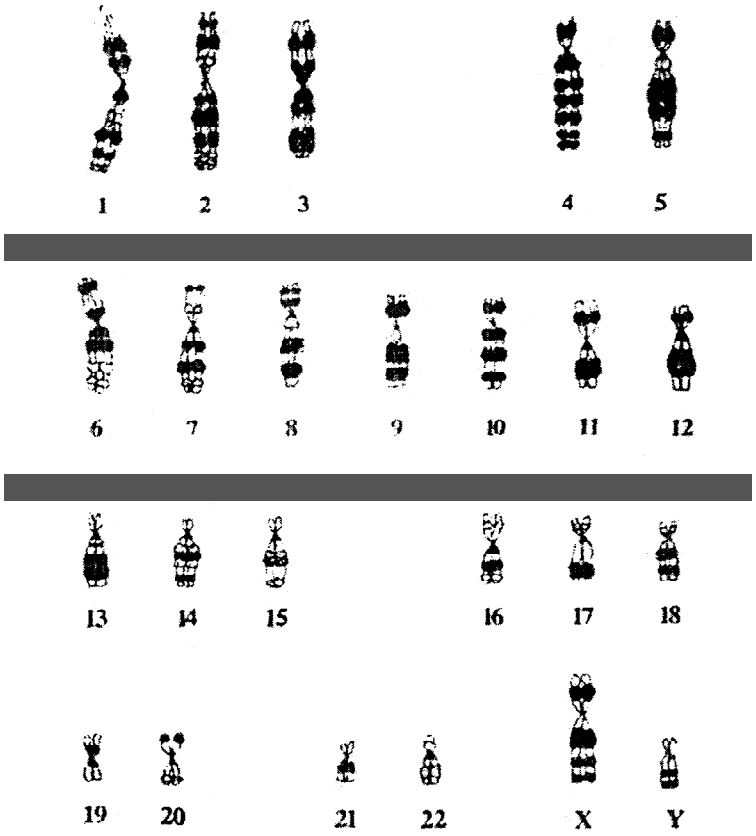


Fig. 5.8 Karyotype of human male showing G-banding pattern produced by the trypsin technique. Courtesy of Mrs. Marina Seabright

7. Figure 5.8 presents a male human karyotype showing the distribution of G-bands.

D. Comments

1. If the trypsin procedure does not give satisfactory G-banding, it may be worthwhile to try the urea procedure of Kato and Yosida (1972). Prepare a stock urea solution (*8 M* aqueous) and mix 3 parts of this stock solution with 1 part of Sorensen's phosphate buffer (0.15 M, pH 6.8), heat to 37UC. Treat the slides at 37°C for 10 minutes. Rinse in tap water and stain directly without drying. For old slides, elevate the temperature to 60°C.

2. For combination staining, e.g., Q-banding and C-banding of the same cell, it is advisable to perform Q-banding first. The same slide can be then used for C-banding. It is recommend, however, separate runs for C-banding and G-banding. From each sample, one can always prepare a sufficient number of slides for slides for al treatments as well as some for storage.

5.3 Numerical and structural abnormalities of human chromosomes—syndromes.

5.3.1 Introduction

Fifty years has elapsed since Tjio and Levan in 1956 established that human somatic cells have 46 chromosomes. There technique employing treatment of cells grown in culture with colchicine and hypotonic solution, was applied by numerous workers over the next half-dozen years, demonstrating the chromosomal findings in such disorders as trisomy 21, trisomy 13, trisomy 18, Turner's syndrome, and Klinefelter's syndrome. During this same period, the significance of the Barr body was realized, and lyonization of the X chromosome was hypothesized.

The next few years saw the application of autoradiography for better identification of chromosomes, the delineation of more subtle chromosome disorders such as the cri-du-chat (5p—) syndrome, Wolf-Hirschhorn (4p—) syndrome, 18p—, 18q—, 13q—and a host of other anomalies. Amniocentesis was employed for prenatal diagnosis.

Following 6 years have been marked by advances in our understanding of heterochromatin, chromosome banding, the identification of the Y body, and documentation of previously unrecognized alterations (small translocations, inapparent inversions etc.).

From the first conference held in Denver in 1958 to the one, convened in Paris in 1971, chromosome nomenclature has developed parri passu with advances in our understanding of these disorders.

Gross human chromosome abnormalities are not rare. Over 25% of human abortuses lost before the eighth week of pregnancy have abnormal karyotypes. Large surveys on newborns have consistently shown that about 0.5% exhibited aneuploidy (Sergovich et al., 1969; Lubs and Ruddle, 1970).

5.3.2 Autosomal abnormalities of human chromosome—syndromes

A. 4p— Syndrome (Wolf-Hirschhorn Syndrome)

The syndrome described independently by Wolf *et al.* (1965) and Hirschhorn *et al.* (1965), results from partial deletion of the short arm of one of the late replicating no. 4 chromosomes. It is much less common than the 5p— syndrome. About 25 cases have been described to date, all sporadic Translocation has not been demonstrated with the possible exception of the case described by Wilson *et al.* (1970). Parental age has been somewhat increased (Fryns et al., 1973). A 4r syndrome has been described (Carter *et al.* 1969).

The disorder is characterized by severe psychomotor and growth

retardation. Birth weight is usually about 2000 gm in spite of normal gestation time. Fetal activity is diminished. Most infants are hypotonic.

The skull is microcephalic and often there is cranial asymmetry. In a few cases, midline scalp defects have been noted (Hirschhorn et at., 1965; Wolf *et al,* 1965 Miller *et al,* 1970). Hemangioma on the brow is frequent. A prominent glabella and ocular hypertelorism are almost constant features. Divergent strabismus, antimongoloid obliquity of the palpebral fissures have been noted in about half the cases. Iris coloboma has been found occasionally.

The ears have narrow external canals and are low set and simplified in from. The nose is misshapen or beaked with a broad base. The philtrum is short with a down-turned Cleft lip or, especially, cleft palate and micregnathia have been noted in most cases.

Males comonly exhibt cryptorchidism and, especially, hypospadias. Absent uterus and streak gonad have been described. Congenital heart malformations, most often atria! or ventricular septal defects, have been noted in about 50% (Wolf *et al.*, 1965, Arias *et al.,* 1970, Guthrie *et al.,* 1971) and may result in death during the first year.

In several patients there has been dimpling of the skin over the sacrum and elsewhere, such as shoulders, elbows, or knuckles. The pelvic and carpal bones are late in ossification. Pseudoepiphyses are seen in the phalanges and at the base of each metacarpal.

B. 5p—Syndrome (Cri-du-Chat Syndrome)

Described initially by Lejeune *et al.* (1963), over 150 examples have been documented to date. The syndrome is present in about 1% of institutionalized individuals with intelligence quotients less than 35. The syndrome results from deletion of 35—55% of the short arm of one of the early replicatig B group chromosomes (German *et al,* 1964; Miller et al, 1969). Most decletion are thought to occur as a result of two breaks. If these occur in the short arm, an interstital deletion results. If deletion occurs in both arms, a ring chromosome is produced (Rohde and Tompkins. 1965). Maternal age is not elevated. About 70% of those identified at birth are females: however, most older patients have been male (Breg *et al,* 1970). The reason for this discrepancy is not evident. Mosaicism has also been described; patients having all the stigmata of the full-blown syndrome (Zellweger, 1966; Mennicken *et al,* 1968). About 10—15% result from translocation (Warburton and Miller, 1967). Pericentric inversion has also been described (Faed *et al,* 1972).

As the name implies, the syndrome is characterized by a catlike, weak, shrill cry in infancy caused by hypoplasia of the larynx (Ward *et al.*, 1968). However, the cry usually disappears with time, even within a few weeks of age

(Gordon and Cooke, 1968; Breg *et al,* 1970). The cry, almost one octave higher than normal, is quite monotone in quality (Schroeder *et al,* 1967).

The infant face is characterized by microcephaly, round form, hypetelorism, antimonogoloid obliquity of. palpebral fissures, epicanthus, bilateral alternating strabismus, broad nasal bones, and low-set ears. Most patients have mild micrognathia. However, the rouness of the face and the ocular hypertelorism disappear with age. The face becomes thin and the philtrum short. Permature graying of the hair has been noted in about 30%. Dental malocclusion is common (Breg *et al,* 1970, Gordon and Cooke, Niebuhr, 1971).

There is usually severe mental retardation (I. Less than 25), failure to thrive, and hypotonia in infancy. Birth weight is usually less than 2500 gm in spite of normal gestation time. Adult height usually ranges from 124 to 168 cm. (49 to 66 inches) Various musculoskeletal anomalies have included hypotonia, flat feet, mild scoliosis, large frontal sinuses, small ilia, syndactyly, and short metacarpals and metatarsals (Mennicken *et al,* 1968).

Dermatoglyphic alterations include simian creases in about 35%. Eight or more whorls have been noted in about 30%.

C. Group C deletion, trisomy, trisomy mosaicism, and partial trisomy

There are seven pairs of C-group chromosomes and hence many possible types of trisomy, partial trisomy, or deletion involving this group. These states, with few exceptions, have not been clinically recognized and, thus, are probably lethal. The use of newer banding techniques on the chromosomes of abortuses may shed light on this question (Hirschhorn *et al,* 1973).

There is a paucity of information concerning anomalies of chromosome no. 6 documented by banding, de Grouchy *et al.* (1968) described a child with bulbous nose, preauricular tubercle, hernias, hypospadias, undescended tested, deep acromial dimples, and psychomotor retardation.

Deletion of the long arms of chromosome no, 7 was reported by Shokeir *et al.* (1973). The child exhibited psychic and somatic retardation, urinary malformations, flexion contractures at the elbows, and low finger ridge count. The most striking aspects of trisomy 8 or trisomy 8 mosaicism syndrome are mental retardation, abnormally shaped skull, reduced joint mobility, various vertebral anomalies, supernumerary ribs, strabismus, absent patellae, short neck, long slender trunk, cleft palate, and marked palmar and plantar creases (Oikawa *et al.,* 1969; Lejeune *et al,* 1969; Emberger et al, 1970; Riccardi *et al*.)

Deletion of the short arms of chromosome no. 9 was noted by Alfi *et al.* (1973) and probably by Kistenmacher and Punnett (1970). The patients had trigonocepha, mental retardation, ocular hypertelorism, anteverted nostrils, malformed pinnas. long philtrum, short neck, hypertonia, congenital heart disease,

and an incrreased number of digital whorls.

Trisomy for the short arm of no. 9 was defined by Rethore *et al.*, (1973) who reviewed earlier cases. Clinical features included mental retardation, microcephaly, enophthalmos, hypertelorism, mongoloid palpebral fissures, bulbous nose, abnormal pinna, hypoplasia of the phalanges, and abnormal finger creases.

Partial trisomy for the long arm of chromosome no. 10 was reported by de Grouchy *et al*., (1972). Facial dysmorphia was evident with microcephaly, large forehead, flat round face, arched and wideset eyebrows. Antimongoloid palpebral fissures, microphthalmia, cleft palate, small nose with depressed bridge, malformed pinnas, short neck, micrognathia, various skeletal anomalies (osteoporosis, various rib abnormalities, scoliosis), congenital heart disease, and genitourinary defects. All patients had severe mental retardation.

Partial trisomy for the short arm of chromosome no. 11 was described by Falk *et al*. (1973) and Sanchez *et al*. (1974). Findings common to both cases were mental retardation, marked frontal bossing, nystagmus, antimongoloid palpebral fissures, strabismus, broad fingers or toes, and cleft lip and/or palate.

D. Trisomy 13 syndrome (Patau's syndrome, trisomy D_1)

Trisomy 13 was first recognized by Patau *et al.* (1960), although Bartholin in 1657 may have given the first description of the clinical features (Warburg, 1960). The phenotype is so striking that diagnosis is usually made on clinical ground before the karyotype has been made. The incidence has been estimated to be about 1 per 6000 births (Conen and Erkrnan, 1966).

Arhinencephaly, apneic spells, seizures, feeding difficulties, severe mental retardation, and deafness are common. Any of the holoprosencephalic states (cyclopia, ethmocephaly, cebocephaly, and premaxillary agenesis) may be associated with trisomy 13 (Conen *et al.*, 1966; Fujmoto *et al.*, 1973). Moderate microcephaly with sloping forehead and wide sagittal suture and fontanels have been noted in over 60%.

Microphthalmia or iris coloboma with retinal dysplasia, ocular hypertelorism, and malformed pinnas occur in about 80% (Cogan and Kuwabara, 1964). Capillary hemangiomas in the glabellar region and localized scalp defects in the parieto-occipital area have been described in about 75%. Cleft lip and/or cleft palate and micrognathia have been noted in 60-70% (Conen *et al.*, 1966; Taylor, 1968).

Musculoskeletal abnormalities include postaxial polydactyly of the hands or feet with overlapping flexed fingers (about 75%) with hyper convex narrow fingernails. The calcaneus is often prominent and frequently there are rockerbottom feet.

At least 80% have congenital heart defects, genital anomalies include cryptorchidism (over 90%) in males and bicornuate uterus (about 50%) and hypoplastic ovaries in females.

Polymorphonuclear neutrophils frequently (25—80%) have nuclear projections in cases of trisomy 13 owing to primary nondisjunction. Excellent ultrastructural study of the projections has been carried out (Waltzer *et al,* 1966, Lutzner and Hecht, 1966). Fetal hemoglobin, Hb-Gower and other hemoglobins have been elevated but there is good evidence that those changes disappear with age and merely represent general delayed maturity (Marden and Yunis, 1967).

DNA replication studies have demonstrated that the D-group chromosome involved is number 13, which is the longest and the latest of the pairs to replicate (Yunis and Hook, 1966).

1. Trisomy 13 Caused by Primary Nondisjunction

About 75% of cases of 13 trisomy are caused by primary nondisjunction. There is no sex predilection. The mean age for mothers of infants with 13 trisomy caused by this type is elevated (32.4 years), far higher than for cases caused by translocation or mosaicism (Magenis *et al.*, 1968; Taylor *et al.*, 1970).

There have been several examples of 13 trisomy occurring with other chromosomal abnormalities in the same sibship (Klinefelter's syndrome, Turner's syndrome, Down's syndrome and triploidy), but this may be chance association (Visfeldt, 1969).

2. *Translocation D_1*

About 20% of the cases of trisomy 13 are caused by translocation, far more common than occurs in Down's syndrome (Magenis *et al.*, 1968; Taylor *et al.*, 1970). In at least 85%, the translocation has occurred between two D chromosomes. Maternal age is not elevated (25.6 years). There appears to be definite male predilection. Fertility and intelligence in balanced carriers are quite variable (Wilson, 1971).

3. *Mosaicism*

About 5% of the cases of Dj trisomy are caused by mosaicism. About half of these examples are caused by an extra chromosome 13 in proportion of the cells. The remainder result from a complex assortment of chromosomal abnormalities (Magenis *et al.*, 1968; Taylor *et al.*, 1970).

As in translocation D_1 trisomy, the age of the mother of a D_1 trisomy mosaic is not elevated (25.4 years) in contrast to mothers of G_1 trisomy mosaics. The clinical stigmata, as expected, are less severe than in those of children with classic trisomy 13 (Bain *et al.*, 1965).

4. Partial Trisomy

Partial trisomy for the distal segment of the long arm of chromosome 13 was documented by banding techniques by Taysi *et al.* (1973) and by Escobar *et al.* (1974). The latter authors reviewed several case reports which had been documented prior to the advent of banding. Common clinical characteristics included psychomotor retardation, seizures, microcephaly, *frontal bossing,* open anterior fontanel, short neck, inguinal and umbilical hernias, polydactyly, rocker bottom feet, distal axial triradius, and elevated fetal hemoglobin. Life expectancy over a year was frequent. Absent were cleft lip and palate, sloping forehead, microphthalmia, and neutrophil drumsticks, common findings in trisomy 13.

E. Dq—and Dr syndromes

Over 60 case reports have been published in which the patient had deficiency of part of the long arm of a D-group chromosome (Dq—) or in which a D-group chromosome was replaced by a ring (Dr) (Lejeune *et al.*, 1968; Gilgenkrantz *et al.*, 1971; Niebuhr and Ottosen, 1973). Although these cases may represent a heterogeneity, there is good evidence to suggest that most involve No. 13 (wilson *et al.*, 1973).

Only a few examples of Dr have been described in which the chromosome has been identified as no. 15 (Jacobsen, 1966; Emberger *et al.*, 1971). The phenotype in these cases was not striking : short stature, mental retardation, and microcephaly. Mean survival has been 39 months for Dq— cases and 89 months for Dr examples (Taylor, 1970).

All patients have exhibited mental and somatic retardation and many have been hypotonic.

Musculoskeletal abnormalities have included bilateral hip dislocation, focal lumbar vartebral agenesis, inguinal hernia, coxa valga, and synostosis of the fourth and fifth metacarpals.

F. Trisomy 18 syndrome (Edwards syndrome)

In 1960, Edwards *et al.* and, almost simultaneously, Patau *et al.* 1961) described a new syndrome associated with the presence of an extra chromosome in the E group which was subsequently shown to be a no. 18 chromosome (Yunis *et al,* 1964).

The most constant features of this syndrome, noted in over 75% of the cases, include : deveelopmental retardation, failure to thrive, feeding difficulties, hypertonia, limited hip abduction, flexion deformities (usually ulnar deviation) of fingers, short sternum, congenital heart disease (ventricular septal defect—90%, patent ductus arteriosus—70%, and atrial septal defect—20%), short

dorsiflexed halluces, rockerbottom feet, calcaneovalgus deformity of feet, and cryptorchidism (Weber and Sparkes, 1970).

Craniofacial anomalies almost always present include prominent occiput, low-set malformed pinnas, and micrognathia. Severe anomalies found at autopsy, apart from the cardiac anomalies noted above include Meckel's diverticulum, heterotopic pancreatic tissue, thin diaphragm with eventration, and various renal anomalies.

Dermatoglyphic alterations are frequent. Over 85% of finger prints are simple arches. Over 30% have a simian palmar crease and over 40% have a single flexion crease in the fifth finger.

Trisomy 18 has an uncommon but yet definite association with aplasia of the radius and thrombocytopenia.

1. Trisommy 18

The incidence of trisomy 18 in trhe more recent surves has varied from 1 per 3500 to 1 per 7000 births (Taylor, 1968, Benady and Harris, 1969; Garfinkel and Porter, 1971). Mean maternal age is elevated, 32 years (Taylor, 1968).

There is a 3 : 1 female predilection caused, in large part, by a greater male fatality rate during the first few weeks of life (Weber, 1967).

The mother often exhibits small weight gain during pregnancy and indicates that fetal movements were feeble. Most examples are postmature. Mean birth witht is less than 2300 gm. The placenta is often small with umbilical artery, and hydramnios has been noted in over 50%.

Thrity percent fail to survive more than 1 month, 50% succemb by 2 months, and less than 10% live more than 1 year. Mean survival time is about 70 days (females—134 days, males—15 days).

2. Double Trisomies

Double primary nondisjunction has been observed in 5—10% of cases (Hamerton, 1971). Mean survival time for douuble trisomies has been 3 weeks. Maternal age is markedly increased in this group.

3. Trisomy 18 Caused by Translocation

Translocation is usually sporadic but examples of familial trnslocation have been recorded (Hamerton, 1971). Means maternal age is lower than for those with trisomy 18 caused by nndisjunction.

G. 18p—syndrome

Deletion of the short arms of chromosome 18 is associated with a variable phenotype. Maternal age is elevated. There is a 2 : 1 female sex predilection (Parker *et al,* 1973).

Mental retardation is a constant feature but of variable degree. Birth weight is low and somatic growth retarded. There is no characteristic facial dysmorphia. Frequently, however, hypertelorism, epicanthal folds, strabismus, and ptosis of lids are noted. The ears are low-set, large, floppy, and poorly formed.

H. Trisomy 21 (Down's syndrome)

Langdon Down (1866) first extensively described the syndrome which has received his name, calling it "Mongoloid idiocy" or "Mongolism." In 1959. Lejeune demonstrated that the disorder was associated with an extra chromosome in the G group. In 1960, Polani et al. described translocation Down's syndrome, and in 1961, clarke et al discovered rnosaicism for an extra G-group chromosome. Yunis et al (1965), by means of autoradiography, identified the chromosome as one of the no. 22 chromosomes, although by this time the term trisomy 21 had been so extensively employed that it has remained.

The incidence of trisomy 21 is between 1 and 2 per 1000 live births among various populations (Mikkelsen, 1971). Over 95% of the cases are caused by nondisjunction, the remainder resulting from translocation.

The skull is brachycephalic with shortening of the anteroposterior diameter and flattening of the occiput in about 75%. The cephalic index (normally 0.75—0.80) is usually greater than 0.80 and may exceed 1.00. (Roche et al., 1961). In infants with trisomy 21, the fontanels are larger than normal and closure is late. In those over 10 years of age, a patent metopic suture is found in 65% of males (normal—9%) and in 40% of females (normal—12%). An extremely common feature (over—90%) is absence of frontal and spehenoid sinuses and hypoplasia of the maxillary sinuses (Spitzer *et al.*, 1961, Betlejewski *et al.* 1964). There is poor development of the bones of the middle face, producing a relative prognathism and ocular hypotelorism (Gerald and Silverman, 1965).

The profile is flattened owing to hypoplasia of the nasal bones. The palpebral fissures are oblique, the outer canthus being slightly higher than the inner. Epicanthal folds are extremely common. Speckled iris (Brushfield's spots) and lens opacity are present in about 85% and 60% of patients, respectively.

Various other anomalies include missing or malformed teeth (especially maxillary lateral incisors and mandibular second premolars), delayed eruption, increased periodontal destruction, and malocclusion (Cohen and Cohen, 1971).

The hands are characteristically short and broad, the fifth finger usualy being abbreviated and clinodactylous, and having a single flexion crease in about 20% of the cases. There is usually greater space than normal between the hallux and the rest of the toes.

Hypotonia, especially marked in infancy, improves with age. loints are usually hyperextensible. The penis and scrotum are usually small and about

25% have cryptorchidism Public hair is straight. Congenital cardiac anomalies is present in about 40% Down Syndromic persons.

Diastasis recti, duodenal atresia, or umbilical hernia occur in about 10% (Butterworth *et al.*, 1964).

Radiographic changes include reduced iliac and acetabular angles in the young infant (Nicolis and Sacchetti, 1963) and hypoplastic middle phalanx of the fifth finger.

Intelligence quotients range from 25 to 70, most Down's syndrome patients 3 years of age or less having I.Q. s of 50—59 but slipping with increasing age to 25-49 (Penrose and Smith, 1966).

Dermatoglyphic anomalies include distal axial triradius in the palm (over 80%), bilateral simian creases (30%), single flexion crease in fifth finger (20%), 10 ulnar loops (30%), and hallucal arch tibial (70%) or small loop distal (30%) patterns (Preus and Fraser, 1972).

Because of susceptibility to respiratory infection, early mortality used to be great. With the introduction of antibiotics, the means survival age is almost 20 years. There is a twentyfold increased association with acute leukemia (Conand Erkman. 1966).

Numerous attempts have been made to established specific biochemical alterations. However et aL, (1965) found decreased blood serotonin and increased galactose phosphate uridyltransferance leukocyte alkaline with Down's syndrome.

1. Trisomy G_1 due to Primary Nondisjunction

As discussed above, about 95% of cases of Down's syndrome are sporadic primary trisomics, resulting from nondisjunction which is age dependent. This occurs at the first meiotic division in the mother (Robinson, 1973). If the mother is less than 20 years of age at time of conception, the risk of producing a child with trisomy 21 is about 1 per 2500 live births. This risk gradually increases until 35 years, after which there is a more marked increase in frequency such that a mother over 45 years has about 1 chance in 50 or less of having a child with Down's syndrome.

2. Association of Down's Syndrome with Other Primaiy Nondisjunctions

Individuals with trisomy 21 have been occasionally (about 1 per 200) found to have another extra chromosome (double primary nondisjunction), the most frequent type being 48,XXY, G+ (Hamerton *et al,* 1965; Taylor and Moores, 1967). Other forms such as 48,XXX,G+ and 48,XYY,G+ have also been described (Yunis *et al,* 1964, Uchida *et al.*, 1966). This association is much higher than might be expected by chance.

3. Translocation Down's Syndrome

Down's syndrome patients born to young mothers as well as those with affected relatives often have the extra G] chromosome attached to another chromosome. This has been designated translocation and comprises about 3.5% of cases of Down's syndrome. It may be sporadic or familial. Translocation Down's syndrome is not age dependent. About 8% of Down's syndrome patients born to mothers less than 30 years of age have exhibited translocations as opposed to 1.5% born to mothers over 30 years old. It is widely accepted that the short arms of acrocentric chromosomes have nucleolar organizers and that these points are likely to break, producing a high frequency of structural chromosome aberrations.

In familial translocation Down's syndrome, one of the parents has 45 chromosomes instead of the normal 46. One of the small G-group chromosomes is "missing" since it has been translocated to another chromosome. The parent carrying the translocation chromosome is phenotypically normal, since no significant amount of genetic material has been lost in the translocation process. In most cases, the extra G(chromosome is phenotypically normal, since no significant amount of genetic material has been lost in the translocation process.

4. Down's Syndrome Mosaicism

Patients having two different cell populations, one trisomic for chromosome G; and another normal, constitute about 2—3% of patients with Down's syndrome. This condition is usually suspected when the phenotypic expression of trisomy 21 is not fully expressed or when the intelligence of the patient is higher than expected. In addition, they may have children with Down's syndrome (Weinstcin and Warkany, 1963). Individuals having trisomy Gt mosaician may vary in phenotype from typical trisomy 21 to normal. There is no age dependency (Richards, 1969). One cannot correlate the percentage of trisomic blood cells with intelligence. Richards (1969) found about 20% more trisomic cells in fibroblasts than in lymphocytes.

If mosaicism is found in one of the parents of a child with Down's syndrome, meiotic study of ovary or testis should be carried out. There is evidence that if half the cells are abnormal, about 25% of the childern will have Down's syndrome (Mikkelsn,1971a).

I. Nonmongoloid "Trisomy G"

Several cases of nonmongoloid "trisomy G" have been published (Uchida et at, 1968; Al-Aish, 1969; Lozzio, 1969; mikkelsen, 1969). Some have been designated as having trisomy 22 to contrast with trisomy 21 (Down's syndrome). At this point in time, within this group, with two possible exceptions cited

below, there seems to be no characteristic phenotype and it would appear likely that some of these represent centric fragments that may come from several different chromosomes.

J. G deletion syndromes

There are at least two relatively distinct phenotypes presumably representing monosomy or deletion of a portion of the long arm of two different G-group chromosomes (Warren *et al*., 1973).

1. The G_1 Deletion Syndrome (Antimongolism)

This syndrome consists of mental and growth retardation, hypertonia, nail anomalies, skeletal malformations, cryptorchidism, hypospadias, inguinal hernia, pyloric stenosis, thrombocytopenia, eosinophilia, and hypogammaglobulinemia. Facial and oral manifestations include microcephaly, large low-set ears, antimongoloid obliquity of palpebral fissures, highly arched or cleft, and micrognathia. Dermatoglyphic analysis has shown a marked increase in radial loops (Schindeler and Warren, 1973).

2. The G_1 Deletion Syndrome

This syndrome has less distinctive features: severe mental retardation, hypotonia, soft tissue syndactyly of the second and third toes, and clinodactyly of the fifth finger. Facial and oral manifestation include large, low-set ears, epicanthal folds, ptosis of eyelids, highly arched palate, and bifid uvula 1971; Stoll *et al.*, 1973). Dermatoglyphc analysis has shown a marked increase in whorls, a decrease in ulnar and radial loops, a distal axial triradius, and hypothenar patterns (Schindeler and Warren, 1973).

K. Triploidy

Triploidy, as noted later in this chapter, is a frequent cause of fetal wastage prior to the eighth intrauterine week. Diploid/triploid mosaicism is occasionally compatible with survival and there have even been several examples of pure triploidy.

5.3.3 Sex chromosomal abnormalities

A. Klinefelter syndrome

In 1942, Klinefelter *et al*., described a syndrome in post pubertal males consisisting of small firm testes with tubulan by a linization but with a normal number of Leydigs cells, aerosposrmia, gynecomaztia, elevated urinary gonadotropies and low concentrations of urinary 17-ketosteroids. Several years later, Bradleury etal. (1956) and plunkeft and Barr (1956) noted Chromotin-positive nuclei in the tissues of such patients, and Jacob and strong (1959) described an

XXY sex chromosome complement in chromotin-positive klinefelters syndrome.

Chromatin-positve males have been found to comprise about 2 per 1000 live male births. These also contains XXY, XXYY, XY/XXY, and other rarer forms of Klinefelter syndrome. It has been estimated that about 80% are XXY, 10% are mosaics, and the rest are XXYY and the more unusual types.

1. XXY Klinefelter syndrome

The clinical features of XXY klinefelter syndrome do not become apparent until after puberty. Body proportions usually do not appear remarkably abnormal, however, the lower extremities tend to be and about 60% have a span that exceeds their height by 3 cm or more.

The Prepubertal testes are normal size and microscopic appearance but during adolescence they fail to enlarge and remain small and firm, averaging less than 2cm in length. The seminiferous tubules are usually shrinken, hyalinized, and irregularly arranged. Elastic fibres are absent around the tunica propria of the tubules. Ledig cells are clumped, Rarely spermatogesesis can be demostrated. In nearly all cases the testes descend. The penis is usually of normal size but may be somewhat shorter than normal. The prostate is smather than normal. Gynecomastia develops after puberty in about 50% and facial hair is sparse in about 60-75%. Axillary hair may also be less.

The prepubertal testes are of normal size and microscopic appearance but during adolescence they fail to enlarge and remain small and firm, averaging less than 2 cm in length. The seminiferous tubules are usually shrunken, hyalinized, and irregularly arranged. Those tubules which are not sclerotic are immature and lined exclusively with Sertoli cells. About 50% have a female pubic pattern. Libido and potency are usually decreased. There is some evidence of increased tendency to pulmonary disorders, varicose veins, and, possibly, breast cancer. There is the same frequency of color blindness among XXY patients as in normal females. Although intelligence may be reduced, at least 75% of XXY males have normal intelligence. Personality is usually passive.

The incidence of XXY Klinefelter's syndrome is about 1.3 per 1000 live male births. Meternal age is significantly increased for XXY but not for XXYY,XXXY, or XXXXY patients. About 60% of XXY males are X^MX^MY, while 40% are X^MX^PY. The X^MX^MY state arises from nondisjunction either during oogenesis or at an early postzygotie division. The X^MX^PY condition probably has its origin in nondisjunction during the first meiotic division.

2. XXYY Klinefelter syndrome

Patients with the XXYY variant tend to be about 4 cm taller than average height, more aggressive, and more mentally retarded than those with XXY Klinefelter's syndrome (Borgaonkar *et al,* 1970). Otherwise the phenotype is quite

similar: small firm testes, eunuchoid body build, sparse, body hair, gynecomastia, and elevated gonadotropins. Almost all XXYY males described to date have been mentally retarded and many have been aggressive (Schlegel *et al.*, 1965).

As mentioned above, there is no increase in parental age in contrast to XXY Klinefelter's syndrome. The disorder is most likely due to nondisjunction in both the first and second meiotie divisions during spermatogenesis with production of an XYY sperm. One cannot,however, rule out the less likely possibility of nondisjunction at the second meiotie division in both parents.

Dermatoglyphic studies have shown that digital arch patterns are more common in the XXYY patient than in the XXY individual who, in turn, has more than the normal male.

A child with an XXYYY sex chromosome complement was noted to have mental retardation, lordosis, flexed index and fifth fingers, pes planus, and aggressive personality (Gracey and Fitzgerald, 1967).

3. XXXY Klinefelter's syndrome

Over 25 cases of XXXY Klinefelter's syndromoe have been published (Vormittag and Weninger, 1972). All have been mentally retarded. The phenotype is similar to that of the XXY male but the size of the penis is small (McGann *et al.,* 1970).

Two late-labeling X chromosomes have been demmonstrated. However, two Barr bodies are seen in only a proportion of cells (Vormittag and Weninger, 1972). The condition may arise from successive nondisjunction in either the maternal or paternal meiotie divisions (Pfeiffer and Sanger, 1973). Dermatoglyphic findings have not been consistent. An XXXYY male has been described by Bray and Josephine (1963).

4. XXXXY Klinefelter's syndrome

There have been over 70 cases of 49,XXXXY males published since Fraccaro and Lindsten documented the first example in 1960. Nearly all have been severely mentally retarded intelligence quotients ranging from 20 to 60. A marked difference between the XXY and XXXXY male is thepoordevelopment of the external genitalia in the latter. The penis is always minute and the testes very small and undersended with hypoplastic Leydig cells and absence of term cells. The scrotum is. usually hypoplastic.

Mild nicrocephaly, ocular hypertelorism (90%), myopia (25%), strabismus (50%), mild nongoloid obliquity of palpebral fisures (35%), epicanthus (80%, and short neck with redundant skin on the nape have been noted. Skeletal anomalies present in over half the cases. Congenital heart disorders have been noted in about 20% of cases. Gonadotropins have not been elevated.

Autoradiographic evidence has shown three heavily labeled X chromosomes

(Hsu and Lockhart, 1965). Three Barr bodies may be found in a proportion of interphase nuclei (Miller and Warburton, 1968).
Parental age is not elevanted. Postzygotic nondisjunction in an XXY zygote appears to be the cause for the XXXXY state, all the X chromosomes coming from the mother (Murken and Scholz, 1967; Race and Sanger, 1969).

5. XX Klinefelter syndrome

Less than 30 cases have been published of males having 46,XX karyotypes. They exhibit many of the stigmata of Klinefelter's syndrome and, sence, will be considered here (Anderspn *et al.*, 1972).

All have small testes, are infertile, and rarely shave. About 70% have gynecomastia and elevated gonadotropin levels. Plasma testosterone levels are very low (Neuwirth *et al.*, 1972). The penis and scrotum have been small in about half the cases. All are of normal intelligence and have normal skeletal proportions.

6. Klinefelter's syndrome Mosaicism

About 15% of patients with Klinetelter's syndrome have been found to have two or more chromosomally distinct cell populations. In each of these individuals, one of the cell populatins generaly has an XXY, XXXY, or XXXXY sex chromosome constitution while the other is XX or XY. The clinical expression of mosaicism for Klinefelter's syndrome depends on the type of sex consitution present at a critical time of development. Thus, one can find, for example, and XY/XXY mosaic who is phenotypically normal, provided the XY cells exerted the predominant genetic effect.

In a study of XY/XXY mosaics, Gordon *et al.* (1972) found that only one-half exhibited azospermia and about one-third had gynecomastia and elevated gonadotropins. About one-quarter had germinal epithelium. Among 6 patients with XX/XXY mosaicism, Ferguson-Smith (1969) noted comparable findings.

B. XYY syndrome

Although the presence of an extra Y chromosome had been described as early as 1961 (Sandberg et al., 1961), interest was markedly aroused by a finding of a disproportionately high percent (usually 2—4%) of such individuals in prisons and mental hospitals (Casey *et al,* 1968; Jacobs *et al.,* 1968; Marinello *et al.*, 1969; Hook 1973). It was soon noted that most XYY patients are excessively tall and not uncommonly midly mentaly retarded (mean intelligence quotient —90) (Valentine et al., 1971). However, the frequencey of the condition among newborn male infants is about 1 per 700 births (Ratcliffe *et al.,* 1970) and few of these individuals lead other than quite routine lives. The adult height of an XYY individual is usually over 180 cm while XYY children are usually above the 90th

percentile in height by 6 years of age. Leg length and trunk length are increased but the leg/trunk ratio is normal (Keutel and Dauner, 1969). Muscle weakness (especially of the pectoralis major) and poor coordination are commonly noted. Phenotypical alterations are subtle: mild facial asymmetry, mild pectus excavatum, and mild scapular winging. The ears tend to be long and often there is a bony chin point. Most have exhibited normal sexual development (Court Brown, 1969). There are no characteristic dermatoglyphic aherations (Hubbell *et al,* 1973).

The disorder probably arises from paternal nondisjunction during the second meiosis. Retarded intelligence (I.Q.—70), impulsive aggressive behaviour, bilateral simian creases, clinodactyly of fifth fingers, retarded bone age with pseudoepiphyses at the bases of the metacarpals and metatarsals, and lack of patellar epiphyseal calcification were described by Schoepflin and Centerwall (1972).

Ridler *et al.* (1973) noted low normal intelligence, behavior problems with aggressive outbursts, repeated pulmonary infections, hypotrophic testes, sparse body hair, and acne in a 48,XYYY patient. Conversely, Hunter and Quaife (1973) decribed no stigmate other than sterility.

C. Turner's syndrome

In 1938, Turner described a syndrome in postpubertal female^consisting of sexual infantilism, short stature, webbed neck, and cubitus valgus. Albright *et al.* (1942) showed that these patients had an elevated urinary excretion of gonadotropins, and Wilkins and Fleischmann (1944) described "streak" gonads devoid of ovarian folicles in such cases. Polani *et al.*, (1954) and Wilkins *et al.*, (1954) demostrated that most cases are chromatin-negative, and Ford *et al.* (1959) first described the XO karyotype.

Turner's syndrome has been estimated to occur 1 per 2500 female births (Maclean *et al.*, 1964; Mikamo, 1968) and has been frequently noted in abortuses. Parental age is not increased.

Variation in phenotype has led to some confusion concerning nomenclature. Since the most common features are short stature, streak gonads, and X monosomy or short arm loss of X chromosomal material, all patients with these features are classified here as examples of Turner's syndrome (Yunis, 1965). Deletion of the long arm of the Y chromosome has occasionally been associated with the Turner syndrome. Cases with streak gonads and sexual infantilism but of normal or increased suture and normal female or male sex chromosome complement will be referred to as having "pure gonadal dysgenesis" or, more accurately, XY or XX gonadal dysgenesis.

Primary amenorrhea and sterility are almost constant features of the XO Turner syndrome although exceptions have been noted. Breast development is

poor, the chest is broad with seemingly widely spaced, hypoplastic, at times, inverted nipples. The external genitalia are infantile and public hair is sparse.

The histological pattern of the dysgenetic gonad found in Turner's syndrome consists of long streaks of white wavy connective tissue stroma without follicles. Follicles are present, however, in fetal and infantile ovaries of patients with Turner's syndrome (Gordon and O'Neill, 1969).

Adult height is usually less than 57 inches (144 cm). Various skeletal anomalies include cubitus valgus (about 75%), short fourth metacarpals (about 65%), deformity of medial fibial condyle (about 40%) osteoporosis (about 50%), hypoplasia of first cervical vertebra (about 40%), and small carpal angle.

Birth weight is below the 3rd percentile in about half the cases. In infants, excess skin on the nape and peripheral lymphedema have been noted in 15—50% of the cases. During embryonic life, neck blebs or cystic hygroma are common (Singh and Carr, 1966; Rushton *et al.,* 1969). Toenails are frequently hypoblastic. With age, the excess skin on the nape metamorphoses into pterygium colli and, with improvement in deep lymphatic circulation, the peripheral lymphedema gradualy disappears. Increased numbers of cutaneous nevi are found in about 60%.

Epicanthal folds, ptosis of upper eyelids, prominent ears, and micrognathia are common facial features. The hairline is low at the nape.

Thyroid antibodies are elevated in XO Turner's syndrome but less frequently than in the X-iso X mosaic and glucose intolerance occurs with grater frequency in patients with Turner's syndrome and in their parents than in the normal population (Rimoin, 1973).

1. X Monosomy

Patients with an XO sex complement appear to comprise about 60% of the cases of Turner's syndrome. Furthermore, they appear to be more severely affected clinically than other forms of the disorder.

2. XO/XX and XO/XXX Mosaicism

Patients with Turner's syndrome may have two different cell populations, one having an XO sex constitution, the other a normal XX sex complement. Such individuals are called XO/XX mosaics and constitute about 7% of the cases of Turner's syndrome. The two cell population types may appear in every tissue of the body or only in certain ones.

Presumptive evidence for mosaicism lies in a discrepancy between sex chromatin pattern and karyotype, or through observing a low percentage of chromatin-positive nuclei (5—15%) in phenotypic females. The clinical spectrum of XO/XX mosaicism is wide and may vary from cases quite typical of Turner'

syndrome with many associated anomalies to cases with normal gonads and normal stature. About 20% menstruate (Ferguson-Smith, 1969). In contrast to patients with XO Turner's syndrome who are prone to aortic coarctation, those with XO/XX karyotypes are likely to have pulmonic stenosis with or without atrial spetal defect (Nora *et al.*, 1970), being similar to patients with Noonan's syndrome.

The usually accepted explanation for XO/XX mosaicism is loss of an X chromosome during cleavage in the early embryo.

About 5% of the cases of Turner's syndrome are XO/XXX mosaics. Clinically they resemble the XO/XX mosaic. Patients having three stem lines XO/XX/XXX have been reported but are quite similar phenotypically to XO/XX mosaics.

3. Isochromosome X (XXqi)

About 20% of patients having Turner's syndrome have an X isochromosome, i.e., replication of the long arm of the late replicating X chromosome. They exhibit short stature, sexual infantilism, primary amenorrhea, and skeletal anomalies.

The Barr body and polymorphonuclear neutrophils drumsticks are larger than normal. Drumsticks are also more nuemerous (Taft *et al.*, 1965; Sparkes and Motulsky, 1967).

4. Short and Long Arm Deletion of an X chromosome

Deletion of the short arm of an X chromosome (XXp—) results in the Turner phenotype. They are as short as individuals with the XO Turner syndrome but are less likely to have associated malformations (Atkins *et al.*, 1965).

Deletion of the long arm of an X chromosome (XXq—) is far less likely to be asociated with short stature. The girl was 159 cm tall and exhibited no stigmata of Turner's syndrome. She never menstruated and her ovaries were not palpable. As expected, her Barr bodies were smaller than normal. Xg1 studies showed that the Xpi was of maternal origin. However, Turner's syndrome has been reported in association with XXq—.

5. Y deletion

At least a dozen cases of Turner's syndrome associated with a dicentric Y Chromosome have been published (Armendares *et al.*, 1972; Cohen *et al.*, 1973). All have short stature, female phenotype, and most have associated anomalies. A patient with long arm isochrorriosome Y had drumstiicks have been noted in polymorphonuclear leukocytes.

5.4 Human genome

5.4.1 The Human Genome Project

As the recombinant DNA, gene cloning and DNA sequencing technologies improved in the 1970s and early 1980s, scientists began discussing the possibility of sequencing all 3 x 109 nueleotide pairs in the human genome. These discussions led to the lauching of the **Human Genome Project** in 1990. The initial goals of the Human genome project was to construct a detailed physical map of the entire human genome, and to determine the nueleotide sequences of all 24 human chromosomes by the year 2005. Scientists soon realized that this huge undertaking should be a worldwide effort. Therefore, an international **Human Genome Organization** (HUGO) was organized to coordinate the efforts of human geneticists around the world.

James Watson, who with Francis Crick, discovered the double-helix structure of DNA, was the first director of this ambitious project, which was expected to take nearly two decades to completle and to cost in excess of $3 billion. In 1993. Francis Collins, led the research teams that identified the cystic fibrosis gene, replaced Watson as director of the Human Genome Project. In addition to work on the human genome, the Human Genome Project has served as an umbrella for similar mapping and sequencing projects on the genomes of several other organisms, including the bacterium *E. coli*, the yeast *S. cerevistae*, the fruit fly D. *melanogaster*, the plant A. *thaliana*, and the worm *C. elegans.*

5.4.2 Bacterial Genomes

In 1995, *Haemophilus influerzae* was the first bacterium to have its genome sequenced in its entirety. By mid-2001, the complete sequences of 32 bacterial genomes were abailable in the public databases (collections of the sequences of genes, chromosomes and genomes). The genomes range in size from 580,070 bp for *Mycoplasma genitalium*, which is thought to have the smallest geneme of any self-replicating organism, to 4,411,529 bp for *Mycabacterum tuberculosis*, which causes more human deaths than any other infections bacterium, to 4,539,221 bp for *Escherichia coli*, the best-known cellulalr microorganism. The genome of *M. genitalium* is of special interest because it may aproximate the "minimal gene set" for a self-replicating organism—the smallest set of genes that will alow an organism to reproduce itself. Of course, the genome of *M. tuberculosis* is of great interest becauseof the pathogenicity of this organism and the hope that a complete understanding of its metabolism will suggest ways to prevent tuberculosis in humans: The need for new ways to combat this pathogen has been enhanced by the recent evolution of antibiotic-resistant strains of *M. tuberculosis*.

Of the bacterial genomes sequenced to date, the genome of *E. coli* (Fig. 5.9)

has undoubtedly caused the most excitement among geneticists. *E. coli* is the most planet.Geneticists, biochemists, and molecular biologists have utilized *E. coli* as the prefered model organism for decades. Most of what is known about bacterial genetics was learned from research on *E. coli.*

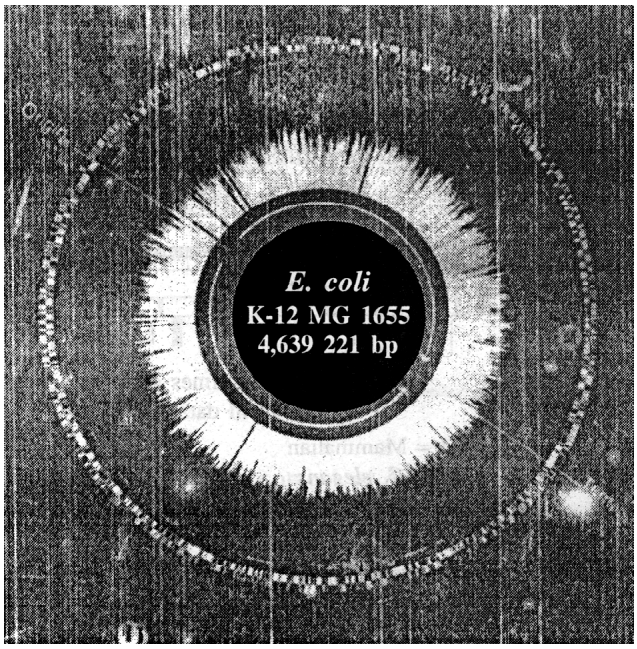


Fig. 5.9 Sequence-based map of the chromosome of *Escherichia coll.* The blue arrows mark the halves of the chromosome travered by the two replication forks. The outer concentric circle gives the position of genes encoding proteins similar to bacteriophagc proteins. The second concentration circle shows the location of genes that are transcribed clock-wise (gold) from one strand or counterclockwise (yellow) from the complementary strand. The sunburst in the center is a histogram in which the length of each ray is proportional to the randomness of codon usage within each coding sequence

The *E. coli* genome contains 4288 putative proteincoding sequences or genes. About one-third of these are well-studied genes encoding known products, whereas 38 percent are of unknown function, Putative protein-coding sequences that are not knwn to encode proteins are called **open reading frames** or **ORFs**. An **ORF** is a nucleotide sequence that begins with a translation-initiation signal (usually ATG). continues with a sequenced of base triplets specifying amino acids, and ends with one of the three translation-termination singnals.

The average distance between genes (size of intergenic regions) in the E. coli genome is 118 bp. Known and putative genes specifying proteins and stable RNAs make up 87.8 percent and 0.8 percent of the genome,respectively, and noncoding repetitive elements account for 0.7 percent of the genome. Thus, 10.7 percent of the genome must involve regulatory sequences and sequences with unknown functions.

Once the complete sequence of a bacterial genome is available, it can be searched using computers for similarities with other sequenced genomes. Such sequence comparisons can often be used to gain inferences about gene function. Because so much is known about gene function in *E. coli*, comparisons with other secuenced bacterial genomes are often very informative. For example, a comparison of the genomes of *Treponema pallidum*, the parasitic spirochete that causes syphilis, and *E. coli* shows that *T. pallidum* contains the genes that encode proteins involved in DNA replicartion and repair, transcription, and translation, but carries few genes encoding biosynthetic enzymes and trtanspoort protiens.

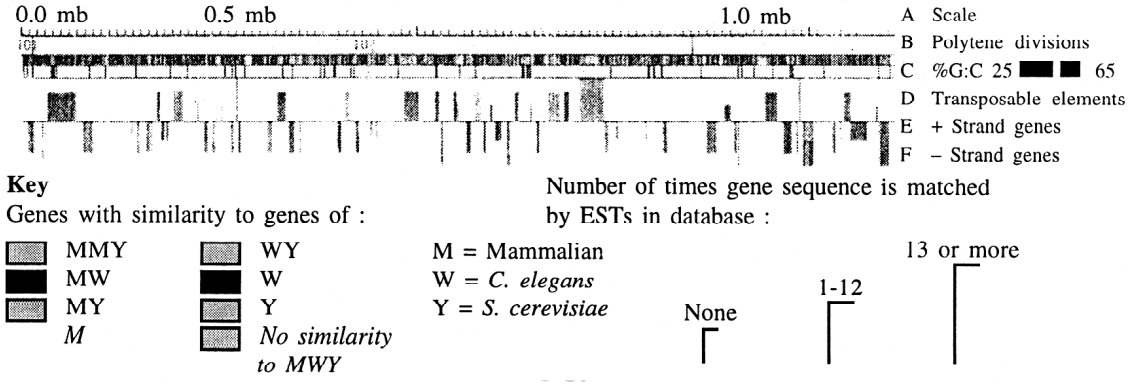


Fig. 5,10 Sequence-based map of chromosome 4 of *Drosophila.* The top line (A) gives map position in megabase paris (1 mb = 1 million hase pairs). The second panel (B) gives the polytene chromosome band number. The third panel (C) shows the percetage of G:C base pairs, and the fourth (D) shows the positions of transposable genetic elements. The bottom two panels show the positions of genes where transcription occurs with the plus strad as template (F) or with the minus strand as template (F). The color of each gene box in panels F and F indicates its similarity to genes of mammals, *C. elegans,* or *S. cerevisiae,* as shown by the key below the map, and the height of each gene box in panels E and F indicates the frequency of the sequence in the expressed-sequenced tag (EST) database for *Drosophila*

Transposable genetic elements make up about 10 percent of the *Arabidopsis* genome, far less than the 50 to 80 percent of the corn genome estimated to be derived from transposable elements.

The next challenge is to determine the functions of the *Arabidopsis* gene products. There is where this little weed really shines as a model system. It is idealy suited for genetic dissections of biological processes. The goal or the *Arabidopsis* research community is to determine the function of all 25,498 genes in the nect 10 years.

5.4.3 The Human Genome

Recall that the initial goals of the Human Genome project were (10 to map all the human genes, to construct detailed physical maps of all 24 chromosomes, and to sequence the entire genome by 2005. All of these goals are likely to be achieved ahead of schedule. With two first-draft sequences of the human genome already published in February 2001 a complete sequence of the euchromatic portion of the human genome will certainly be available long before 2005.

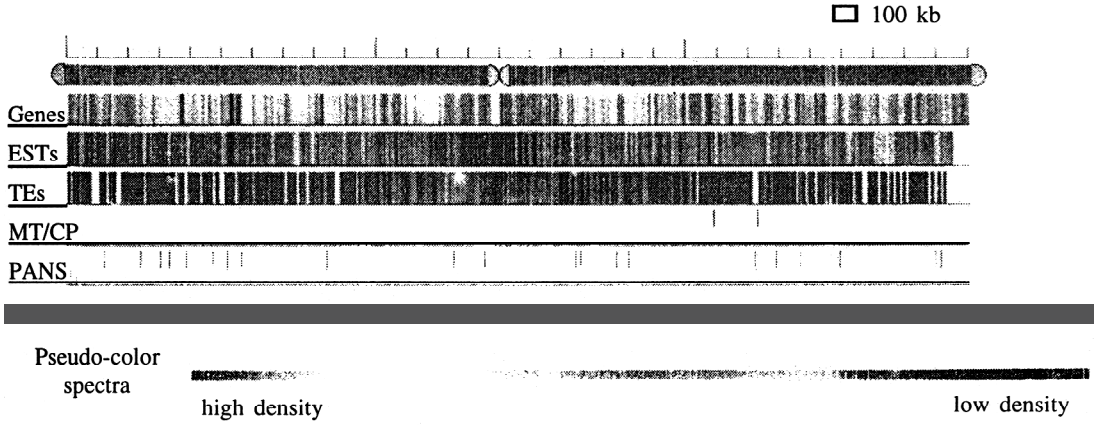


Fig. 5.11 Sequence-based map of chromosome 1 of *Arabidopsis.* Distance in 100-kb units is shown at the top. The chromosome is represented by the top bar, with sequenced regionsin red and the unsequenced centromere and telomerees in blue. The densities of genes (top panel below chromosome), matches to expressed sequence tags (ESTs; second panel), transposable small nuclear RNAs (bottom panel) are color coded, with red representing the highest density and dark blue the lowest density

Progress in mapping the human genome has been excellent. Complete physical maps of chromosomes Y and 21 and detailed RFLP maps of the X chromosome and al 22 autosomes were published in 1992. By 1995, the genetic map contained markers separated by, on average, 200 kb. A detailed micro satellite map of the human genome was published in 1996, and a comprehensive map of 16,354 distinct loci was released in 1997. All of these maps have proven invaluable to researchers cloning genes based on their locations in the genome.

Unfortunately, the resolution of genetic mapping in humans is quite low—in the range of 1-10 mb. The resolution of fluorescent *in situ* hybridization (FISH) is also approximately 1 mb. Higher resolution mapping (down to 50 kb) can be achieved by **radiation hybrid maping,** which is a modification of the somatic-cell hybridization mapping procedure. Standard somatic-cell hybridization involves the fusion of human cellls and rodent cells growing in culture and the correction of human gene products with human chromosomes retained in the hybrid cells.

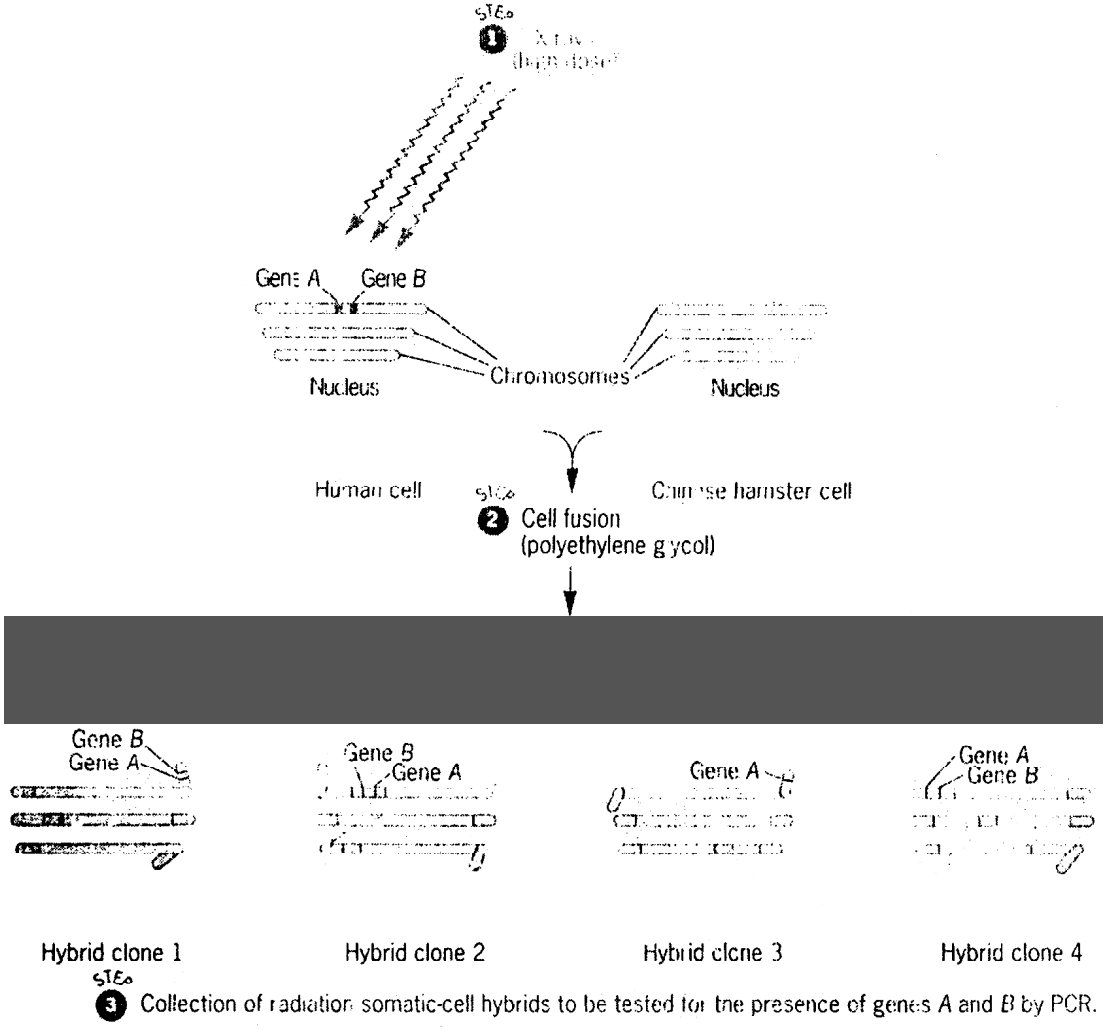


Fig. 5.12 The use of radiation somatic-cell hybrids for high-resolution mapping of the human genome. The rationale behind radiation hybrid maps in that the probability of an X ray-induced break between genes *A* and B is directly proportional to the distance between them on the DNA molecule. Note that genes *A* and *B* have remained together in hybrid clones 1, 2, and but were separated by an X ray-enduced break during the formation of hybrid clone 3

Radiation hybrid mapping is done by fragmenting the prior to cell usion (Fig. 5.12). The irradiated human eels are then fused with Chinese hamster (or other rodent) cells growing in culture,usually in the presence of a chemical such as polyethylene glycol to increase the efficiency of cell fusion. The human—Chinese hamster somatic-cell hybrids are then identified by growth in an appropriate selection medium.

Many of the human chromosome Chromosome 1 fragments become integrated into the Chinese hamster chromosomes during this process and are transmited to progeny cells just like the normal genes in the Chinese hamster chromosomes The polymerase chain reaction is the used to screen a large panesl of the seleted hybrid cells for the presence of human genetic markers. Chromosome maps are constructed based on the assumption that the probability of an X ray-induced break between two markers is directly proportional to the distances separating them in chromosomal DNA.

Several groups have used the radiation hybrid mapping procedure to construct high-density maps of the human genome. In 1997, Elizabeth Stewart and coworkers published a map of 10,478 STSs based on radiation hybrid mapping; their map of human chromosome 1 is shown in Fig. 5.13.

Whereas the gene maping work advanced quickly, progress towards sequencing the human genome initially lagged behind schedule. However, that all changed rapidly beginning inl998. During May of 1998, J. Craig Venter announced that he had formed a private company, Celera Genomics, with the goal or sequencing the human genome in just three

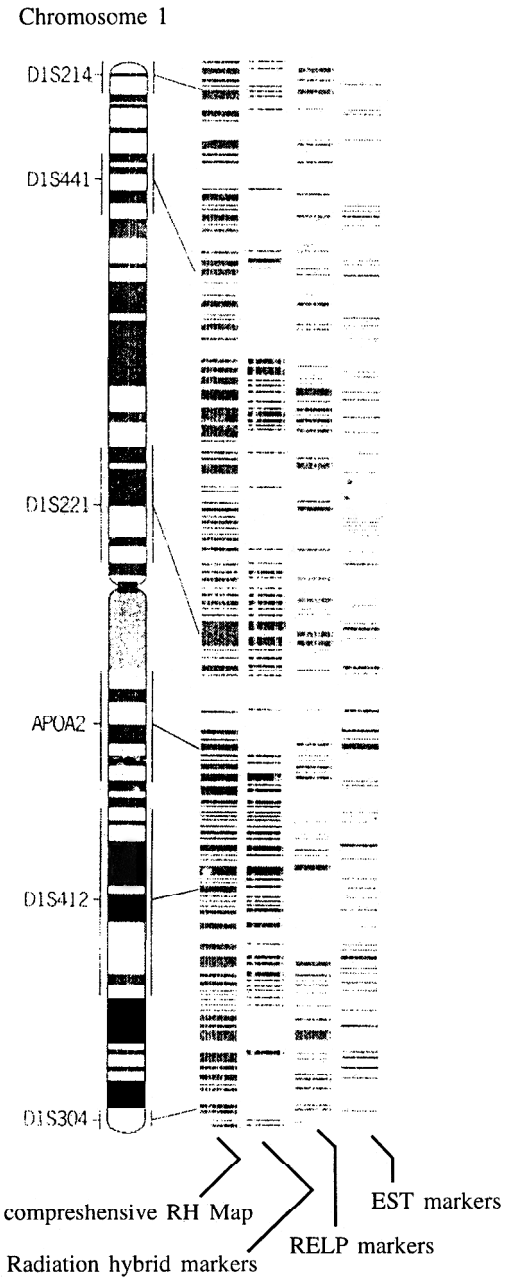


Fig. 5.13 A high-resolution radiat on hybrid map of human chromosome 1. The cytogenetic map of chromosome is shown on the locations of the comprehensive radiatipon hybrid map showing all the markers (red lines), the high confidence radiation hybrid markers (blue lines), the high confidence radiation hybrid markers (blue lines), the highconfidence radiation hybrid markers (blue lines), the high confidence radiation hybrid markers (blue lines), the RFLP markers (green lines), and the ESTs (parple lines)

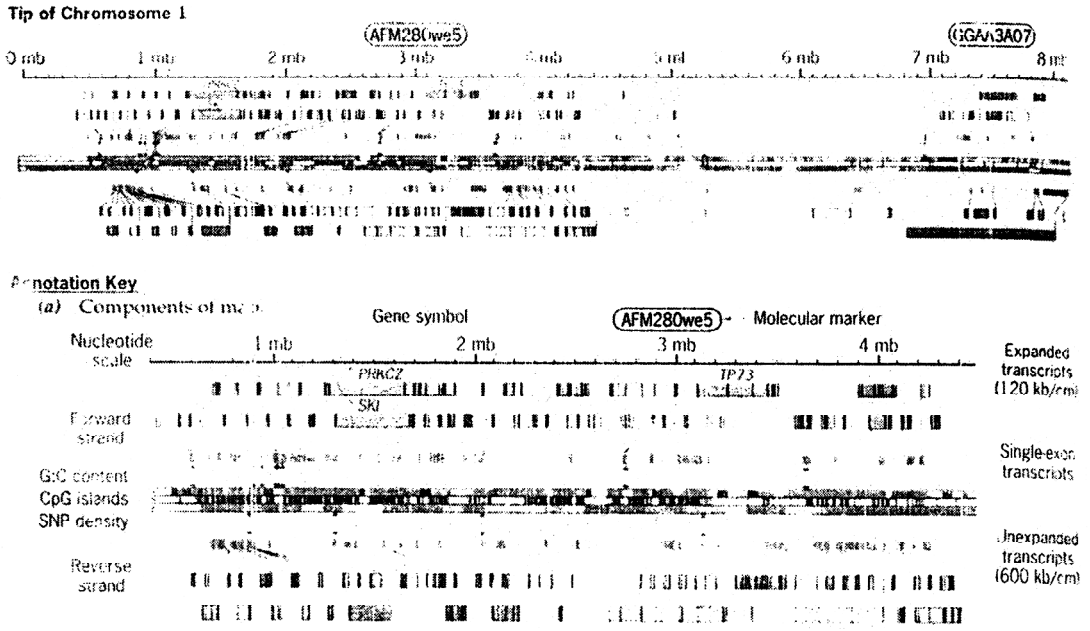


Fig. 5.14 A nnortated, sequence-based map of an 8 mb segment of DNA at the tip of human chromosome 1, assembled by rescarel ers at Celera Genomics. The top line gives distances in mb. The next three panels show predicted transcripts from one strand of DNA (the "forward strand"), whereas the bottom three panels show predicted transcripts specified by the other strand of DNA (the "reverse strand"). The middle three panels give the G:C content, the positions of CpG islands, which occur upstream of genes, and the density of single nucleotide polymorphisms (SNPs), respectively. The annotation key below the map of chromosome 1 shows the components of the map, the color code for gene product functions, and the color codes for G:C content and SNP density

years. Shortly thereafter, the leaders of the public Human Genome Propject's sequencing laboratories announced that they had revised their schedule and planned to complete the sequence of the human genome by 2003—two years earlier than originally poroposed. From this point in time, everything accelelrated.

The complete sequence of the first human chromosome—small chromosome 22—was published in December 1999. The complete sequence of human chromosome 21 followed in May of 2000. Then, with the interventionof the White house, Venter, of Celera Genomics, and Francis Collins, Director of thepublic Human Genome Project, agreed to publish first drafts of the sequence of the human genome at the same time. The Celera and public sequences were both published in February 2001. Figure 5.14 shows an annotated, sequence-based map of an 8mb segment at the tip of the short arm of human chromosome 1. This map ilustrates the positions and orientations of known and predicted genis in one small portion of the human genome. For summar maps of the entire human genome, see the February 15,2001, issue of *Nature* and the February 16, 2001, issue of *Science.*

The amount of information in these first drafts of the human genome was quite overwhelming including the sequence of over 2650 megabase pairs of DNA (over 2,650,000,000 bp). The human genome is more than 25 times the size of the previously sequenced Drosopbila and Arabidopsis genomes, and it is eight times the sum of al previously sequenced genomes.

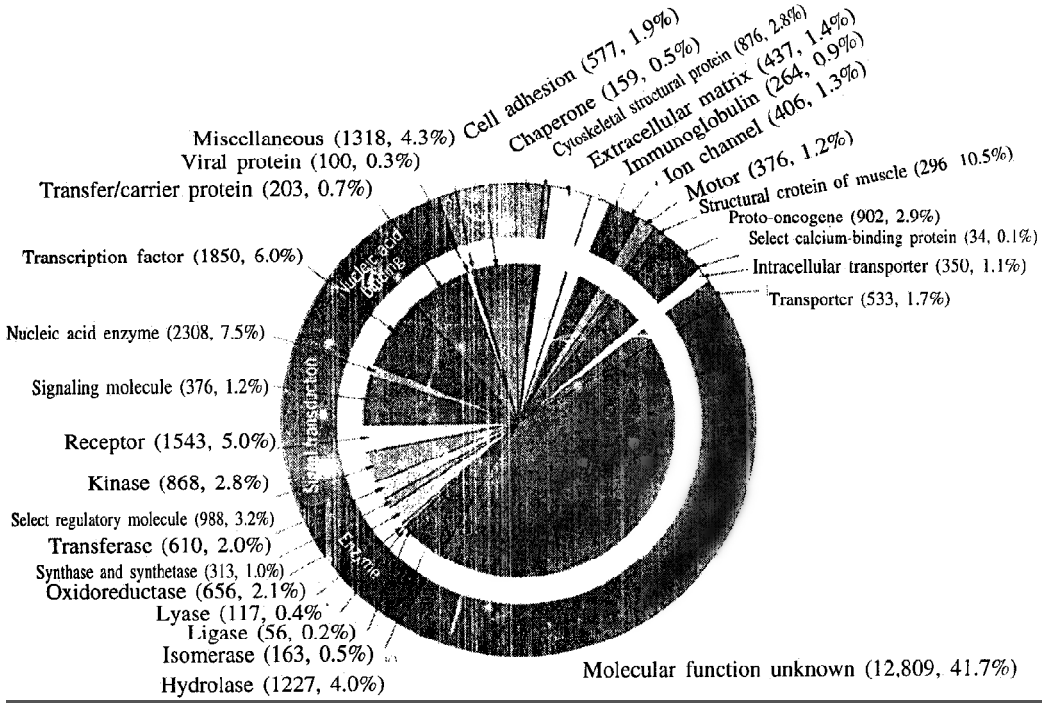


Fig. 5.15 Functional classification of the 26,283 genes predicted by Celera Genomics' first draft of the sequence of the human genome, Each sector gives the number and percentage of gene products in each functional class in parentheses. Note that some classes overlap: a proto-encogene, for example, may encode a signaling molecule

The sequence of the human genome provided one surprise: there appear to be only about 30,000 to 35,000 genes rather than the estimated 50,000 to 120,000 genes suggested by earlier studies. The distribution of functions for the 26,383 genes predicted by the Celera sequence is shown in Figure 5.15. About 60 perecent of the predicted proteins have similarities with proteins of other species whose genomes have been sequenced (Figure 5.16). Over 40 percent of the predicted human proteins share similarities with *Drosopbila* and *C. elegans* Proteins. The picture is quite different for mamilies of closely related proteins, which tend to perform important basic cellular functions. Only 94 of 1278 protein families predicted by the sequence of the human genome are specific to vertebrates. The rest have evolved from domains of proteins in distant ancestors, including prokaryotes and unicellular eukaryotes.

On average, there is one gene per 60 to 85 kh in the human genome, although there is some clustering of highly expressed genes in euchromatic regions of specific chromosomes. Exons make up only 1.1 percent of the genome, whereas inttrons make up 24 percent, with 75 percent of the genome being intergenic DNA. Of the intergenic DNA, at least 44 percent is derived from transposable genetic elements. The initial drafts of the human genome are far from complete, and the immediate goal will be to fill in the gaps in the genome ad produce a finished sequence in the next year or so. The other major goal is to determine the structure and function of the human proteome (all of the proteins encoded by the human genome).

Knowledge of the nucleotide sequences of entire genomes has spawned the development of a new scientific discipline, *bioinformatics*, a fusion of computer science and biology, with the goal of developing new tools with which to analyze

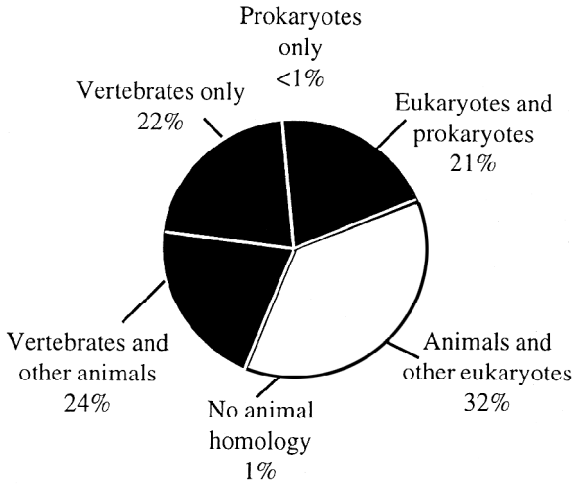


Fig. 5.16 Pie chart showing homology of predicted human proteins to protems of other species for those where homologues were deterced by comoputer searches of the public databases

the welth of data that genomics in providing. These new tools alow scientists to search genome databases for specific sequences or structrural features, to compare various features of different genomes, and to male inferences about the evolution of genomes. Indeed, with the sequence of the human genome aproaching completion, the question being asked is which genomes to sequence next—the mouse genome, the chimpanzee genome, and so on. One point is very clear: comparative genomics is providing unprecedented information about the evolution of species.

Unit 6 Cytogenetic Implications and Consequences of Structural Changes and Numerical Alterations of Chromosomes

Structure

6.1 Introduction

Between 20—50% of human abortuses have been shown to have a chromosome abnormality (Rashad and Kerr, 1965; Thiede and Metcalfe, 1966; Carr, 1967; Larson and Titus, 1970; Kajii *et al.*, 1973). An even wider range has been reported (8—50%), but the studies have not been comparable (Carr, 1971a).

Analysis of over 350 cases compiled from several published series has demonstrated that 45,XO is the most frequent single anomaly, constituting about 20% of the cases. Triploidy is only slightly less frequent, i.e., about 15%, while tetrapoidy has been demonstrated in 5%. Trisomies, as a group, have been found in about 50% of abortuses (E, 15%; G, 15%; D, 10%; C, 5%; A, B, F, 5%). The remainder are mosaics, or translocations (Carr, 1965, 1967; Inhorn 1967; Larson and Titus, 1970). only rarely is autosomal monosomy found (Kajii *et al.*, 1973).

It should be pointed out that within the E group, trisomy 16 and not trisomy 18 largely comprises this number (Carr, 1967; Waxman et a/., 1967). Use of recently developed banding techniques have shown trisomies for chromosomes 2, 3, 4, 6, 7, 8, 9, 10, 14, 15, 16, 18, 21 and 22 (Lauritsen et ah, 1972 Kajii *et al.*, 1973).

It is likely that findings in spontaneous abortuses do not necessarily reflect the frequency of clinical anomalies at conception since the more lethal ones probably never survived past a few cell divisions. This may explain the absence of viable trisomic states for A, B, most C, 14, 15, 16, 17, and F chromosomes being a more likely explanation than low frequency of meiotic or mitotic error for these chromosomes. The apparent high lethality of the 45,XO embryo cannot be explained. There is suggestions that embryos with chromosomal abnormalities are more likely to be aborted earlier than those with normal karyotypes (Carr, 1965, 1967; Szulman, 1965; Dhadial *et al.*, 1970).

The gross appearance of abortuses have some correlation with their chromosome status. A recognizable fetus is most frequent in the 45, XO group. Not uncommonly they can be recognized by cystic hygromas (neck blebs) in older fetuses. Triploid abortuses characteristicaly exhibit hydatidiform degeneration of villi and only rarely contain an embryo (Szulman, 1965; Carr, 1965, 1971b; Singh and Carr, 1967; Boue *et al.*, 1967). Trisomic abortuses do not have any specific phenotype with the possible exception of those having D-group trisomy,which not uncommonly have facial clefts. Presumably most of these are 13 trisomics (Roux, 1970). Kajii *et al.* (1973) found no example, however, of trisomy 13 in their large series.

When analyzed for mean maternal age, polyploid and XO abortuses have been found to be from younger mothers while trisomics have been from older mothers (Carr, 1965, 1971a; Szulman, 1965; Kerr *et al.*, 1966). However, the mean maternal age for 45,XO abortuses has been higher than that of survivors (Dhadial etal., 1970). Arakaki and Waxman (1970) found an increase in mean maternal age in cases of 16 trisomy abortuses.

In those couples having a history of two or more spontaneous abortions. 1 in 26 couples was found to have a translocation. This contrasts with the 0.4% found in the general population (Lucas *et al.*, 1972).

6.2 Chromosomes and cancer

Cancer cells may have bizarre karyotypes which may be hypodiploid, hyperdiploid, triploid, hypertriploid, hypotriploid, etc. Many unusual structural abnormalities have been described. Cells having over 1000 chromosomes have been documented. On the other hand, cancer cells have been described with normal karyotype and no evidence of structural abnormalities. While one may conclude that not all neoplasia is associated with gross chromosomal anomalies, one cannot exclude point mutations, gene deletions or duplications, or hidden rearrangements.

Review of chromosomal alterations concerning specific tumors is be yond the scope of this review and the reader is referred to Cervenka and Koulischer (1973).

A. Acute leukemias

In no form of acute leukemias have any specific chromosome abnormalities been described and, in at least half the cases, normal karyotypes have been found (Sandberg *et al.*, 1968; Krogh-Hensen, 1969; Whang Peng *et al.*, 1969). Furthermore, ostensibly identical clinical types of acute leukemia may minifest different chromosmoe patterns. Karyotypic changes, when present, are confined to the leukemic cells of the marrow or other organs. Long-term culture of

leukocytes from the blood of patients with acute leukemias is rarely successful. When an abnormal karotype is discerned, it seems to exhibit more hypoploid cell lines while acute lymphoblastic leukemias have more hyperploid lines. Karyotype analysis alone cannot be employed either for diagnosis or for prognosis concerning survival (Cervenka and Koulischer, 1973). During remission, aneuploid cells may disappear from the marrow only to reappear on relapse (Sandberg and Hossfeld, 1970).

B. Chronic myelogenous leukemia

In 1960, Nowell and Hungerford found delection of part of the long arm of a G chromosome associated with chronic myelogenous leukomia (CML). This unique structural abnormality, termed the Philadelphia chromosome (Ph^1), has been noted in over 90% of patients with CML (De Nava, 1969). Chronic myelogenous leukemia without the ph^1 chromosome and the ph^i chromosome without chronic myelogenous leukemia have been thoroughly reviewed by Cervenka and Koulische 91973).

The Ph^1 chromosome represents deletion with translocation to the long arm of a chromosome no. 9 (see Chapter 3). With.the use of quinacrine mustard fluorescent technique, it has been shown to be a G_{22} chromosome (Caspersson *el al.*, 1971; O'Riordan *et al.*, 1971). Its occurrence is limited to hematopoietic cells of all the granulocytic, erythocytic, and megakaryocytic types (Tough *et al.*, 1963; Clein and Flemans, 1966). Other tissues, such as skin fibroblasts, do not contain the ph1 chromosome.

The best technique for demonstration of the ph! chromosome is by direct study of bone marrow (Sandberg and Hossfeld, 1970).

It is an acquired, not an inherited, characteristic as demonstrated by its presence in only one of monozygotic twins with chronic myelogenous leukemia and not in the healthy co-twin (Jacobs *et al.*, 1966, Kosenow and pfeiffer, 1969). An unusual subgroup of ph'-positive CML patients are males who are missing the Y chromosome in all or in a portion of their marrow cells. However, fibroblasts and blood lymphocytes contain the Y chromosome (Lawler and Galton, 1966; Pedersen, 1968). Two or more ph1 chromosomes appearing in marrow cells either heralds or accompanies the transformation of CML to a blastic phase (Smalley, 1966).

A ph^1-like chromosome has been found in a small porportion of marrow cells of patients with acute myeloblastic leukemia, polycythema, thrombocytopenia, myeloid metaplasia with myelofiborsis, and erythroleukemia (Sandberg and Hossfeld, 1970). Khan (1973) reported two ph'-like chromosomes in acute myeloid leukemia.

C. Solid tumors

Most malignant tumors have aneuploid karyotypes ranging from hypodiploidy to extreme hyperdiploidy. Human tumor cell populations are clonal in nature, some tumors having but a single clone, others of two or more. No consistent cytogenetic findings have been described, but various markers have been noted, for example, a missing no. 22 chromosome in meningiomas (Zaftg and Singer, 1967; Mark *et al.*, 1972) and microchromosome in various neurogenic tumors (Cox *et al.*, 1965; Levan *et al,* 1968, Kucheria, 1968). Metastatic cells tend to have a higher ploidy and more variability in chromosome number (Sandberg *et al.,* 1967). In about 50% of the cases, abnormal ("marker") chromosome have been found in metastatic cancer cells. In general, there is a tendency toward relatively few chromosomes with distally placed centromeres, i.e., fewer B-, D-, and G-group chromosome and more A-3, C-group, and E 16 chromosomes (Atkin, 1970). Manolov and Manolova (1972) described a market band in a chromosome 14 in Burkitt's lymphoma.

Precancerous lesions, largely of the uterine cervia, have shown that dysplastic lesions exhibit chiefly pseudo-or near diploid karyotypes while carcinoma in situ shows an increase in ploidy and aberrations. Invasive carcinomas exhibit near diploid patterns, showing that progression does not depend on high chromosome counts (Atkin *et al,* 1967). Benign tumors have normal diploid karyotypes.

D. Waldenstrom's macroglobulinemia

Waldenstrom (1944) described a disorder characterized by intractable anemia and increased amounts of macroglobulin in serum, accompanied by fatigue, epistaxis, gingival hemorrhage, disturbances in vision, moderate lymphadenopathy, high sedimentation rate, and bone marrow lymphocytosis (Kok *et al.*, 1963). It is presently classified in the group of gammopathies. The disease usually appears after age 40 and is more frequent in males. Its relationship to lymphosarcoma and leukemia is not clear but patients diagnosed as having the disease sometimes develop chroniclymphatic leukemia or lymphoma.

Bottura *et al.* (1961) first described the presence of 47 chromosomes in about 50% of the cells, the supernumerary being about the size of an A-group chromosome. This finding was son confirmed by German *et al*. (1961) andBenirschke *et al*. (1962), who employed the term "W" chromosome.

The morphology of the marker chromosome is not constant. It usualy has been large with the centromere varying from metacentric to subterminal, but in some cases it has been as small as an F-group chromosome (Spengler *et al.,* 1966). The marker has been noted in both marrow and in peripheral cells in form 0—50% of cells (De Nava, 1969).

The abnormality is apparently acquired. Spengler *et al.* (1966) demonstrated the marker in one monozygotic twin who had Waldenstrom's macroglobulinemia but not in his normal co-twin. Interesting also are the findings of Lustman *et al.* (1968), who described an affected female with the marker whose otherwise healthy son had a normal karyotype but had an eleyated γ-globulin peak. Elves and Brown (1968) described the marker in 4 of 6 relatives of a patient with the disorder. Only one of the individuals had an elevated γ_1 fraction.

6.3 Diseases associated with spontaneous chromosome aberrations

At least seven inherited diseases have been found to be associated with spontaneous chromosome aberrations and increased frequency of leukemia or other neoplasias. The chromosome aberrations consist of gaps (achromatic regions), chromatid and chromosome breaks, fragments, reunion or translocation figures, ring chromosomes, and dicentric chromosomes. It should be emphasized that chromosome breakage may be very rarely seen in cells of ostensibly normal people. The enzyme deficiencies in the inherited disorders may result either in increased frequency in which openings appear in the DNA strands or in decreased speed with which such breaks are healed. Higurashi and Conen (1973) demonstrated greater *in vitro* chromosomal sensitivity in several of these disorders.

Fanconi's anemia, inherited as an autosomal recessive trait, is characterized by generalized skin pigmentation, pancytopenia with marrow hypoplasia, thumb and radius anomalies, hypogenitalism, and microcephaly (Fanconi, 1967). In 1964, Schroeder *et al*. noted that more than 40% of analyzed metaphases from peripheral blood cultures of patients with Fanconi's anemia exhibited chromatid gaps and breaks and chromosomal rearrangements. Direct bone marrow preparations have shown about 10% aberrant metaphases, usually involving B and C group chromosomes (Hirschman *et al.*, 1969; Shahid *et al.*, 1972). Of 41 cases subsquently studied, 36 were found to have similar findings. Among 170 known cases, four have terminated in leukemia and one had skin cancer (Swit and Hirschhorn, 1966; Swift, 1971). Heterozygotes have an increased frequency of leukemia (Gmyrek *et al.*, 1967; Swift, 1971). Occasionaly quadriradials and dicentric forms are noted but far less frequently than in Bloom's syndrome (vide infra).

Bloom's syndrome, consisting of growth retardation, sensitivity to.sunlight, and telagiectatic erythema, was reported by German (1969) to have chromosome breaks. Of 35 cases, four were found to have subsequently developed leukemia

or cancer, especially gastrointestinal. Cell lines with an abnormal karyotype have been described in clutured fibroblasts from a patient with Bioom's syndrome (Rauh and Soukup, 1968).

Quadriradial figures, i.e., a four-armed figure derived from two chromosomes, each arm consisting of sister chromatids of one of two homologous chromosomes. The autosomes most often involved are No. 1 and either No. 19 or 20. Asymmetric dicentric chromosomes, triradials, and abnormal new monocentric chromosomes can also be found. Heterozygotes may have the same types of figures, but less frequently than the homozygote.

Atxia-telangiectasia inherited as an autosomal recessive trait is characterized by retarded growth, progressive cerebellar ataxia, telangiectasia especialy about the face and bulbar conjunctiva, increased sonopulmonary infections, and decreased immunoglobulins (especially IgA and IgE). Approximately 10% develop lymphomas (pferiffer, 1970). Heche *et al.* (1966) reported a high frequency on in vitro chromosome breakage, a finding suported by Groop and Flatz (1967), pfeiffer (1970), and German (1972).

Lesser well-documented associations are wth glutathione reductase deficiency anemia, pernicious anemia, Kostmann's agranulocytosis (Schroeder and kurth, 1971), and possibly xeroderma pigmentosum (German *et al.*, 1970).

Matsaniotis *et al*. (1966) found approximately 20% aberrant cells from direct bone marrow preparations of a baby with Kostmann's agranulocytosis. Krogh-Jensen and Friis-Moller (1967) and Bottura and Continho (1988) described in vivo demonstration of chromosomal aberration in untreated pernicious anemia. The evidence for dominatntly inherited glutathione reductase deficiency is less solid and seems to depend on the stage of the disease (Hampel *et al.*, 1969).

German *et al*. (1970) detected a tendency toward the formation of pseudodiploid clones in cultured fibroblasts from a patient with xerox deama pigmentosum, an autosomal recessively inherited disorder, in which there is a proclivity toward development of skin cancer. Failure DNA repair following untraviolet light exposure has been demonstrated. Repair failure results from deficiency of ultraviolet-specific endonuclease.

Unit 7 Microbial Genetics

Structure

7.1 Introduction

Main constituent in microbial genetics are the bacteria. Bacteria reproduce asexually. However **parasexual** reproduction is also found among bacteria. In fact, genetic information is transferred from one bacterium to another by three totally distinct processes—**transformation, conjugation and transduction.**

Gene transfer in Bacteria :

Transformation (Definition) : This is the process by which a donor DNA molecule is taken up from the external environment and incorporated into the genome of a recipient cell.

Conjugation (Definition): This is the process by which bacterial cells make direct contact with each other, and DNA is transferred from one cell (the donor) to the other (the recipient cell).

Transduction (Definition) : This is the process by which DNA is transferred from one bacterial cell to another by a bacterial virus, or bacteriophage.

7.2 Bacterial mutation

Infection by the bacteriophage leads to the reproduction of the virus at the expense of the bacterial cell, which is lysed or destroyed. If a plate of *E. coli* is homogeneously sprayed with Tl, almost all cells are lysed. Rare *E. coli* cells, however, survive infection and are not lysed. If these cells are isolated and established in pure culture, all their descendants are *resistant* to Tl infection. This might be argued that the mutations responsible for Tl resistance were "induced" by the presence of the Tl viruses, and that, in the absence of the Tl viruses, the mutations would not have occurred. In 1943 Salvador Luria and Max Delbruck elegantly proved that such Tl-resistant cells result from **spontaneous mutation**.

Bacterial cells that bear spontaneous mutations, such as Tl resistance, can be isolated and established independently from the parent strain by means of various selection techniques. As a result, one can now induce and isolate mutations for almost any desired characteristic. Because bacteria and the viruses that infect them are haploid, all mutations are expressed directly in the descendants of mutant cells, adding to the case with which these microorganisms can be studied.

Bacteria are grown in either a liquid culture medium or in a petri dish on a semisolid agar surface. If the nutrient components of the growth medium are very simple and consist only of a organic carbon source (such as a glucose or lactose) and a variety of inorganic ions, including Na^+, K^+, Mg^{++}, and NH_4^+, present as inorganic salts, it is called **minimal medium.** To grow on such a medium, a bacterium must be able to synthesize all essential organic compounds (e.g., amino acids, purines, pyrimidines, sugars, vitamins, and fatty acids). A bacterium that can accomplish this remarkable biosynthetic feat—one that we ourselves cannot duplicate—is termed a **prototroph.** It is said to be wild type for all growth requirements and can grow on minimal medium. On the other hand, if a bacterium loses, through mutation, the ability to synthesize one or more organic components, it is said to be an **auxotroph.** For example, a bacterium that loses the ability to make histidine is designated as a *bis*$^-$ auxotroph, as opposed to its prototrophic bis$^+$ counterpart. For the bis bacterium to grow, this amino acid must be added as a supplement to the minimal medium. The medium that has been extensively supplemented is referred to as complete medium.

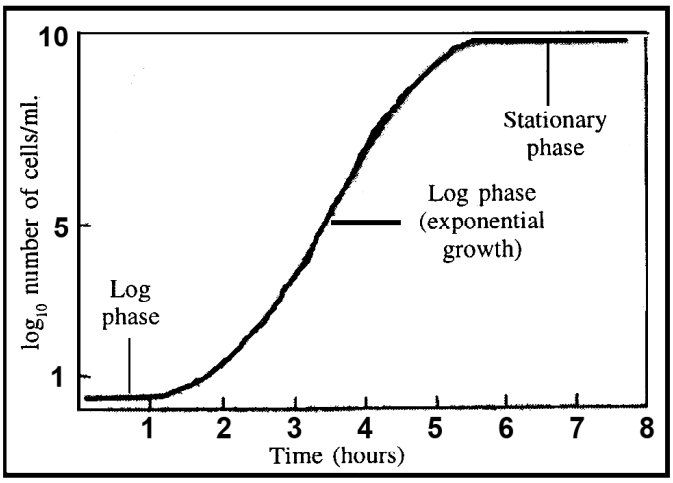


Fig. 7.1 A typical bacterial population growth curve illustrating the initial lag phase, the subsequent log phase where exponential growth occurs, and the stationary phase that occurs when nutrients are exhausted

To study mutant bacteria in a quantitative fashion, an inoculum (e.g., 0.1 ml, 1.0 ml) of bacteria is placed in liquid culture medium. The bacteria exhibit a characteristic growth pattern, as illustrated in Figure 7.1. Initially, during the lag phase, growth is slow. Then, a period of rapid growth follows called the **log phase,** during which cells divide many times with a fixed time interval between cell divisions, resulting in logarithmic growth. When the bacteria reach a cell density of about 10^9 cells per milliliter, nutrients and oxygen become limiting and cells enter the **stationary phase.** As the doubling time during the log phase may be as short as 20 minutes, an initial inoculum of a few thousand cells added to the culture can easily achieve a maximum cell density overnight.

Once cells are grown in liquid medium they can be quantitated. First, the bacteria are plated on (transferred to) semi-solid medium in a petri dish where, following incubation and many divisions, each cell gives rise to a colony visible on the surface of the medium. From the number of colonies that subsequently grow, it is possible to estimate the number of bacteria present in the original culture. If the number of colonies is too great to count, then serial dilutions of the original liquid culture can be made and plated, until the colony number is reduced to the point where it can be counted (Figure 7.2).

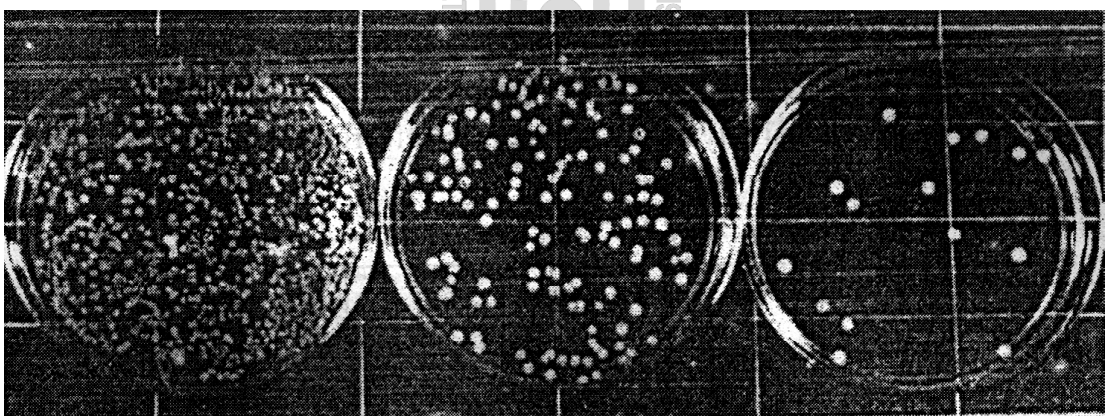


Fig. 7.2 Results of the serial dilution technique and subsequent culture of bacteria. Each of the dilutions varies by a factor of 10. Each colony was derived from a single bacterial cell

For example, assume that the three petri dishes in Figure 7.2 represent dilutions of 10^{-3}, 10^{-4}, and 10^{-5}, respectively (left to right). We select the dish were there are few enough colonies to be accurately counted. Because each colony presumably arose from a single bacterium, the number of a colonies times the dilution factor represents the number of bacteria in the initial milliliter. In this case, the dish farthest to the right contains 15 colonies. Since it represents a dilution of 10^{-5}, we can estimate the initial number of bacteria to be 15×10^5 per milliliter. Calculations such as these are useful in a number of studies.

7.3 Conjugation is one of the methods of genetic recombination in bacteira

7.3.1 Introduction

Development of techniques that allowed the identification and study of bacterial mutations led to detailed investigations of the arrangement of genes on the bacterial chromosome. Such studies began in 1946 when Joshua Lederberg and Edward Tatum showed that bacteria undergo conjugation, a parasexual process in which the genetic information from one bacterium is transferred to and recombined with that of another bacterium. Like meiotic crossing over in eukaryotes, genetic recombination in bacteria provided the basis for the development of methodology for chromosome mapping. Note that the term **genetic recombination,** as applied to bacteria and bacteriophages, leads to the *replacement* of one or more genes present in one strain with those from a genetically distinct strain. While this is somewhat different from our use of genetic recombination in eukaryotes, where the term describes crossing over that results in *reciprocal exchange events,* the overall effect is the same : Genetic informatin is transferred from one chromosome to another, resulting in an altered genotype. Two other phenomena that result in the transfer of genetic information from one bacterium to another, **transformation** and **transduction,** have also served as a basis for determining the arrangement of genes on the bacterial chromosome.

Lederberg and Tatum's initial experiments were performed with two multiple auxotroph strains (nutritional mutants) of *E. coli* K12. As shown in Figure 7.3, Strain A required methionine (met) and biotin (bio) in order to grow, whereas strain B required theonine (thr), leucine (leu), and thiamine (thi). Neither strain would grow on minimal medium. The two strains were first grown separately in supplemented media, and then cells from both were mixed and grown together for several more generations. They were then plated on minimal medium. Any bacterial cells that grew on minimal medium are **prototrophs** (wild-type bacteria that did not need nutritional supplements). It was highly improbable that any of the cells that contained two or three mutant genes would undergo spontaneous mutation simultaneously at two or three gene sites. Therefore, any prototrophs recovered must have arisen as a result of some form of genetic exchange and recombination.

In this experiment, prototrophs were recovered at a rate of $1/10^7$ (10^{-7}) cells plated. The controls for this experiment involved separate plating of cells from strains A and B on minimal medium. No prototrophs were recovered. On the basis of these observations, Lederberg and Tatum proposed that, while the events were indeed quite rare, genetic recombination had occurred.

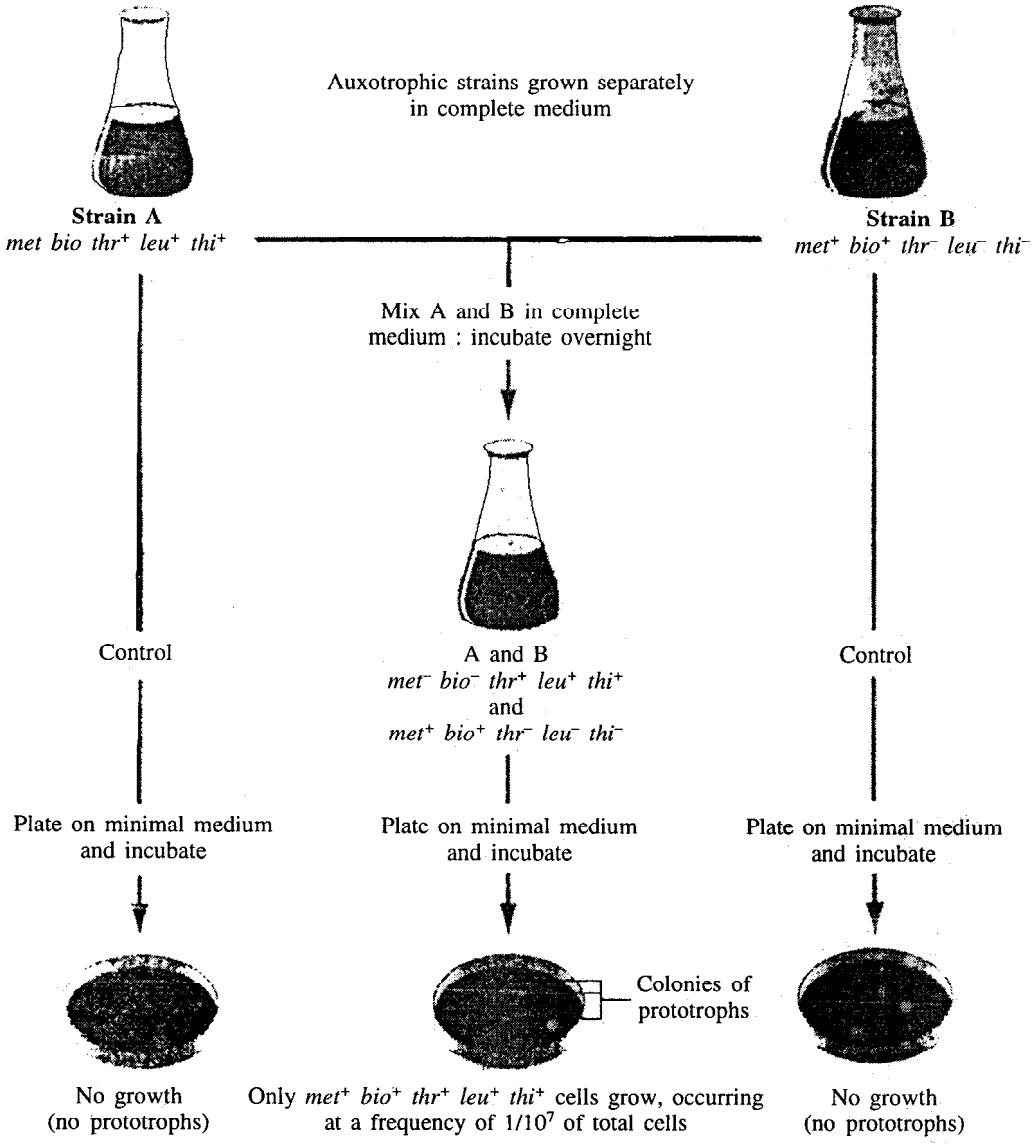


Fig. 7.3 Genetic recombination involving two auxotrophic strains producing prototrophs. Neither auxotrophh will grow on minimal medium, but prototrophs will, suggesting that genetic recombination has occurred

7.3.2 F^+ and F^- Bacterial strains

Lederberg and Tatum's findings were soon followed by numerous experiments designed to elucidate the genetic basis of conjugation. It quickly became evident that different strains of bacteria were involved in a unidirectional

transfer of genetic material. When cells serve as donors of parts of their chromosomes, they are designated as **F^+ cells** (F for "fertility"). Recipient bacteria receive the donor chromosome material (now known to be DNA), and recombine it with part of their own chromosome. They are designated as **F^- cells.**

7.3.3 Conjugation in bacteria

It was established subsequently that cell contact is essential to chromosome transfer. Support for this concept was provided by Bernard Davis, who designed a U-tube in which to grow F^+ and F^- cells (Figure 7.4). At the base of the tube was a glass filter with pores large enough to allow passage of the liquid medium, but too small to allow the passage of bacteria. The F^+ cells were placed on one side of the filter and F" cells on the other side. The medium was moved back and forth across the **filter** so that the bacterial cells essentially shared a common medium during incubation. Samples from both sides of the tube were then plated independently on minimal medium, but no prototrophs were found. Davis concluded that *physical contact is essential to genetic recombination.* Such physical interaction is the initial stage of the process of conjugation and is mediated through a conjugation tube called the F, or sex, pilus. Bacteria often have many pili, which are microscopic tube-like extensions of the cell. Different types of pili perform different cellular functions, but all piii are involved in some way with adhesion. After contact has been initiated between mating pairs via the F pili (Figure 7.5), transfer of DNA begins.

Later evidence established that F^+ cells contained a fertility factor (called the F factor) that confers the ability to donate part of their chromosome during conjugation. Experiments by Joshua and Esther Lederberg and by William Hayes and Luca Cavalli-Sforza showed that certain conditions could eliminate the F factor in otherwise fertile cells. However, if these "infertile" cells were then grown with fertile donor cells, the F factor was regained.

The conclusion that the F factor is a mobile element was further supported by the observation that, following conjugation and genetic recombination, recipient cells always become F^+. Thus, in addition to the *rare* cases of transfer of genes from the bacterial chromosome (genetic recombination), the F factor itself is passed to *all* recipient cells. On this basis, the initial crosses of Lederberg and Tatum (Figure 7.3) may be designated.

STRAIN A		STRAIN B
F^+	X	F^-
DONOR		RECIPIENT

Isolation of the F factor confirmed these conclusions. Like the bacterial chromosome, though distinct from it the F factor has been shown to consist of

a circular, double-stranded DNA molecule, equivalent to about 2 percent of the bacteria] chromosome (about 100,000 nucleotide pairs). Contained in the F factor, among others, are 19 genes, the products of which are involved in the transfer of genetic information (*tra* genes). These include those essential to the formation of the sex pilus.

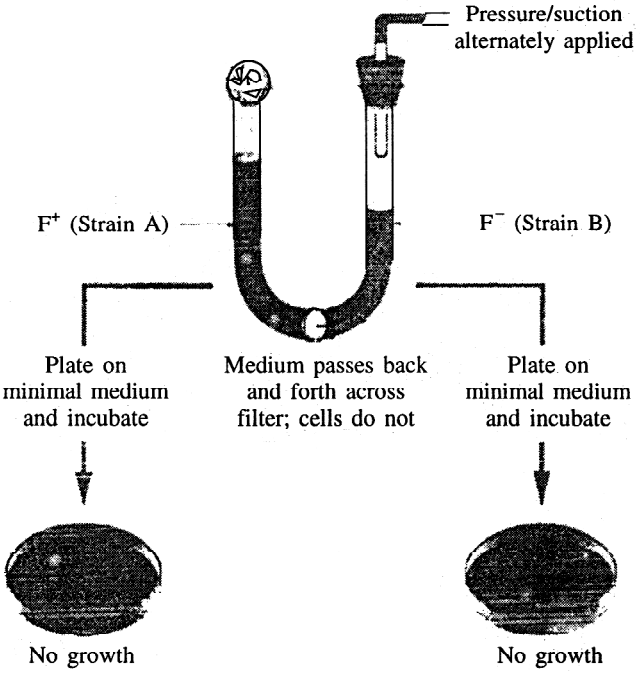


Fig. 7.4 When strain A and B auxotrophs are grown in a common medium but separated by a no genet.c recombination occurs and no prototrophs are produced. This apparatus is called a Davis U-tube

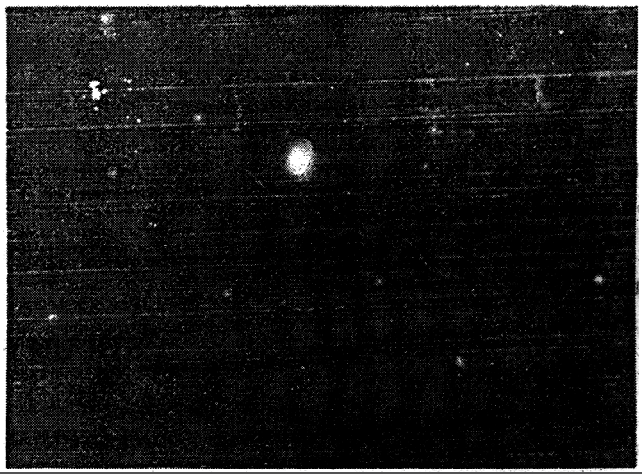


Fig. 7.5 An electron micrograph of conjugation between an F^+ *E coli* cell. The sex pllus linking them is clearly visible

As we soon shall see, the F^- factor is in reality an autonomous genetic unit referred to as a plasmid. However, in our historical coverage of its discovery, we will continue in this chapter to refer to it as a "factor."

It is believed that the transfer of the F factor during conjugation involves separation of two strands of the F factor double helical DNA and the movement of one of the two strands into the recipient. The other strand remains in the donor cell. Both of these parental strands serve as templates for DNA replication, resulting in two intact F factors, one in each of the two cells. Both cells are now F^+ (See in Figure 7.6).

To summarize, *E. coli* cells may or may not contain the F factor. When it is present, the cell is able to form a sex pilus and potentially serve as a donor of genetic information. During conjugation, a copy of the F factor is almost always transferred from the F^+ cell to the F^- recipient, converting it to the F^+ state. The question remains as to exactly how a very low percentage of F^- cells undergo genetic recombina-tion. The answer awaited further experimentation. Subsequent
discoveries not only clarified how genetic recombination occurs but also defined a mechanism by which the *E. coli* chromosome could be mapped.

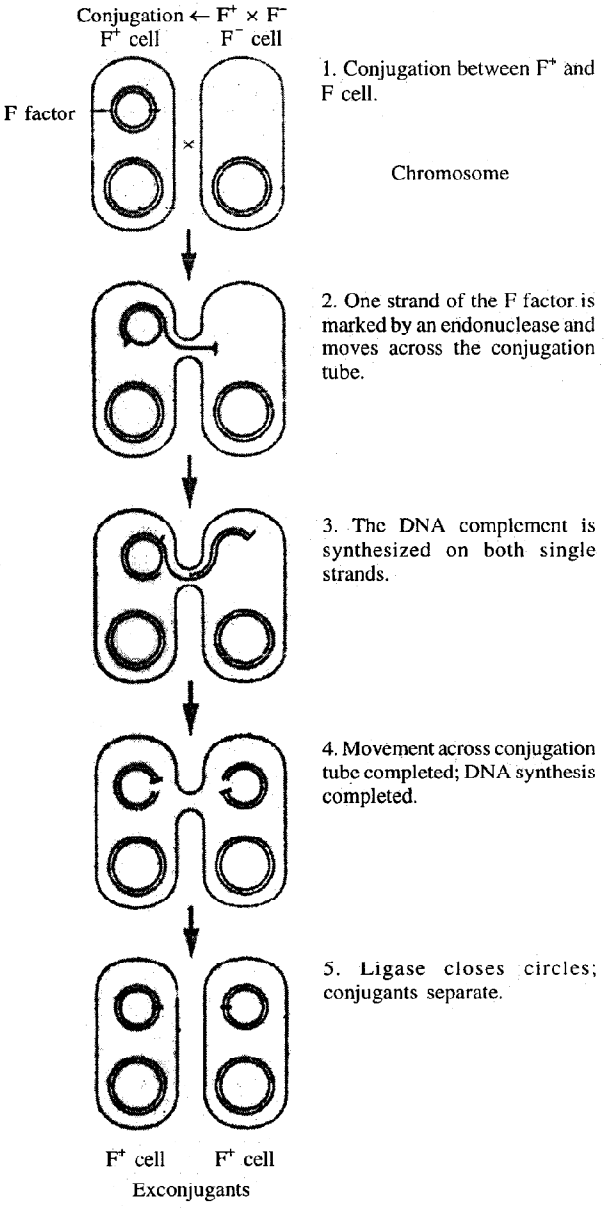


Fig. 7.6 An $F^+ \times F$ mating demonstrating how the recipient F^- cell is converted to F^+. During conjugation, the DNA of the F factor is replicated with one new copy entering the recipient cell converting it to F^+. The black bar has been added to the F factors to follow their clockwise rotation during replication

7.3.4 Hfr bacteria and chromosome mapping

In 1950, Cavalli-Sforza treated an F^+ strain *of E. coli* K12 with nitrogen mustard, a potent chemical known to induce mutations. From these treated cells, he recovered a strain of donor bacteria that underwent recombination at a rate of $1/10^4$ (10^{-4}), 1000 times more frequently than the original F^+ strains. In 1953, William Hayes isolated another strain demonstrating a similar elevated frequency. Both strains were designated **Hfr,** or **high-frequency recombination.** Because Hfr cells behave as chromosome donors, they are a special class of F^+ cells.

Another important difference was noted between Hfr strains and the original F^+ strains. If the donor is an Hfr strain, recipient cells, while sometimes displaying genetic recombination, never become Hfr; that is, they remain F^- In comparison, then,

$F^+ \quad x\ F^- \rightarrow F^+$ (low rate of recombination)
Hfr $x\ F^- \rightarrow F^-$ (higher rate of recombination)

Perhaps the most significant characteristic of Hfr strains is the nature of recombination. In any given strain, certain genes are more frequently recombined

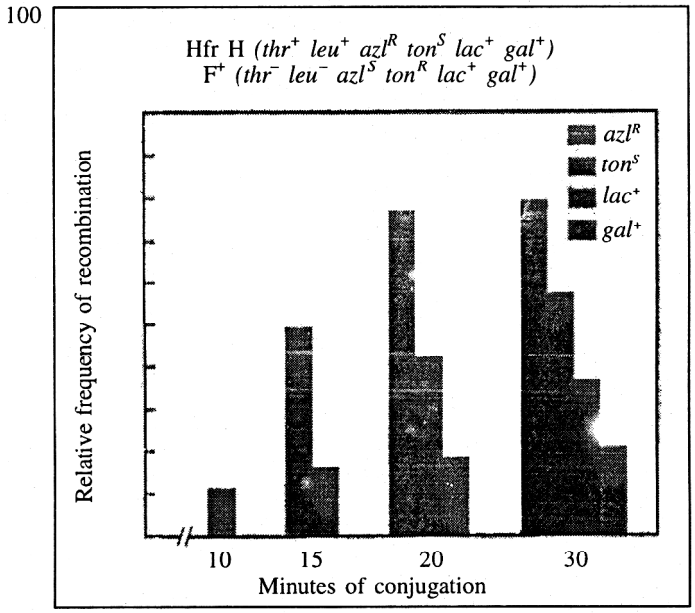


Fig. 7.7 The progressive transfer during conjugation of various genes from a specific Hfr strain of *E, coli* to an F strain. Certain genes *{azi* and *ton)* are transferred sooner than others and recombine more frequently. Others *(lac* and *gal)* take longer to be transferred and recombine with a lower frequency. Others *(thr* and *leu)* are always transferred and are used in the initial screen for recombinants

than others, and some do not recombine at all. This *nonrandom pattern of gene transfer* was shown to vary from Hfr strain to Hfr strain. While these results were puzzling. Hayes interpreted them to mean that some physiological alteration of the F factor had occurred, resulting in the production Hfr strains of *E. coli.*

In the mid-1950s, experimentation by Ellie Wollman and Francois Jacob explained the differences between cells that are Hfr and those that are F^+ and showed how Hfr strains allow genetic mapping of the *E, coli* chromosome. Wollman and Jacob first incubated a culture containing a mixture of an Hfr strain and an F strain. To facilitate the recovery of only recombinants, the Hrf strain was sensitive to an antibiotic while the recipient strain was resistant. At various intervals, the researchers removed samples and placed them in a blender. The shear forces created in the blender seperated conjugating bacteria so that the transfer of the chromosome was effectively terminated. To assay the cells for genetic recombination following the blender treatment, they were grown on medium *containing* the antibiotic in order to ensure the recovery of only recipient cells.

7.3.5 Interrupted mating technique

This process, called the interrupted mating technique, demonstrated that specific genes of a given Hfr strain were transferred and recombined sooner than others. Figure 7.7 illustrates this point. During the first 8 minutes after the two strains were initially mixed, no genetic recombination could be detected. In cells assayed at about 10 minutes, recombination of the azi^R gene could be detected, but no transfer of the ton^S, lac^+, or gat^+ genes was noted. By 15 minutes, 70 percent of the recombinants were azl^R; 30 percent were now also ton^s; but none was lac^+ or *gat.* Within 20 minutes, the lac^+ gene was found among the recombinants; and within 30 minutes, gal^+ was also being transferred. Wollman and Jacob had demonstrated an *oriented transfer of genes* that was correlated with the length of time conjugation was allowed to proceed.

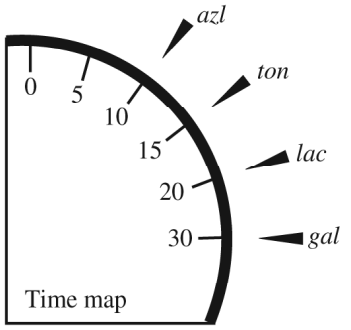


Fig. 7.8 A time map ot the genes studied in the experiment depicted in Figure 7.7

It appeared that the chromosome of the Hfr bacterium was transferred linearly and that the gene order and distance between genes, as measured in minutes, could be predicted from such experiments (Figure 7.8). This information served as the basis for the first genetic map of the *E. coli* chromosome. "Minutes" in bacterial mapping are equivalent to "map units" in eukaryotes.

Wollman and Jacob repeated the same type of experimentation with other Hfr strains, obtaining similar results with one important difference. Although genes were always transferred linearly with time, as in their original experiment, which genes entered first and which followed later seemed to vary from Hfr strain to Hfr strain [Figure 7.9(a)].

When they reexamined the rate of entry of genes, and thus the different genetic maps for each strain, a definite pattern emerged. The major difference between each strain was simply the point of the origin and the direction in which entry proceeded from that point [Figure 7.9(b)].

(a)

Hfr strain	(Ear liest)														(latest)
H	*thr*	–	*leu*	–	*azi*	–	*ton*	–	*pro*	–	*lac*	–	*gal*	–	*thi*
1	*lac*	–	*pro*	–	*ton*	–	*azi*	–	*leu*	–	*thr*	–	*thi*	–	*gal*
2	*pro*	–	*ton*	–	*azi*	–	*leu*	–	*thr*	–	*thi*	–	*gal*	–	*lac*
7	*ton*	–	*azi*	–	*leu*	–	*thr*	–	*thi*	–	*gal*	–	*lac*	–	*pro*

(b)

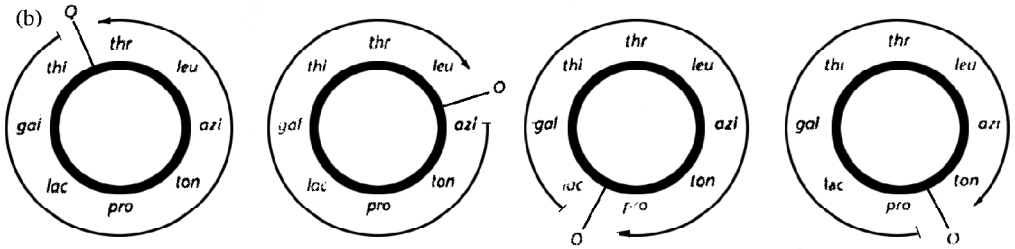


Fig. 7.9 (a) The order of gene transfer in four Hfr strains, suggesting that the *E. coli* chromosome is circular, (b) The point where transfer originates *(O)* is identified in each strain. Note that transfer can proceed in either direction, depending on the strain. The origin is determined by the point of integration into the chromosome of the F factor, and the direction of transfer is determined by the orientation of the F factor as it integrates

To explain these results, Wollman and Jacob postulated that the *E. coli* chromosome is circular. If the point of origin *(O)* varied from strain to strain, a different sequence of genes would be transferred in each case. But what determines *O*? They proposed that in various Hfr strains, the F factor integrates into the chromosome at different points and its position determines the *O* site. One such case of integration is shown in Figure 7.10 (Step 1). During conjugation between this Hfr and an F cell, the position of the F factor determines the initial point of transfer (Step 2 and 3). Those genes adjacent to *O* are transferred first. *The F factor becomes the last part to be transferred* (Step 4). Apparently, conjugation rarely, if ever, lasts long enough to allow the entire chromosome to pass across the conjugation tube (Step 5). This proposal explains why recipient cells, when mated with Hfr cells, remain F^-.

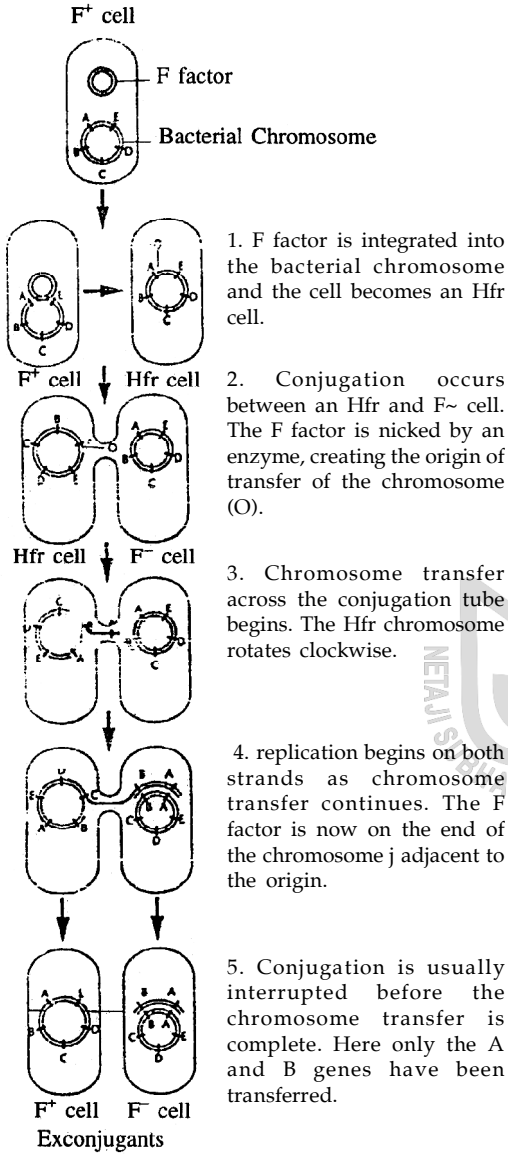


Fig. 7.10 Conversion of $F^{4"}$ to an Hfr state occurs by the integration of the F factor intot he bacterial chromosome. The point of integration determines the origin *(O)* of transfer, during conjugation, the F factor, now integrated into the host chromosome, is nicked by an enzyme, initiating transfer of the chromosome at that point. Conjugation is usually interrupted prior to complete transfer. Above, only the A and B genes are transferred to the F cell, which may recombine with the host chromosome

Figure 7.10 also depicts the way in which the two strands making up a DNA molecule unwind during transfer, allowing for the entry of one of the strands of DNA into the recipient (Step 3). Following replication, the entering DNA now has the potential to recombine with its homologous region of the host chromoso'me. The DNA strand that remains in the donor also undergoes replication.

The use of the interrupted mating technique with different Hfr strains has provided the basis for mapping the entire *E. coli* chromosome, Mapped in time units, strain K12 (or *E. coli* K12) is 100 minutes long. Over 900 genes have now been placed on the map. In most instances, only a single copy of each gene exists.

7.3.6 Recombination in $F^+ \times F^-$ matings : A review

The above model has helped geneticists to better understand how genetic recombination occurs during the F^+xF^- matings. Recall that recombination occurs much less frequently in them than in $Hfr \times F^-$ matings, and that random gene transfer is involved. The current belief is that when F^+ and F^- cells are mixed, conjugation occurs readily and that each F^- cell involved in conjugation with an F^+ cell receives a copy of the F factor, *but that no genetic recombination occurs.* However, at an extremely low frequency in a population of F^+ cells, the F factor integrates spontaneously from the cytoplasm to a random point in the bacterial chromosome, converting the F^+

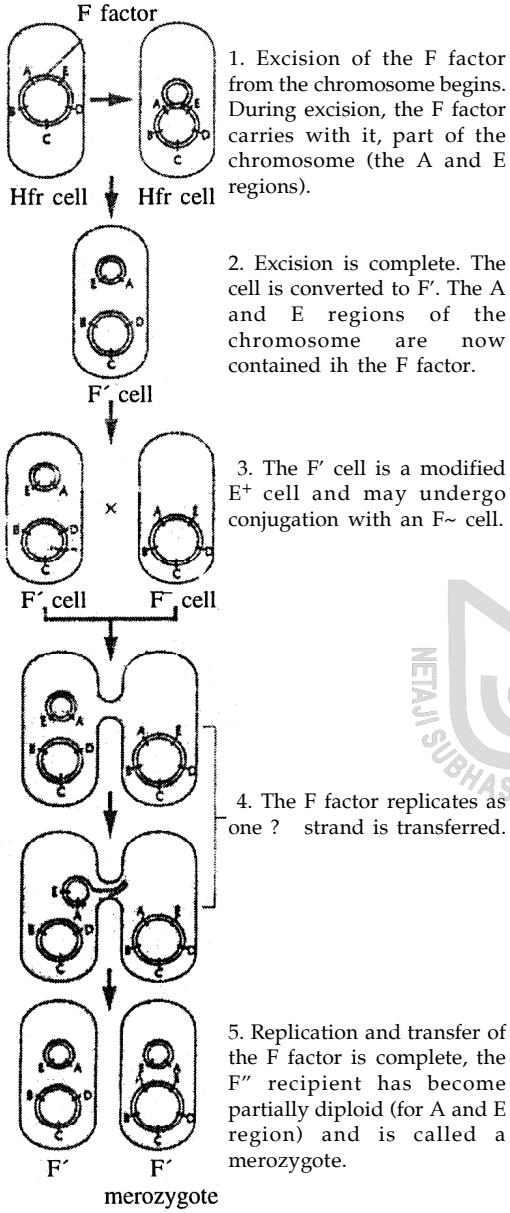


Fig. 7.11 Conversion of an Hfr bacterium to F[1'" and is subsequent mating with an F cell. The conversion occurs when the F factor loses its integrated status. During excision from the chromosome, it carries with it one or more chromosomal genes (A and E) F lowing conjugation with an F~ cell, the recipient cell becomes partially diploid and is called a merozygote. It also behaves as an F~ donor ones

cell to the state as we saw in Figure 7.10 Therefore, in F^+ x F^- crosses, the extremely low frequency of genetic recombination (10^{-7}) is attributed to the rare, newly formed Hfr cells, which then undergo conjugation with F^- cells. Because the point of integration of the F factor is random, a nonspecific gene transfer ensues, leading to the low-frequency, random genetic recombination observed in the F^+ x F^- experiment. Unless the recipient cell simultaneously or subsequently undergoes conjugation with a separate F^+ cell, it will remain F^-. Most often, the recombinants become F^+.

7.3.7 The F′ State and merozygotes

In 1959, during experiments with Hfr strains of *E. coli.* Edward Adelberg discovered that the F^+ factor could lose its integrated status, causing the cell to revert to the F^+ state (Figure 7.11 Step 1). When this occurs, the F factor frequently carries several adjacent bacterial genes along with it (Step 2). Adelberg labeled this condition F′ to distinguish it from F^+ and Hfr. F^+ like Hfr, is thus another special case of F^+. This conversion is described as one from Hfr to F^+.

The presence of bacterial genes within a cytoplasmic F factor creates an interesting situation. An F′ bacterium behaves like an F^+ cell, initiating conjugation with F^- cells (Figure 7.11-Step 3). When this occurs, the F factor, containing chromosomal genes, is transferred to the F^+ cell (Step 4). As a result, whatever chromosomal genes are part of the" F factor are now present in

duplicate in the recipient cell (Step 5), because the recipient still has a complete chromosome. This creates a partially diploid cell called a merozygote. Pure cultures of F′ merozygotes can be established. They have been extremely useful in the study of bacterial genetics, particularly in genetic regulation.

7.3.8 Bacterial recombination is dependent on rec poteins

Once researchers established that a unidirectional transfer of DNA occurs between bacteria, they were interested in determining how the actual recombination event occurs in the recipient cell. Just how does the donor DNA replace the comparable region in the recipient chromosome? As with many systems, the biochemical mechanism by which recombination occurs was deciphered through genetic stuies. Major insights were gained as a result of the isolation of a group of mutations representing genes called *rec.*

The first relevant observation in this case involved a series of mutant genes labeled *recA, recB, recC,* and *recD.* The first mutant gene, *recA,* was found to diminish genetic recombination in bacteria 1000-fold, nearly eliminating it altogether. The other *rec* mutations reduced recombination by about 100 times. Clearly, the normal wild-type products of these genes play some essential role in the process of recombination.

By looking for a functional gene product present in normal cells but missing in mutant cells, researchers subsequently isolated several gene products and showed that they played a role in genetic recombination. The first is called the RecA protein.* The second is a more complex protein called the RecBCD product, an enzyme consisting of polypeptide subunits encoded by three other *rec* genes. The roles of these proteins have now been elucidated *in vitro.* As a result of this genetic research, our knowledge of the process of recombination has been extended considerably. These discoveries underscore the value of isolating mutations, establishing their phenotypes, and determining the biological role of the normal, wildtype gene as a result of subsequent investigation.

7.3.9 F factors are plasmids

In the preceding sections we have examinined the extra-chromosomal heredity unit called the F factor. When it exists autonomously in the bacterial cytoplasm, the F factor is composed of a double-stranded closed circle of DNA [Figure 7.12(a)]. These characteristics place the F factor in the more general category of genetic .structures called plasmids. These structures contain one or more genes—often, quite a few. Their replication depends on the same enzymes that replicate the chromosome of the host cell, and they are distributed to daughter cells along with the host chromosome during cell division.

Plasmids are generally classified according to the genetic information specified by their DNA. The F factor confers fertility and contains genes essential

*Note that the names of bacterial genes begin with lowercase letters and are italicized. The names of the corresponding gene products (proteins) begin with an uppercase letter and are not italicized. For example, the m-4 gene encodes the RecA protein.

for sex pilus formation, upon which genetic recombination depends. Other examples of plasmids include the R and the Col plasmids.

Most R plasmids consist of two components : the RTF (resistance transfer factor) and one or those r-de-terminants {Figure 7.12(b)]. The RTF encodes genetic information essential to transfer of the plasmid between bacteria, and the r-determinants are genes conferring resistance to entibiotics.

The **Col plasmid,** ColEl, derived from *E. coli,* is clearly distinct from R plasmids. It encodes one or more proteins that are highly toxic to bacterial strains that do not harbor the same plasmid. These proteins, called **colicins,** may kill neighboring bacteria. Bacteria that carry the plasmid are said to be colicinogenic. Present in 10 to 20 copies per cell, the plasmid also contains a gene encoding an immunity protein that protects the host cell from the toxin. Unlike an R plasmid, the Col plasmid is not usually transmissible to other cells.

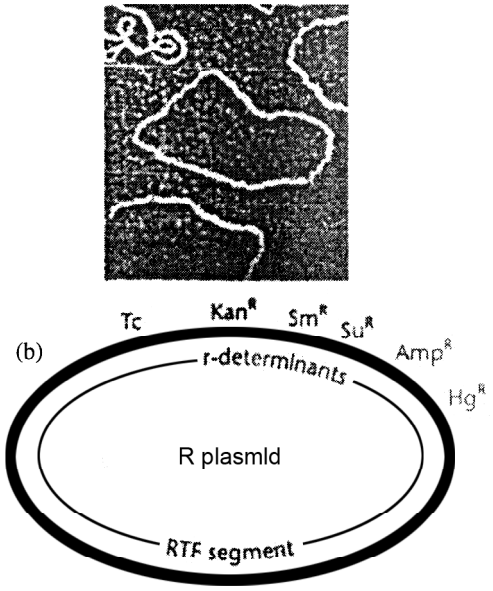


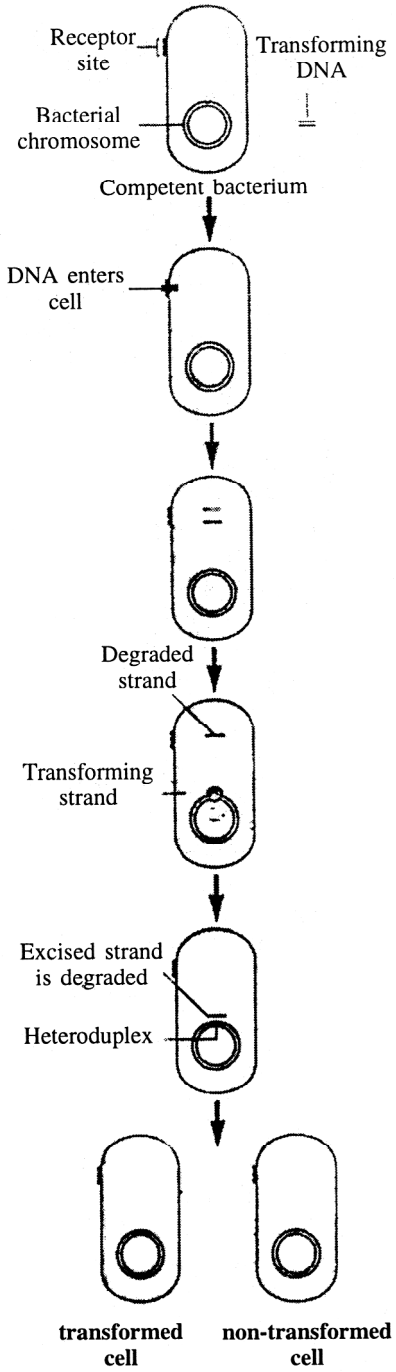
Fig. 7.12 (a) Electron micrograph of a plasmid isolated from *E. coll;* (b) diagramatic representation of an R plasmid containing resistance transfer factors (RTFs) and multiple r-determinance (Tc, tetracycline; Kan, kanamycin; Sm, streptomycin; Su, sulfonamide; Amp, ampicillian; and Hg, mercury)

7.4 Transformation is another process leading to genetic recombination in bacteria

Transformation is another process that provides a mechanism for the recombination of genetic information in some bacteria. In transformation, small pieces of extracel Iular DNA are taken up by a living bacterium, ultimately leading to a stable genetic change in the recipient cell.

7.4.1 The Transformation Process

Transformation (Figure 7.13) consists of numerous steps that can be divided into two main categories: (1) entry of DNA into a recipient cell, and (2) recombination of the donor DNA with its homologous region in the recipient chromosome. In a population of cells, only those in a particular physiological state, referred to as competence, take up DNA. Entry is thought to occur at a limited number of receptor sites on the surface of the bacterial cell. Passage



1. Extracellular DNA binds to the competent cell at a receptor site.

2. DNA enters the cell and strands separate.

3. One strand of transforming DNA is degraded. The other and pairs homologously with the host cell DNA.

4. The transforming DNA is integrated, forming a heteroduplex.

5. Following one round of cell division, a transformed and a non-transformed cell are produced.

Fig. 7.13 Proposed steps leading to transformation of a bacterial cell by exogenous DNA. Only one of the two strands of the entering DNA is involved in the transformation event, which is completed following cell division

across the cell wall and membrane is an active process requiring energy and specific transport molecules. This concept is supported by the fact that substances that inhibit energy production or protein synthesis in the recipient cell also inhibit the transformation process.

During the process of entry, one of the two strands of the invading DNA molecule is digested by nucleases, leaving only a single strand to participate in transformation (Step 2 and 3). The surviving DNA strand aligns with its complementary region of the bacterial chromosome. In a process involving several enzymes, this segment of DNA replaces its counterpart in the chromosome, which is excised and degraded (Step 4).

For recombination to be detected, the transforming DNA must be derived from a different strain of bacteria, bearing some genetic variation. Once integrated into the chromosome, the recombinant region contains one DNA strand from the bacterial chromosome and one from the transforming DNA. Because these strands are not genetically identical, this helical region is referred to as a **heteroduplex.** Following one round of replication, one chromosome is restored to its original configuration, identical to that of the recipient cell, and the other contains the transformed gene. Cell division produces one host cell and one transformed cell (Step 5).

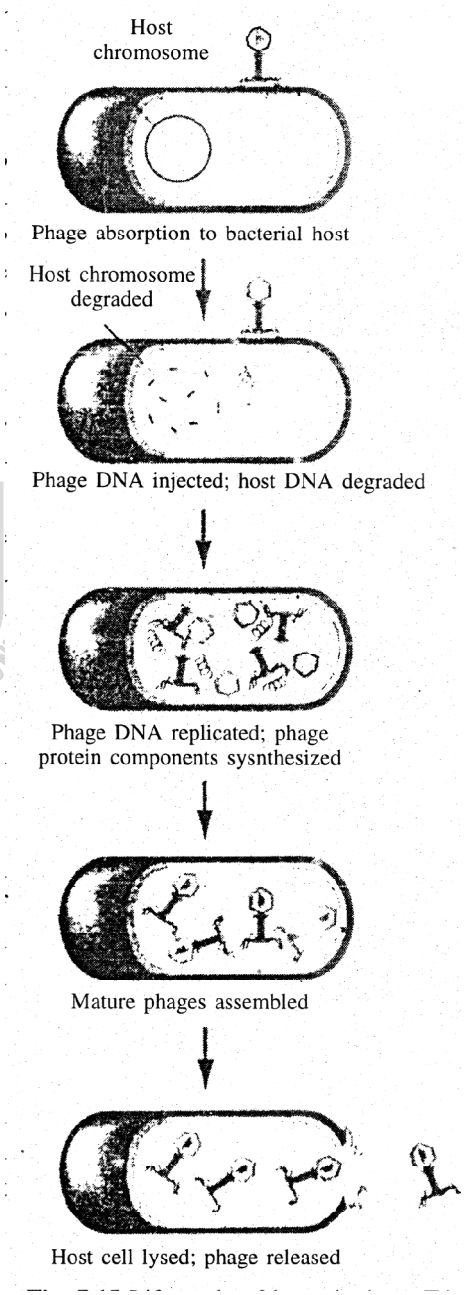


Fig. 7.15 Life cycle of bacteriophage T4

7.4.2 Lysogeny

The relationship between virus and bacterium does not always result in viral reproduction and lysis. As early as the 1920s, it was known that some bacteriophages could enter a bacterial cell and establish a symbiotic relationship with it. The precise molecular basis of this symbiosis is now well understood.

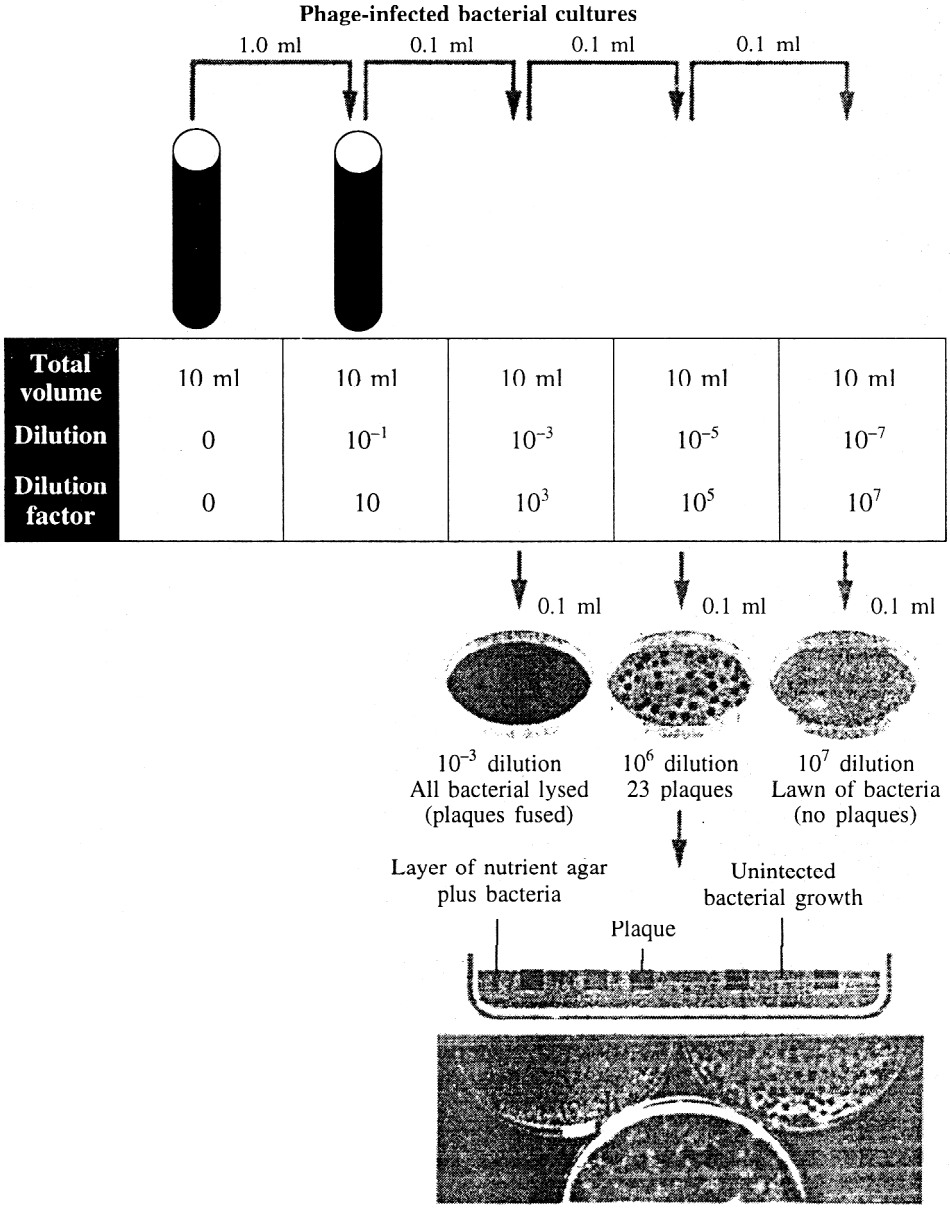


Fig. 7.16 Diagramatic illustration of the plaque assay for bacteriophage analysis. Serial dilutions of a bacterial culture infected with bacteriophages are first made. Then three of the dilutions (10^3, 10^6 and 10^7) are analyzed using the plaque assay technique in each case 0.1 ml of the diluted culture is used. Each plaque represents the initial infection of one bacterial cell by one bacteriophage. In the 10^3 dilution, so many phaes are present that all bacterial are lysed. In the 10^{-5} dilution 23 plaques are produced. In the 10^{-7} dilution, the dilution factor is so great that no phages are present in the 0.1 ml sample, and thus no plaques form. From the 0.1 ml sample of the 10^{-3} dilution, the original bacteriophage desnity can be calculated as 23, 10, 10^5 phages/ml (23.10^4 or 23.10). The photograph illustrates phage T2 plaques on lawns of *E. coli*

Upon entry, the viral DNA, instead of replicating in the bacterial cytoplasm, is integrated into the bacterial chromosome, step that characterizes the developmental stage referred to as lysogeny. Subsequenly, each time the bacterial chromosome is replicated, the viral DNA is also replicated and passed to daughter bacterial cells following division. No new viruses are produced and no lysis of the bacterial cell occurs. However, in response to certain stimuli, such as chemical or ultraviolet-light treatment, the viral DNA may lose its integrated status and initiate replication, phage reproduction, and lysis of the bacterium. (Fig, 7.15)

Several terms are used to describe this relationship. The viral DNA integrated into the bacterial chromosome is called a **prophage.** Viruses that can either lyse the cell or behave as a prophage are called **temperate.** Those that can only lyse the cell are referred to as virulent. A bacterium harboring a prophage has been lysogenized and said to be lysogenic; that is, it is capable of being lysed as a result of induced viral reproduction. The viral DNA, which can replicate either in the bacterial cytoplasm or as part of the bacterial chromosome, is sometimes classified as an **episome.**

7.5 Transduction is virus-mediated bacterial DNA transfer

In 1952, Norton Zinder and Joshua Lederberg were investigating possible recombination in the bacterium *Salmonella typhimurium.* Although they recovered prototrophs from mixed cultures of two different auxotrophic strains, subsequent investigations revealed that recombination was occurring in a manner different from that attributable to the presence of an F factor, as in *E. coll* What they discovered was a process of bacterial recombination mediated by bacteriophages and now called transduction.

7.5.1 The Lederberg-Zinder experiment

Lederberg and Zinder mixed the *Salmonella* auxotrophic strains LA-22 and LA-2 together and, when the mixture was plated on minimal medium, they recovered prototroph cells. LA-22 was unable to synthesize the amino acids phenylalanine and tryptophan *(phe^- trp^-),* and LA-2 could not synthesize the amino acids methionine and histidine (met^- his^-). Prototrophs (*phe' trp' met' his')* were recovered at a rate of about $1/10^5$ (10^{-5}) cells.

Although these observations at first suggested that the recombination involved was the type observed earlier in conjugative strains of *E. coli,* experiments using the Davis U-tube soon showed otherwise (Figure 7.17). The two auxotrophic strains were separated by a glass-smtered filter, thus preventing cell contact but allowing growth to occur in a common medium. Surprisingly, when samples were removed-foam both sides of the filter and plated independently on minimal medium, prototrophs were recovered only from the side of the tube containing LA-22 bacteria.

Since LA-2 cells appeared to be the source of the new genetic information (*phe*$^+$ and *trp*$^+$), how that information crossed the filter from the LA-2 cells to the LA-22 cells allowing recombination to occur, was a mystery. The unknown source was designated simply as a **filterable agent (FA).**

Three subsequent observations were useful in identifying the FA :

1. The FA was produced by the LA-2 cells only when they were grown in association with LA-22 cells. If LA-2 cells were grown independently and that culture medium was then added to LA-22 cells, recombination did not occur. Therefore, LA-22 cells play some role in the production of FA by LA-2 cells and do so only when the two share common growth medium.
2. The presence of DNase, which enzymatically digests DNA, did not render the FA ineffective. Therefore, the FA is not naked DNA, ruling out transformation.
3. The FA could not pass across the filter of the Davis U-tube when the pore size was reduced below the size of bacteriophages.

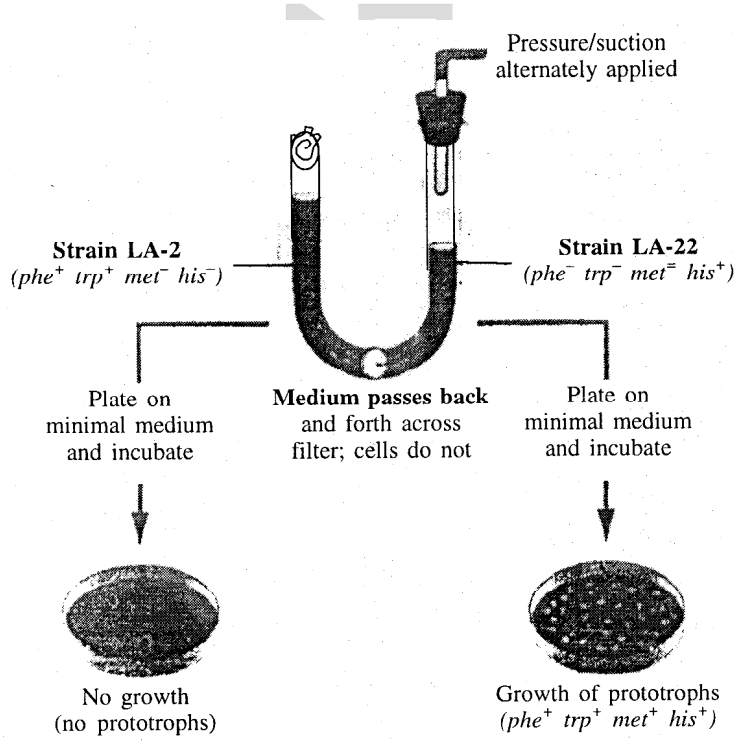


Fig. 7.17 The Lederberg-Zinder experiment using *Solmonelid*. After placing two auxotrophic strains on opposite sides of a Davis U-tube. Lederberg and Zinder recovered prototrophs from the side containing the LA-22 strain but not from the side containing the LA-2 strain. These initial observations led to the discovery of the phenomenon called transduction

Added by these observations and awareof temperate phages that could lysogenize *Salmonella,* researchers proposed that the genetic recombination event was mediated by bacteriophage P22, present initially as a prophase in the chromosome of the LA-22 *Salmonella* cells. It was hypothesized that rarely P22 prophages migt enter the vegetative or Iytic phase, reproduce, and be released by the LA-22 cells. Such phages, being much smaller than a bacterium, were then able to cross the filter of the U-tube and subsequently infect and lyse some of the LA-2 cells. In the process of lysis of LA-2, the P22 phages occasionally packaged in their heads a region of the LA-2 chromosome. If this region contained the *phe*$^+$ and *trp*$^+$ genes, and if the phages subsequenly passed back across the filter and infected LA-22 cells, these newly lysogenized cells would behave as prototrophs. This process of transduction, whereby bacterial recombination is mediated by bacteriophage P22, is diagrammed in Figure 7.18.

7.5.2 The nature of transduction

Further studies revealed the existence of transducing phages in other species of bacteria. For example, *E. coli* can be transduced by phages PI and X. *Bacillus* ***subtilis*** and *Pseudomonas aeruginosa* can be transduced by the phages SPO1 and Fl 16, respectively. The details of several different modes of transduction have also been established. Even though the initial discovery of transduction involved a temperate phage and a lysogenized bacterium, the same process can occur during the normal Iytic cycle. Sometimes a small piece of bacterial DNA is packaged

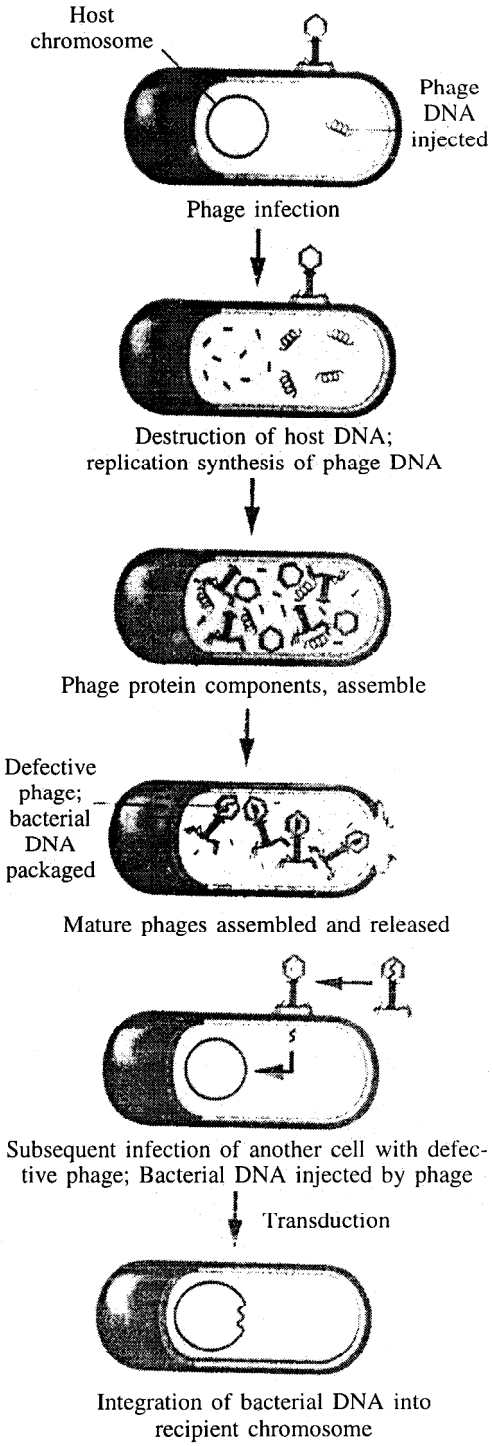


Fig. 7.18 Generalized transduction

along with the viral chromosome so that the transducing phage contains both viral and bacterial DNA. In such cases, only a few bacterial genes are present in the transducing phage. However, when *only* bacterial DNA is packaged, regions as large as 1 percent of the bacterial chromosome may become enclosed in the viral head. In either case, the ability to infect is unrelated to the type of DNA in the phage head, making transduction possible.

When bacterial rather than viral DNA is injected into the bacterium, it can either remain in the cytoplasm of recombine with the homologous region of the bacterial chromosome. If the bacterial DNA remains in the cytoplasm in the cytoplasm, it does not replicate but is transmitted to one of the progeny cells following each division. When this happens, only a single cell, partially diploid for the transduced genes, is produced—a phenomenon called **abortive transduction.** If the bacterial DNA recombines with its homologous region of

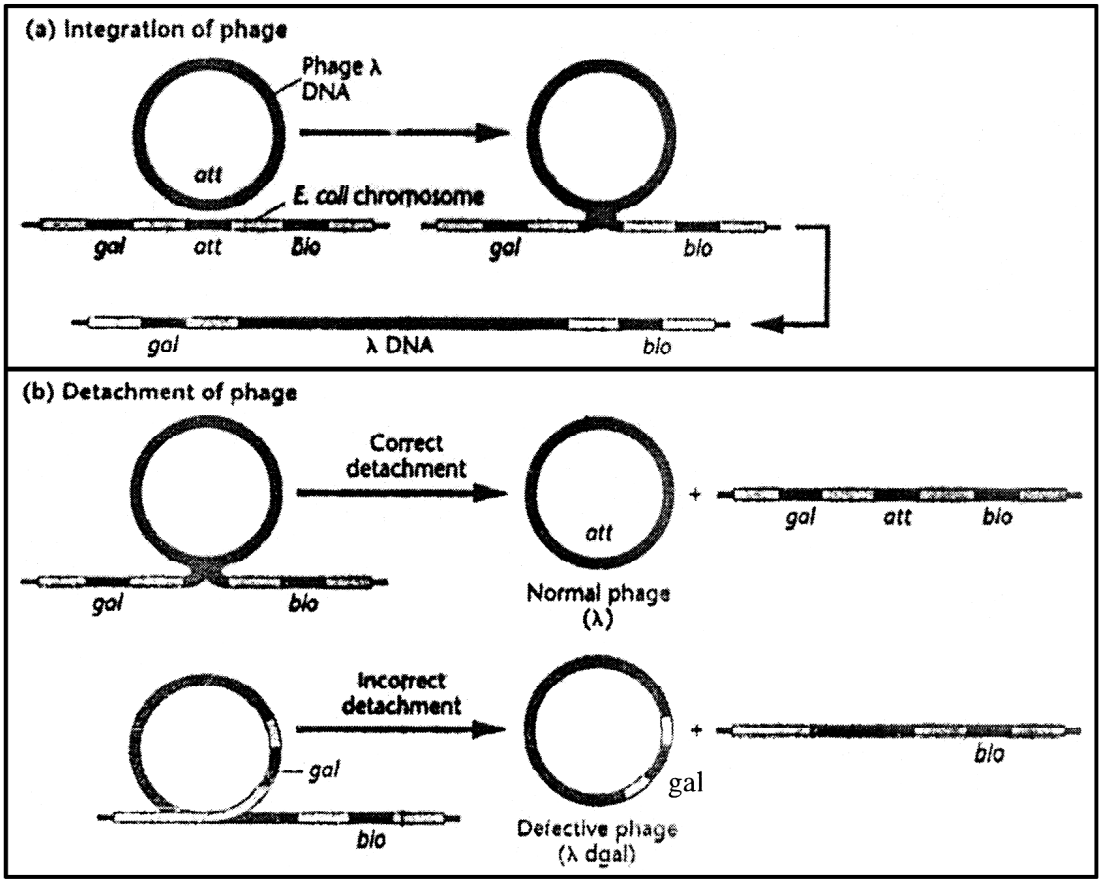


Fig. 7.19 The production of defective phage λ*dgal,* which can result in specialized transduction following another round of infection of *E. coli.* If detachment occurs correctly, no transduction will result

the bacterial chromosome, the transduced genes are replicated as part of the chromosome and passed to all daughter cells. This process is called **complete transduction.**

7.5.3 Transductionand mapping

Like transformation, generalized transduction has been used in linkage and mapping studies of the bacterial chromosome. The fragment of bacterial DNA involved in a transduction event is large enough to include numerous genes. As a result, two genes that are closely aligned (linked) on the bacterial chromosome may be simultaneously transduced, a process called contrasduction. Two genes that are not close enough to one another along the chromosome to be included on a single DNA fragment require two independent events in order to be transdaced into a single cell. Since this occurs with a much lower probability than cotransdoction, linkage can be determined.

By concentrating on two or three linked genes, transduction studies can also determine the precise order of these genes. The closer to each other linked genes are, the greater the frequency of contransduction. Mapping studies involving three closely aligned genes can be executed. The analysis of such an experiment is predicated on the same rationale underlying other mapping techniques.

7.5.9 Specialized transduction

In some instances, only certain genes are recombined, a situation called **specialized transduction.** This is in contrast to generalized transduction described above, where all genes have an equal probability of being recombined. One of the best examples involves transduction of *E. coli* by the temperate phage λ.

In this case, transduction is restricted to the *gal* (galactose) or *bio* **(biotin)** genes. The reason why transduction in volves only these genes became clear when it was learned that the λ DNA'always integrates into the region of the *E. coli* chromosome between these two genes at a site called *att* [Figure 7.19(a)]. Phage λ DNA that is integrated in the bacterial chromosome can subsequenly detach from it, reproduce, and lyse the host cell [Figure 7.19(b)]. Sometimes the excision process occurs incorrectly and carries either the *gal* or *bio E, coli* genes in place of part of the viral DNA [Figure **7.19(b)].** The resulting phage chromosome is defective because it has lost some of its own genetic information, but it is nevertheless replicated and packaged during the formation of mature phage particles. The virus can subsequently inject the defective chromosome into another bacterial cell.

In this case of specialized transduction, the defective phage chromosome

is integrated into the bacterial, chromosome and is replicated along with it. Such bacterial cells contain the defective phage DNA, making them diploid for the *gal* or *bio* genes. The presence of this transducing *gal′* or *bio′* DNA causes these auxotrophs to revert to a *gal′* or *bio′* phenotype.

7.5.5 Bacteriophages undergo intergenic recombination

Around 1947, several research teams demonstrated that genetic recombination also occurs in bacteriophages. These studies relied on the discovery of numerous phage mutations that could be visualized or assayed. Before considering recombination in these bacterial viruses, we will briefly introduce several of the mutations that were studied.

Phage mutations often affect the morphology of the plaques formed following lysis of bacterial cells. For example, in 1946 Alfred Hershey observed unusual T2 plaques on plates of *E. coli* strain B. Where the normal T2 plaques are small and have a clear center surrounded by a diffuse (nearly invisible) halo, the unusual plaques were larger and possessed a more distinctive outer perimeter (Figure 7.20). When the viruses were isolated from these plaques and replated on *E. coli* B cells, the resulting plaque appearance was identical. Thus, the plaque phenotype was an inherited trait resulting from the reproduction of mutant phages. Hershey named the mutant *raid lysis* (r) because the plaques were larger, apparently resulting from a more rapid or more efficient life cycle of the phage. It is now known that in wild-type phages, reproduction is inhibited once a particular-sized plaque has been formed. The mutant T2 phages are able to overcome this inhibition, producing larger plaques.

Unit 8 Cytogenetic effects of Ionizing and Non-ionizing Radiations

Structure

8.1 Introduction

Various agents shown to cause chromosome breaks have been termed "clastogens" by Shaw (1970). These include physical agents (X-rays, ultraviolet light, cold shock, magnetic fields, and sound waves), biological agents (certain genes, viruses and protozoa), and a host of chemical agents. It should be emphasized that most of these clastogens produced these effects *in vitro* by the addition of the agent to cultured lymphocytes and/or fibroblasts for varying times and concentration. In but few cases is there evidence for *in vivo* chromosome breakage.

8.2 Radiation

Survivors of the atomic bomb blasts in Japan have developed leukemia in proportion to the amount of radiation received. Furthermore, increased numbers of chromosome breaks and rearrangements have been found in lymphocytes of nonleukemic survivors (Bloom *et al.,* 1967). Similar anomalies (translocations and inversions) have been demonstrated in lymphocytes of individuals who have received X-ray therapy to the spine or injections of thorotrast (Buckton *et al.,* 1982; Court Brown *et al.,* 1967).

Although ultrasound can effect chromosomal breaks *in vitro,* there is no evidence that it does *so in vivo* (Macintosh and Davey, 1972).

Fibroblasts cultured from skin in the path of X-radiation have manifested chromosome abnormalities (Engel *et al.,* 1964; Visfeldt, 1966). Marrow cells may exhibit abnormalities even after many years following primary exposure (Goh, 1971). Leukemia is also more likely to develop in individuals who have received chronic exposure to radiation (Lewis, 1970). Maternal irradiation before and during the reproductive period increases the incidence of chromosomally abnormal conceptuses. However, most are nonviable and lost early in pregnancy (Alberman *et al,* 1972).

8.3 Viruses

There is insufficient evidence currently available to directly implicate viruses in effecting human chromosome abnormalities. However, a number of investigators have studied the effect of S V40 virus on cultured human fibroblasts and have observed altered growth patterns to the haphazard growth and alterations, within several months an emergence of a few stable heteropliod cells becomes evident (Moorhead, 1970).

Numerous other viruses (Rous sarcoma, vaccinia, rubella, herpes zoster, poliomyelities, influenza, polyoma, etc.) have been shown to produce chromosome breaks in infected cells *in vitro.* Furthermore, several viruses can produce abnormalities in metaphase chromosome in circulating lymphocytes during natural human infections (measles, chicken pox, mumps, and hepatities) (Moorhead, 1970). The effects have been of at least three types: single breaks, pulverization, and fusion and spindle abnormalities. The mechanism is unknown but may be related to addition of the virus like particles (Epstein-Barr virus) have been demonstrated, often exhibit a long submetacentric marker (Gripenberg *et al.,* 1969).

8.4 Chemical clastogens

Over 200 drugs or chemical shown to cause chromosome breaks *in vitro* (shaw, 1970) can be grouped into several categories with a few illustrations in each group: (a) nucleic acid related compounds (6-mercaptopurine and 5-fluorodeoxyuridine), (b) antibiotics (mitomycin C, streptomycin, actinomycin D and daunomycin), (c) central nervous saystem drugs (meprobamate, chlopromazine, mesealine, lysergic acid diethylamide, and scopolamine), (d) food derivatives and additives (caffeine, cyclamate, theobromine, and theophylline), (e) air and water pollutants (chloramine T and ozone), (f) pesticides (captan and thioTEPA), (g) alkylating agents (nitrogen mustards and cytoxan), (h) mitotic poisons (colchicine), (i) photodynamic dyes (acridine orange and neutral red), (j) antifolic compounds (methotrexate and aminopterin), (k) organic solvents (benzene and mercaptoethanol), (1) inorganic substances (lead and arsenic), and (m) miscellaneous comopounds (Imuran and piperazine).

However, it should be emphasized that few chemical clastogens have been implicated in chromosomal breakage *in vivo.* To cite but a few examples: Ambient exposure to benzene has been noted to be associated with both chromosome breakage and subsequent development of leukemia (Tough and Court Brown, 1965; Hartwich *et al.,* 1969). On the other hand, LSD, while producing chromosome breaks *in vitro,* has not been shown to be effective *in vivo* (Stenchever and Jarvis, 1970).

Unit 9 Molecular Cytogenetic Techniques

Structure

9.1 FISH, GISH

9.1.1 *In situ* hybridization technique

Under normal temperature and ionic conditions DNA remains in a duplex state by the base pairing through the hydrogen bonds. By heating in buffer solution or by increasing pH the two strands can be separated. But if again the temperature is lowered or pH is reduced then the separated strands will join again and reassociate.

This fact was shown by Julius Marmur and Paul Doty in 1960. This type of reassociation of DNA strands in called molecular hybridization or nuclear hybridization. It may also take place between the complementary strands of DNA or RNA or between DNA and RNA.

Hybridization technique which can be used to localize specific nucleic acid fragments that reside in their original site *(in situ)* within cells, is known as *in situ* hybridization.

In situ hybridization is a version of hybridization analysis, in which an intact chromosome is examined by probing it with a labeled DNA molecule.

For this technique to work, DNA in the chromosome must be made single stranded and denatured by breaking the hydrogen bonds between the base pairs. Only then the DNA or RNA probe can be hybridizid with the chromosomal DNA. For this purpose, without destroying the chromosome morphology a dry preparation is made into a glass microscope-slide and then treated with formaldehyde.

The position where the hybridization occurs provides the information about the map location of this gene, thus physical mapping of chromosome could be done by this method.

To locate the region of hybridization two types of markers viz., radioactive marker and fluorescent marker are used. Besides these other types of markers are also used for this purpose.

To prepare a DNA profile, the nucleotides are synthesized in which one of the phosphorus atoms replaced by 32_P or 33_P. One of the oxygen atom in the phosphate group is replaced with 35_S or one or more of the hydrogen atom is replaced with 3_H. Radioactive nucleotides act as substrates for DNA polymerases and so are incorporated into a DNA molecule by any strand synthesis reaction catalyzed by a DNA polymerase.

The labeled nucleotides or individual phosphate groups can also be attached to one or both ends of a DNA molecule by the reaction catalyzed by T4 polynucleotide kinase or terminal deoxy nucleotidyl transferase.

The rodioactive signal can be detected by scintillation counting, but for most molecular biology, the position of the hybridization is detected by exposure of an X-ray sensitive film (autoradiography) or radiation sensitive phosphorescent screen (phosphorus imaging). 32_P has a high emission energy and the resolution is lower. But low emission isotopes such as 35_S or 3_H give less sensitivity but greater resolution.

9.1.2 Fluorescent *in situ* hybridization (FISH)

Concept

To solve these problems in 1980's non radioactive fluorescent DNA labels were developed. These labals combine high sensitivity with high resolution and are ideal for *in situ* hybridization. The different fluorolabels of different colours have been disigned for the probes and it is possitive to hybridize the chromosome and the individual hybridization signals enable the location of relative position of the probe sequence to be mapped.

Fluorescent dyes (Fluorochromes) live quinacrine (Q) and quinacrine mustard (QM) are used to obtain specific patterns of cross striations or bands appear with alternate fluorescent and non fluorescent bands. The bands obtained with quinacrine are called Q bands while those obtained with quinacrine mustard are called Q M-bands. Bands are obtained when fluorescent deyes attach to the specefic regions of the chromosomes (See Fig. 9.2).

Technique

During the *in situ* hybridization a sample of dividing cells is dried on a microscope slide and treated with formaldelyde so that the chromosomes become denatured but do not lose their characteristic metaphase morphologies. The probe is added to the denatured chromosomes, which will be hybridized with the complementary DNA region of the chromosome. The position at which the probe hybridizes to the chromosomal DNA is visualized by detecting the fluorescent signal emitted by the labeled DNA.

Problem

If a probe be a long fragment of DNA, then a potential problem is that it is likely to contain repetitive DNA sequences and so may hybridize many portions of the chromosome. If these sequences are not blocked then the probe will hybridize non specifically to any copy of these genome and will repeat in the target DNA. To block the repeat sequences, the probe is pro-hybridized with a DNA fraction enriched for repetitive DNA.

Advanteage of FISH over other methods

Health and environmental issues have meant that radioactive markars have become less popular in recent years and they are now largely superseded by non-radioactive alternatives.

Other drawbacks of radioactive markers in *in situ* hybridization and that the radioactive label has high emission energy (eg., P^{32}) and then it scatters its signal and so gives poor resolution. On the other hand , if H^3 is used, the emission energy is less but its sensitively is so low that lengthy exposures are needed.

Disadvantage

The metaphase chromosomes are highly condensed and a fluorescent signal obtained by FISH is marked by measuring its position relative to the end of the short arm of the chromosome (the flepter value). The two markers having at least 1Mb apart to be resolved as separate hybridization signal (Trask *et al*, 1991). Therefore, using FISH, the highly condensed nature of mataphase chromosomes means that only low resolution mapping is possible. Therefore, FISH provides a fough idea of its map position.

More advanced FISH

In 1996 Heiskanen *et al* solved the problem of FISH and a range of higher resolution FISH technique has been developed. Higher resolution is achieved by changing the condensing pattern of the metaphase chromosome. There are two ways of doing this.

(1) Mechanically stretching of metaphase chromosomes : Centrifugation generates shear forces, which can result in the chromosomes becoming stretched upto 20 times of the normal length. Thus the resolution is significantly improved and markers that are only 200-300 Kb apart can be distinguished.

(2) Taking non-metaplase chromosomes : Attempts have been made to use prophase nuclei because in this stage the chromosomes are not still sufficiently condensed for individual ones to be identified and so provides no advantage.

Interphase chromosomes become more useful because, then the

chromosomes again become less condensed. Using anaphase stage the resolution down to 25 kb is possible but their chromosome morphology is lost, so there are no external reference point against which the position of the probe to be mapped. This technique is used to construct map in small region of the chromosome after obtaining preliminary information.

Fibre-FISH

Interphase chromosomes are most unpacked of all cellular DNA. To improve the resolution of FISH better than 25 kb, it is therefore, necessary to abandon intact chromosomes. This approach is called fiber FISH. In this approach the DNA is prepared by gel stretching or molecular combing. This can distingish markers that are less than 10 Kb.

To carry out gel stretching (Fig. 9.1a), molten agarose containing chromosomal DNA molecules is pipetted into a microscope slide, coated "with a restriction enzyme (Schwartz *et al,* 1993). As the gel solidifies, the DNA molecules become stretched. It is not understood why this happens but it is thought that fluid movement on the glass surface during gelation might be responsible. Once the gel is solidified it is washed with $MgCl_2$ solution, which activates the restriction enzyme. A fluorecent dye such as DAPI (4,6 diamine 2-phenylindolo dihydrochloride) is added which stains the DNA so that the fibres can be seen when the slide in examined with a high power fluorescence microscope. The restriction enzyme cuts the DNA molecule. As the molecules gradualy coil up, the gaps representing the cut sites become visible. The relative positions of the cuts are to be recorded.

In molecular combing (Michalet *et al,* 1997) (Fig. 9.1b), the DNA fibres are prepared by dipping a silicon-coated coverslip into a solution of DNA. It takes 5 minutes to attach DNA to the coverslip by their ends. After that coverslip is removed at a constant speed of 0.3 mm/sec. The force required to pull the DNA molecule through the miniscus causes to line up. Once in the air the surface of the coverslip dries, DNA molecules are arranged as a parallel fibre thus producing a comb of parallel molecules.

9.1.3 Biotin-labelling *in situ* hybridization

Recent advances in nucleic acid technology offer alternative to radioactive-labelling probes. One procedure that is becoming increasingly popular is biotin-labelling of nucleic acid. This is nontoxic, whose half life is longer and can be prepared in adavnce in bulk and stored at -20°C for (repeated) use.

Drosophila salivary gland chromosomes can be hybridized with a biotin labelled nucleic acid probe. After washing, detection can be done by adding a biotin-binding protein called ovidin which is covalently bound to alkaline

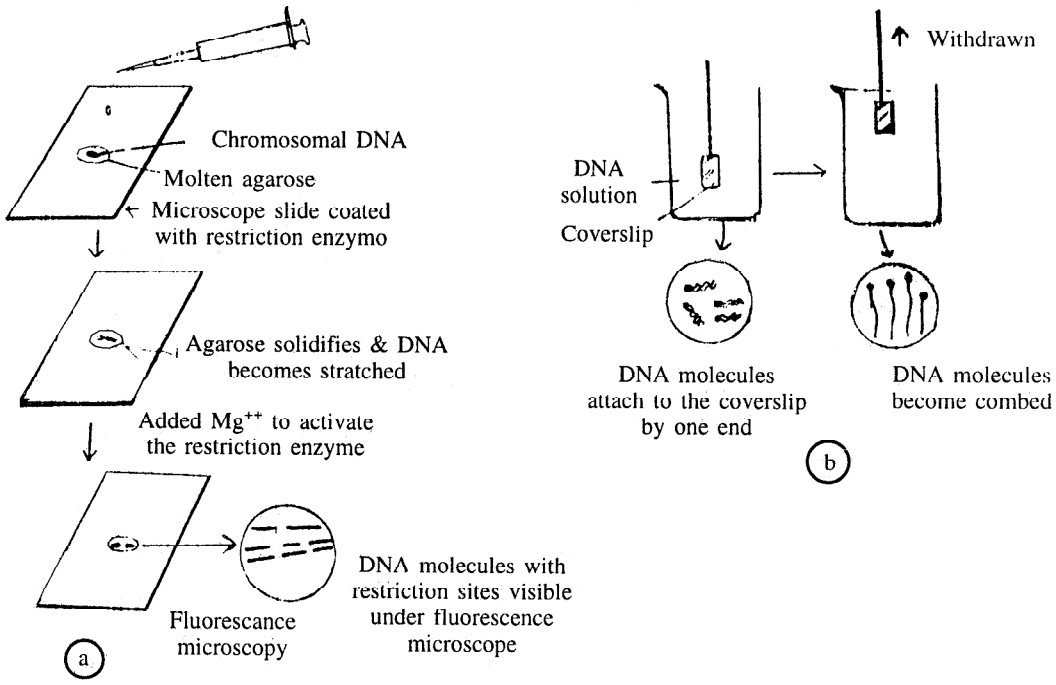


Fig. 9.1 (a) Gel stretching and (b) molecular combing in fibre FISH

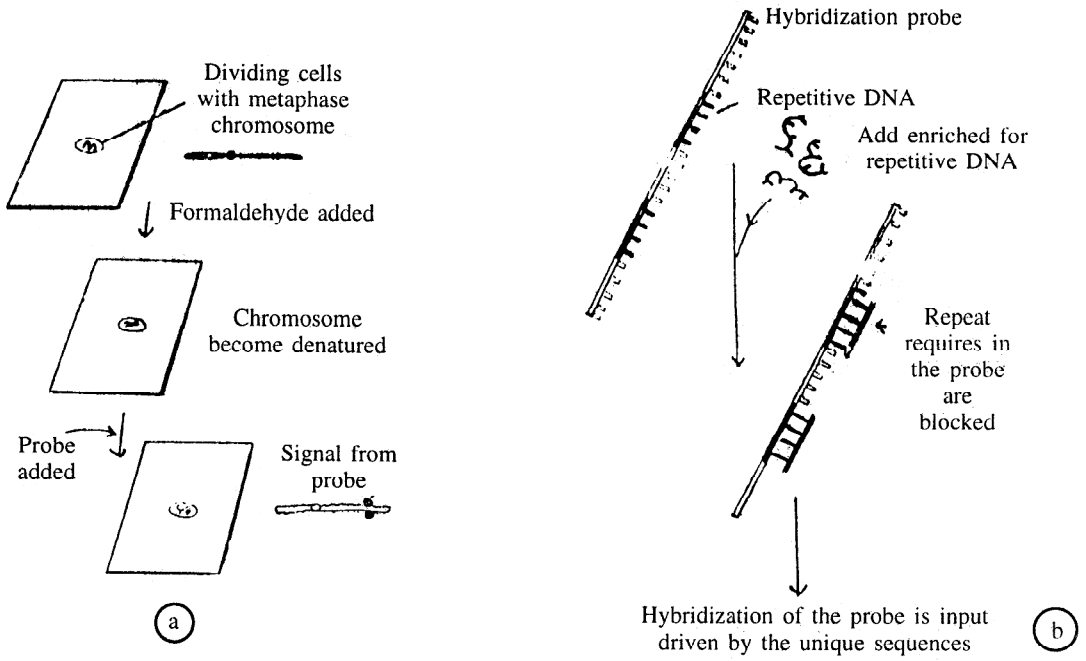


Fig. 9.2 (a) FISH ; (b) Blocking of repetitive DNA sequences in a hybridization probe

phosphatase. After addition of a soluble substrate, the enzyme catalyses a reaction that results in formation of an insoluble blue coloured precepitate at that site of hybridization. The intensity is proportional to the amount of biotin in the hybrid.

In situ hybridization with a biotin labelled probe has been particularly useful in chromosome mapping of DNA clones in *Drosophila* because the logical map of the poly-tene chromosomes of this organism is known at high resolution.

9.2 DNA finger printing

The techniques of DNA diagnosis have found application in a quite different area, the identification of medicine. This is important in areas as diverse as identifying cell cultures, determining family relationships in studies of animal behaviour, immigration problems to identify criminals or murderer, disputed parentage and in forensic medicine.

The most accurate method of identification technique based on recombinant technology is called DNA finger printing or DNA typing or DNA profiling.

Principle

DNA finger printing is based on **sequence polymorphism** that occurs in human genome and the genome of other organisms. The sequence polymorphisms are slight sequence differences, usually single base pair changes, that occur from individuals to individuals once in every few hundred base pairs on average. Each difference from the consensus human genome sequence is generally present in only a fraction of the human population but every individual has some of them. These polymorphic locus is called minisatellite or VNTR (variable numbers tandem repeat) thus forming a haplotype which shows mendelian inheritance among their offsprings. This locus is made up of a variable number of identical sequences joint together in tandem.

One family of minisatellites in the human genome share a common "Core" sequence, The core is G-C rich sequence of 10-15 bp showing on asymmetry of purine/ pyrimidine distribution on the two strands. These repeats are written as (C-A) n (G-T)n, occur in 100000 blocks in every genome and appear to be uniformly distributed throughout the genome (value of n varies from 1 to 40). The successful application of these (C-A)n (G-T)n repeats has led to the use of a variety of other di-, tri- **and** tetranucleotide sequences for mapping.

Technique

(1) DNA of the sample is first isolated whose DNA finger printing has to be made. Usually in forensic case the DNA **is** prepared from dried stains, sperms

in vaginal swabs that had been stored for as long as two years. Sufficient DNA can be isolated from freshly pulled hair roots, polymorphism in mitocondrial DNA and class II HELA gene DQ have been analysed from the shed hairs of several months old containing less than 1 ng of DNA.

(2) Sufficient quantities of intact DNA from forensic samples will always be a problem. PGR may have a great impact in this area. From a small quantities of DNA PCR can produce a large number DNA. These are used as probes. This probe is more redioactive.

(3) The DNA from the individual whose DNA is to be compared with the forensic sample is isolated and are cut into fragments by restriction enzymes.

(4) DNA fragments after digestion of DNA from the genome are first separated according to their size by agarose gel electrophoresis. These ds DNA are denatured by soaking the gel in alkali to make it ss DNA.

(5) The DNA fragments are transferred to nito-cellulose paper by the southern blot technique. The paper is then immersed in a solution containing a radioactively labeled DNA probe. Fragments to which the probe hybridizes are revealed by autoradiography.

Detection of forensic problem :

The power of DNA fingerprinting was demonstrated by Alec Jeffreys in 1985, when a man had been aceured of two rape murders committed three years apart and had made a confussion. Lastly the real murderer was caught and DNA finger printing confirmed the identification.

DNA from a semen sample obtained from a rape and murder victim was analysed along the DNA samples from the victim and two suspects.

Each of the DNA samples was cleaved into fragments and separated by gel electrophoresis.

Radioactive DNA probes were used to identify a small subset of these fragments that contained sequences complimentary to the probe. The sizes of the fragments identified varied from one individual to the next. The different patterns for the three individuals (victim and two suspects) tested. One rape suspects DNA exhibits a banding pattern identical to that of the semen sample taken from the victim.

More than one probe may be used to make a positive identification.

9.3 Separation of chromosomes by flow cytometry

The dividing cells with condensed chromosomes are carefully broken open so that a mixture of intact chromosomes is obtained. The chromosomes are then stained with fluorescent dye. The amount of dye that a chromosome binds

depends on its size. Thus larger chromosomes bind more dye and fluoresence more brightly than smaller ones. The mixture of chromosomes is diluted and passed through a fine aperture, producing a stream of droplets, each one containing a single chromosome. The.droplets are passed through a detector that measures the amount of fluorescence and thus identifies which droplits contain the particular chromosome (being sought). An electric change is applied to these droplets by a charger and then the droplets reach the electric plates, the changed ones are deflected into a separate beaker.

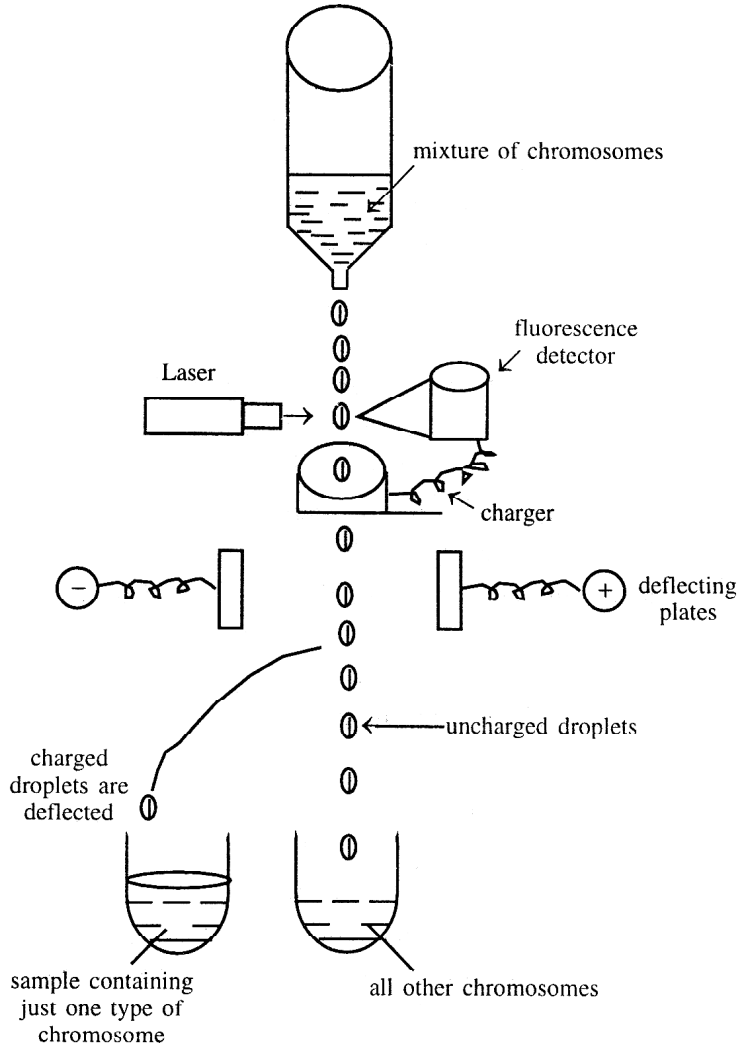


Fig. 9.3 Flow cytometry

If two chromosomes are of equal size as in human chromosome number 21 and 22, then the dye, tioechst 33258 and chlomomycin-A preferably bind to the A-T and G-C rich DNA respectively and by this differential staining activity, these chromosomes can be distinguished properly (Fig. 9.3).

9.4 Chromosome painting

This is a method for visualizing each of the chromosomes in distinct bright colours and thus it simplifies greatly the distinction between chromosomes of similar size and shape and the karyotyping of the chromosome. Usually such a painting is done at the stage in the cell cycle (mitosis) when chromosomes are specially compact and easy to visualize, usually at mitosis. But sometimes, selective chromosome painting may be done in the interphase stage to see its orientation in the nucleus.

Method

(1) The probes which are used for chromosome painting are specific for sites scattered along the length of each chromosome.

(2) The probes are labeled with one of two dyes that fluorence at different wave lengths. For example, DNA molecules derived from chromosome-1 are labeled with one specific dye combination, chromosome-2 with another and so on.

(3) The labeled probe can hybridize only the chromosome from which it was derived, each chromosome is differently labeled. After the probes are hybridized to chromosomes the excess is removed, the sample is placed in a fluorescent microscope in which a detector determines the fraction of each dye present at each fluorescing position in the microscopic field. The information may be conveyed, to a computer and a special program assigns a false colour image to each type of chromosome.

Use

Chromosome painting is very useful in karyotyping the chromosomes. It can be done in interphase stage to locate the specific chromosomes and their arrangement in the nucleus. A combination of chromosome banding with FISH, called multicolor FISH can detect chromosomal translocations which are associated with certain genetic disorder and specific types of cancers. For example, in chronic myeloid leukemia (CML), the lymphocytes contain the Philadelphia chromosome(small), which is produced by the translocation of chromosome no-22 and chromosome no-9. The translocations can be detected by classical banding analysis technique.

Unit 10 Genome Analysis

Structure

10.1 C-value paradox

The haploid DNA content in an individual is described as its C-value.

The anomalies of the gene contents by two different methods one on the basis of knowledge about the rate of mutation per locus and other on the basis of general method used for DNA content, is called C-value paradox.

DNA content in eukaryotic cells are much higher than that in the prokaryotic cells and a wide range of variations are observed among different species even among same species. The content of DNA also depends on the number of chromosomes in the cells (i.e., ploidy of the chromosome). Example : See Table 10.1.

Table 10.1 DNA content of some organisms

Class of organisms	Species	**Haploid DNA (Picogram)**	**Dalton**	Base pairs
Phages	∅ × 174 T4	2.6×10^{-6} $20.7 \times lO^{-5}$	1.7×10^{6} 126×10^{6}	5400 200000
Animal virus	Adenovirus	21.7×10^{-6}	13×10^{6}	21000
Prokaryotes	*E. coli*	4.4×10^{-3}	2.7×10^{9}	$4.2 \times I0^{6}$.
Unicellular-eukaryotes	*S. cerevisial*	$14 \times lO^{-3}$	8.5×10^{9}	$1.4 \times I0^{7}$
Multicellular enkaryotes	*D. melanogaster*	0.18	0.11×10^{12}	0.17×10^{9}
	Homo sapiens	2.8	18×10^{12}	2.8×10^{9}

C-value paradox takes its name from the inability to explain the content of a genome in terms of an anticipated function. There are two aspects of the

paradox. First there are huge variations in C-values between certain species whose apparent complexity does not vary correspondingly. There can be rather substantial variations even between certain closely related species.

The range of C-values is found in different evolutionary phyla. There is some increase in the minimum genome size that is found in each phylum as the complexity increases.

For example, in prokaryotes, the genome size are very small. In eukaryotes, a vast increase occurs in genome size. In yeast, *Saccharomyces cerevisiae* has a genome size of 2.3 × $]0^7$ bp, only 5 times greater than that of *E. colt.*

The modest increase in genome size just over two folds is adequate to support the slime mold, D *discoideum,* able to live in either unicellular or multicellular modes.

Another increase in complexity is necessary to produce the first fully metazoan organisms. For example, *C. elegans* has a DNA content of 8×10^7 bp. Then any close relationship between complexity of the organisms and content of DNA is obscure, although it is necessary to have a genome of more than $I0^8$ bp, to make an insect of more than 4×10^8 bp, to assemble an echinoderm of more than 8×10^8 bp, to produce a bird or amphibians and more than 2×10^8 bp to develop into a mammals.

In some cases the spread of genome size is quite small. For example, birds, reptiles and mammals, all show a little variation within the phylum, with a range of genome size in each case about two fold. But in other cases, there is quite a wide range of values, often more than 10 folds. This reflects some surprizing discrepencies between genome size and complexity of the organism.

An extraordinary C-value is found in amphibian where the smallest is below 10 bp while the largest are almost 10^{11} bp. It is hard to believe that this could reflect a 100 fold variation in number of genes in different amphibians.

There are some cases where rather closely related species show surprising variations in total genome size. For example two amphibian species may have 10 fold increase where morphologies are very similar. Yet if the gene number is roughly similar most of the DNA in the species with the larger genome cannot be concerned with coding for protein. So the question, what could be its function?

The second aspect of C-value paradox is the apparent absolute excess of DNA compared with the amount that could be expected to code for proteins.

Actually, eukaryotic DNA has an excess length and the excess is encountered because genes are much larger than the sequences needed to code for proteins. For example, human haploid cell has DNA amount equal to 1.8 x $]0^{12}$ dalton = 87 cm of DNA which is equal to 2.8×10^9 base pairs, then this genome could contain approximately as many as 3×10^6 genes assuming 1000 bp per gene coding for nearly 300 amino acids.

However, the number of genes estimated in humans on the basis of the rate of mutation per locus, as estimated by Muller (1967), the frequency of deleterious mutations per locus in human is 10^{-5} to 10^{-6} in each generation. If the number of gene is 3×10^6 (as calculated in general method), then it will yield 30 deleterious mutaions in each generations at the rate of 10^{-5} mutations per locus. This will be an unbearable genetic load. The actual frequency of deleterious mutations per generation per individual has been estimated to be 0.5 against an expected frequency of 30. This mutation frequency at the rate of 10^{-5} per locus will be an estimate of 5×10^4 genes.

Thus it is supposed that the actual number of genes in human should be 50000 and not 3 million an estimated from DNA content.

This anomalous situation has been described by some workers as C-value paradox.

10.2 Satellite DNA

Large proportions of DNA in eukaryotes has been shown to be present in the form of multiple copies of identical DNA sequences, thus is called repetitive DNA or Satellite DNA. The remaining DNA in the cell is found in the form of single copy of DNA sequences which is known as unique DNA.

When the denatured DNA (single stranded) is led to reassociate, then it was observed that from the heterogenous populations, the smaller molecular weight DNA associate easily. Britten and his Coworkers (1966, 68) have demonstrated that many vartebrate DNAs reassociates easily when it is broken into small pieces. This observation gave rise to the hypothesis that certain short sequences of bases are repeated hundred times in DNA, this is the **satellite DNA.**

This repetivive DNA generally contributes at least 20% DNA and can reach upto 90% in some cases. It is believed that these repetitive sequences do not carry any genetic informations and therefore do not form genes, but play some other structural or regulatory role.

Repetitive DNA consists of short identical genes which are repeated in tandem, several hundred or thousand times. Such DNA is found in the region of the chromosome adjacent to the centromere. In many case the base compositions of the repeating sequences are unlike that of the rest of the DNA. It is, therefore easy to separate repetitive DNA by ultracentrifugation.

The satellite DNA can be isolated by density gradient centrifugation in neutral caesium chloride as they have distinctive bouyant densities. The fractions can be separated as a band from the main band of DNA, this band is called satellite band.

The satellite band in found on the left of the main band if lighter and on the right side if havier than the DNA of the main band.

In *Drosophila virilis,* there are three highly repetitive DNA each consisting of a repeating sequence of seven nucleotide pair and about 25%. of the DNA is satellite DNA (See Table 10.2).

Table 10.2 Satellite DNA of *Drosophila virilis* (Gall *et al.,* 1974)

Repetitive DNA	Bauyont density	Repeat sequence
Sat DNA-I	1.692	5' ACAAACT 3' 3' ṪĠṪṪṪĠȦ5'
Sat DNA-II	1.688	5' ATAAACT 3' 3' ṪȦṪṪṪĠȦ 5'
Sat DNA-III	1.671	5' ATAAATT 3' 3' ṪȦṪṪṪȦȦ 5'

In human, 30% of DNA is repetitive and in designated as sat I, II and III.

In mouse, 10% of the DNA is highly repetitive and renaturates within a few second, 20% is moderately repetitive and reassociates at an intermediate rate, 70% is-single copy DNA which renaturates very slowly. There are about a million copies of repetiting sequences of about 300 bp.

In prokaryotes, the repeated base sequence in not found.

Two remarkable features of satelite DNA are—

(I) Remakable (relative) uniformity within the same species.

(II) Great variability between closely related species.

The satellite DNA often lies in heterochromatic region of chromosomes and its location can be demonstrated by cytological hybridization by incubating the cells in the radioactive solution and is determined by a autoradiography.

The function of highly repetitive DNA is unknown. This can replicate but cannot transcribe RNA for protein synthesis. This is probably because the short sequences lack promoter sites on which RNA chains can be initiated by RNA polymerase. Repetitive DNA is, therefore, inert and is partly dispensible.

In the African clawed toad, *Xenopus laevis,* the genes for 40s precursor RNA which give rise to 28s and 18s RNA are repeated about 450 times. The genes are tandemly arranged and are separated by a spacer region of about 5000 bp. Genes for 5s **r** RNA are also separated by spacer regions and are arranged

in clusters of 100 to 1000 repetitive units at the ends of the most of the 18 chromosomes.

10.3 Complexity

Complexity is the total length of different sequences of DNA present in a given proparation. The double stranded DNA is denatured and converted into single stranded DNA by heating the DNA solution. This is accompanied with increase in optical density, which is called hyperchromicity.

Again when it is allowed to cool, the single stranded DNA is transformed into a double stranded DNA, again the optical density is decreased, this is called hypochromicity. The 50% resaturation is achieved usually at a specifie temperature which is called melting temperature (Im). The formation of double stranded DNA **is** actually measured over different values of a parameter which is described as C_0. t (concentration x time). It is the product of DNA concentration and time of incubation in a reassociation reaction.

Complexity of the genome can be described under two heads—

A. Kinetie complexity

The reassociation of DNA in the solution depends on the random collisions betwe en the complementary strands, which follow the second order of **kinetics,** since concentration of both the complementary strand will influence the rate of reaction. The rate of reaction, when expressed through differential calculus **is** as follows :—

$\frac{dc}{dt} = -Kc^2$ where, C = Concentration of single stranted DNA at time 't'

K = reassociation rate constant

or $\int \frac{dc}{c^2} = -K \int dt$

or, $-\frac{1}{C} = -Kt + A$ A is constant

or, $\frac{1}{C} = Kt - A$

When t = o, then C = C_0

Now, $\frac{1}{C_0} = KXO - A$

or, $A = -\frac{1}{C_0}$

The equation is

$$\frac{1}{C} = Kt + \frac{1}{C_0}$$

When the reaction is half complete then time is ($t_{1/2}$) and in $t_{1/2}$ time the concentration is C. Then—

$$\frac{1}{C} = Kt_{1/2} + \frac{1}{C_0}$$

or, $\frac{C_0}{C} = k.C_0.t_{1/2} + 1$

We know $C = \frac{1}{2}C_0$

then, $1 + K.C_0.t_{1/2} = 2$

or, $K = \frac{1}{C_0.t_{1/2}}$

or, $C_0.\ t_{1/2} = \frac{1}{k}$

So, during reassociation of DNA occurs at the rate constant K (nt. moles/lit/see) in equal to the receprocals of $C_0.t_{1/2}$

$C_0.\ t_{1/2}$ is the product of DNA concentration and time (t) required to proceed to half completion of the reaction; it is directly proportional to the unique length of reassociating DNA.

The $C_0.t_{1/2}$ of a reaction indicates the total length of different sequence that are present. This is described as the complixity. It is usually given in base pairs, but can be expressed in daltons or any other mass unit.

A higher $C_0.\ t_{1/2}$ means, slower reaction and lower $C_0.\ t_{1/2}$ means faster reaction.

If there is no repetitive DNA (because the repetitive DNA reassociates faster), the $C_0.\ t_{1/2}$ of a reaction will be directly prortional to the DNA content. In view of this $C_0.\ t_{1/2}$ will indicate the length of all the different sequences in a genome, which will be less than the length of the total DNA in a genome when

there is repetition. This is called kinetie complexity.

Kinetie complexity is the complexity of a DNA component measured by the kineties of the DNA association.

This can be calculated by knowing the $C_0 . t_{1/2}$.

For example, *E. coli* has a genome = 0.004 pg DNA = 4.2×10^6 bp with $C_0 . t_{1/2} = 4$.

The Kinetie complexity of the genome

$$= \frac{C_0 . t_{1/2} \text{ of DNA of any genome} \times \text{base pair}}{C_0 . t_{1/2} \text{ of } E.coil \text{ genome}}$$

$$= \frac{4.2 \times 10^6 \times C_0 . t_{1/2} \text{ of the genome}}{4}$$

In eukaryotes, the genome contain more than one pure components. For example. calf DNA has two component, each with characteristics $C_o.t_{1/2}$ value. In wheat, more than two such components are found. Proportion of each component is determined by using the formula $1 - \frac{C}{C_0}$, where C = concentration at $t_{1/2}$ for corresponding component.

Form this proportion, chemical complexity on the component can be determined.

B. Chemical Complexity : Chemical complexity is the amount of a DNA component measured by chemical assay.

For example, if the genome size is 12×10^8 bp and the component represents 25% of the genome then the chemical complixity of this component is 3×10^8 bp.

Chemical complixity = size of the genome × % of the component in this group.

If the kinetic complexity is known from earlier equation, then repetition frequency (f) of repetitive DNA component can be determined using the following formula——

$$f = \frac{\text{chemical complexity}}{\text{kinetie complexity}} = \frac{C_0 . t_{1/2} \text{ of non repetitive DNA}}{C_0 . t_{1/2} \text{ of repetitve DNA}}$$

Following table shows the reassociation of a eukaryotic genome starting at a C_0. t of 10^4 and terminating at C_0.t of 10^4. Reaction falls into three types of components and their results are as follows—

	Fast Component	Intermediate Component	Slow Component
% of genome	25	30	15
$C_0 \cdot t_{1/2}$	0.0013	1.9	630
ECinetie complexity (bp)	340	6×10^5	3×10^8
Repetition frequency	500000	350	1

There is a good relationship between the kinetic complexity and chemical complexity of eukaryotic genome.

Usually *E. coli* is used as a standard. Its components are taken to identical with the length of the genome. Thus, complexity of any DNA can be determined by compairing its $C_0 \cdot t_{1/2}$ with that of standard DNA of know DNA complexity.

$$\frac{C_0 \cdot t_{1/2} \text{ of (DNA of any genome)}}{C_0 \cdot t_{1/2} (E.coli \quad \text{DNA})} = \frac{\text{Complexity of nay genome}}{4.2 \times 10^6 \text{ bp}}$$

According to the table, the slow component represents 45% of the total DNA, so the concentration in the reassociation reaction is 0.45 of the measured concentration (total amount of DNA).

If DNA were isolated as a pure component, free of other fractions, it would renature with $C_0 \cdot t_{1/2}$ of $0.45 \times 630 = 283$.

Suppose that under these conditions *E. coli* DNA reassociates with a $C_0 \cdot t_{1/2}$ of 4.0, Comparing these two clues, we see :

$$\text{The kinetie complexity} = \frac{C_0 \cdot t_{1/2} \text{ of DNA of any genome} \times 4.2 \times 10^6}{4}$$

$$= \frac{0.45 \times 630 \times 4.2 \times 10^6}{4}$$

$$= 3 \times 10^8 \text{ bp}$$

$$\text{Then the whole genome is} = 3 \times 10^8 \times \frac{1}{0.45} = 6.6 \times 10^8 \text{ bp}$$

$$\text{First component complexity} = \frac{0.0013 \times 0.25 \times 4.2 \times 10^6}{4} = 340$$

$$\text{Second component complexity} = \frac{1.9 \times 0.30 \times 4.2 \times 10^6}{4} = 6 \times 10^5$$

Reversing the argument if we took three DNA preparations, each containing a unique sequence of the appropriate length 340 bp, 6×10^5 bp and 3×10^8 bp respectively and mix them in the proportions = 25 : 30 : 45, each would renature as though it was a single component, together the mixture would display the same kineties as those determined for the whole genome.

Non repetitive DNA complexity can estimate the genome size. The complexity of the slow components comprise sequences that are unique in the genome upon denaturation each single stranded sequence is able to renaturate only with the corresponding complementary sequences. It is usually the major component in eukaryotis. It is called non-repetitive DNA. According to the table the compexity of non-repetitive DNA is 3×10^8 bp. If this fraction is unique and represent 45% of the genome, then the whole genome would have a size of $3 \times 10^8 \div 0.45 = 6.6 \times 10^8$ bp. This provides an independent assessment of genome size. The value is approximately 7×10^8, obtained from the result of chemical complexity.

Eukaryotic genomes certainly contain repetitive sequences. Intermediate component occupies 30% of the genome. According to chemical complexity the total amount is $0.30 \times 7 \times 10^8 = 2.1 \times 10^8$ bp.

But kinetic complexity of this component is only 6×10^5 pp.

Thus repetition frequency = $= \frac{2.1 \times 10^8}{6 \times 10^5} = 350$

Thus intermediate components behaves as though consisting of a sequence of 6×10^5 bp that present in 350 copies in every genome. Repetition frequency (f) is the number of copies present per genome.

Highly repetitive DNA takes the name from the very large number of copies of the basic reassociating sequence present. The fast component consists 310 bp long in 500000 copies per genome. Because of the short length of the reassociating unit sometimes this is also referred to as simple sequence DNA.

$$\text{The repetition frequency (f)} = \frac{C_0 . t_{1/2} \text{ of non repetitive DNA}}{C_0 . t_{1/2} \text{ of repetitive DNA}}$$

$$= \frac{630}{0.0013} = 500000 \, (\text{Approx.})$$

Unit 11 Linkage Map, Cytogenetic Mapping

Structure

11.1 Physical and molecular maps

STS, i.e. Sequence Tagged Site is a short DNA sequence generally between 100-500 bp in length that is easily recognisable and occurs in the chromosome only once (i.e. unique).

STS mapping is a physical mapping procedure that locates the positions of sequence tagged sites (STSs) in a genome.

11.1.1 Qualities of STS

(1) Its sequence must be known, so that a PCR (polymerase chain reaction) assay can be set up to test the presence or absence of the STS on different DNA fragments.

(2) It must have a unique location in the chromosome being studied once in the genome. If the STS sequence occurs in more than one position, then the mapping data will be ambiguous. So STSs do not include sequences found in repetitive DNA.

11.1.2 Sources of STS

(1) **Expressed sequence tags (ESTs) :** These are short sequences obtained by the analysis of cDNA clones. cDNA is prepared by converting mRNA into dsDNA. Thus ESTs represent the genes that are expressed in the cell.

(2) **SS4Ps :** These are also used in genetic mapping as well as physical mapping of genes.

(3) **Random genomic sequences (RGS) :** These are obtained by sequencing random pieces of cloned genomic DNA.

11.1.3 Principles of STS mapping

(1) To map a set of STSs, a collection'of overlapping DNA fragments from single chromosome or from the entire genome is needed.

(2) The data from which the map will be derived are obtained by determining fragments which contain STSs. This can be done by hybridization analysis but PCR is generally used which is quieter and automated process.

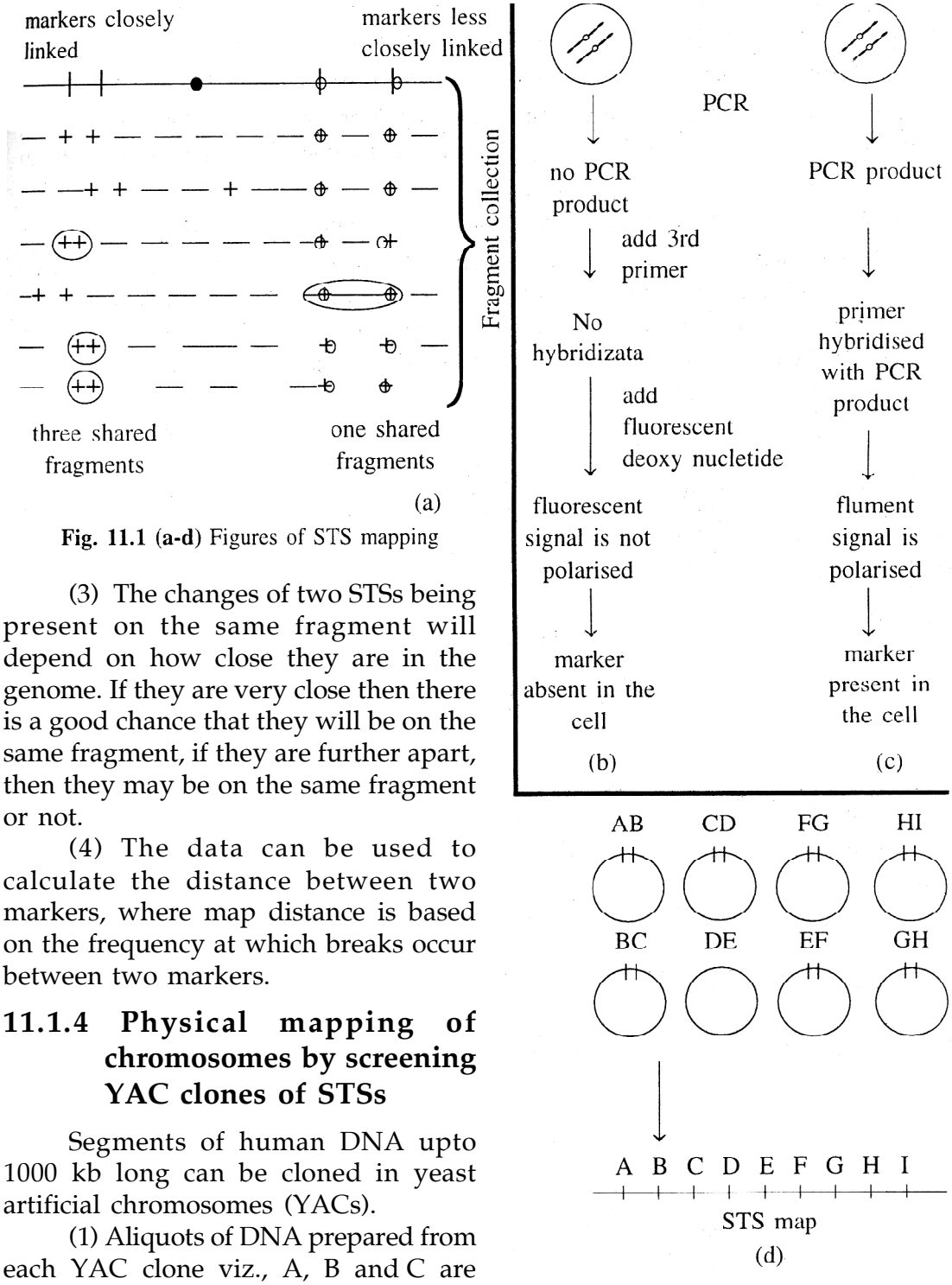


Fig. 11.1 (a-d) Figures of STS mapping

(3) The changes of two STSs being present on the same fragment will depend on how close they are in the genome. If they are very close then there is a good chance that they will be on the same fragment, if they are further apart, then they may be on the same fragment or not.

(4) The data can be used to calculate the distance between two markers, where map distance is based on the frequency at which breaks occur between two markers.

11.1.4 Physical mapping of chromosomes by screening YAC clones of STSs

Segments of human DNA upto 1000 kb long can be cloned in yeast artificial chromosomes (YACs).

(1) Aliquots of DNA prepared from each YAC clone viz., A, B and C are

subjected to PCR amplification using primer pairs (1-6) corresponding to the ends of various STSs. Only those clones containing STSs with ends complementary to particular primers will be amplified.

(2) Electrophoretic analysis then shows that YAC clones contain STSs.

(3) The illustration is very simple showing 6 primer pairs. Clone-A contains STSs no. 1,3 and 5; clone-B contains 2 and 1 and clone-C contains 3, 4, 5 and 6 primer pairs.

(4) The three YAC clones can be ordered showing their relative positions. In the mapping of human chromosome-21, about 1,20,000 clones from two separate YAC libraries were screened. In addition, 14,000 YACs isolated from a library prepared specifically from chromosome-21 were screened individually. By the use of 198 STSs, researchers identified 810 positive clones and ordered them into a contiguous map.

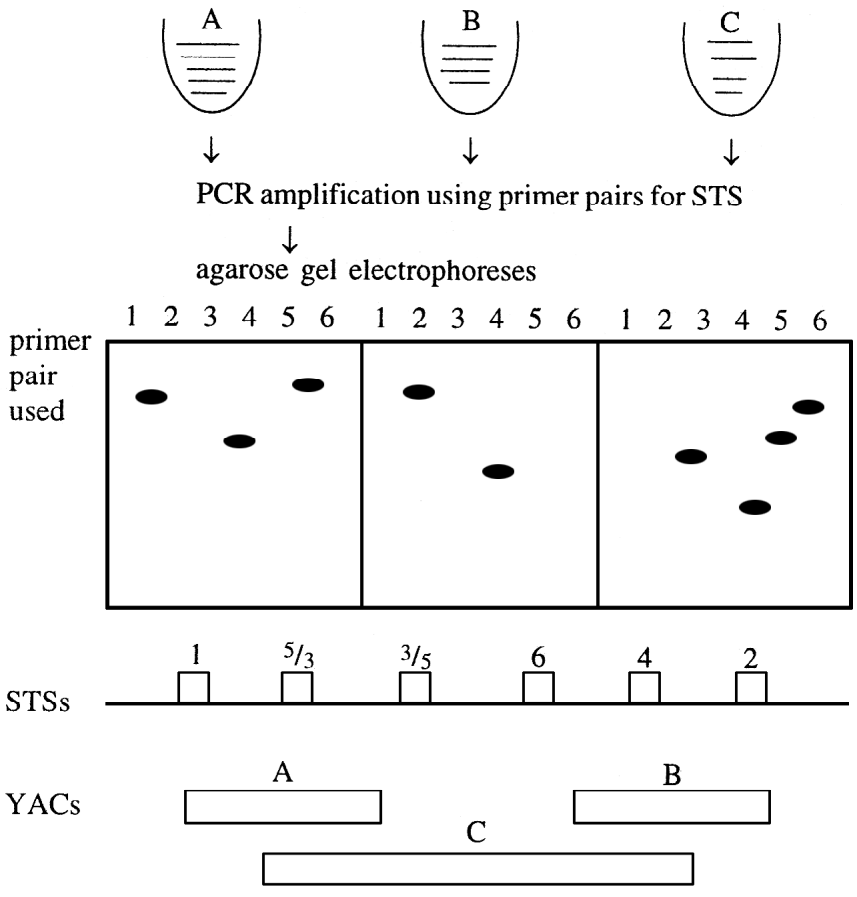


Fig. 11.2 Ordering of contiguous overlapping YAC

11.1.5 Fragments of DNA for STS mapping

At first, collection of DNA fragments sparing the chromosome or genome is required. The collection is called mapping reagent. There are two method as follows—

A. **Radiation hybrid method** :

(1) Human cells are exposed to X-ray doses of 3000-8000 rads, which causes the chromosomes to break randomly into fragments. Higher doses produce smaller fragments.

(2) These fragments can be propagated if the irradiated cell is fused with non-irradiated hamster (or other rodent) cells. Fusion is achieved chemically with polyethylene glycol or by exposure to Sendai virus.

(3) The hamster cell line that is unable to make either thymidine kinase (TK) or hypoxanthine phosphoribosyl transferase (HPRT) is used for the purpose. Those cells incapable of taking up human chromosome fragments are unable to survive in HAT (hypoxanthine, aminopterin and thymidine) medium. Those cells taking up chromosome fragments can synthesize TK and HPRT and are able to grow in HAT. Thus the hybrid cells are collected. These hybrid cells are then used as a mapping reagent in STS mapping.

(4) To map, as many markers as possible are us.ed. A pair of primers is designed for every DNA marker that was to be tested. Each primer pair is specific for one marker and will not give a PCR product with any other part of the genome.

(5) The success or failure of PCR is determined by agarose gel electrophoresis. The presence of a band of the expected size in the gel indicates that the PCR has worked.

(6) Another procedure has been designed where a third specific primer for each marker is added to the reaction mixture along with a fluorescently labelled dideoxy-nucleotide. If PCR has been successful, then the third primer anneals'to the product and is extended by the fluorescent dideoxy nucleotide and emits signals from this.

B. Use of clone library as the mapping reagent for STS analysis

A clone library can also be used as a mapping reagent in STS analysis. The clone library is prepared by using the genome or chromosome which is broken into fragments and are integrated into a high capacity vector.

The single specific chromosome can be isolated by flow cytometry technique and a clone library of a chromosome is made possible.

The data obtained from STS analysis is used for preparation of the physical map.

11.2 Restriction mapping of genes

Restriction mapping of genes is a physical mapping which locates the relative positions on a DNA molecule of the recognition sequences for restriction endonucleases.

Genetic mapping using RFLPs as DNA markers can locate the positions of polymorphic restriction sites within a genome, but very few of the restriction sites in a genome are polymorphic, so many sites are not mapped by this technique. Restriction mapping is very useful to solve a problem but the limitation of the technique is that it is applicable only to relatively small DNA molecules.

Two methods are employed for restriction mapping *viz.,* partial digestion method and double digestion method.

(1) In **partial** digestion method, the circular or linear DNA is treated with a particular restriction enzyme. But the enzyme is prevented from the complete digestion of DNA molecule. So the DNA is incubated for a short time or using a suboptimal incubation temperature. This leads to the partial digestion and will produce many cut products and uncut products. If the DNA molecule is linear, the terminal end of the DNA is labeled with P^{32} by the polynucleotide kinase reaction prior to cleavage and only radioactively labelled fragments are considered in agarose get electrophoresis, the other fragments are ignored/discarded of the molecular weight of each enhanced band is invariably the sum of the molecular weights of two fragments which arc considered to be adjacent. Thus the relative positions of the fragments can be ordered.

(2) Second method of the restriction mapping is the double digestion method, although partial digestion method is usually followed. In double digestion method, three samples of a particular DNA species are taken. Each two of these is treated with two separate restriction enzymes and the third sample is treated with both the restriction enzymes. Thus the three sets of fragments are compared following the agarose gel electrophoresis. The terminal end of the DNA molecule is also labeled with radioactive P^{32} prior to the restriction enzyme cleavage.

11.3 DNA foot printing

When transcription factor binds to a DNA sequence, it protects that sequence from digestion by nucleases. Researchers take advantage of this property by isolating chromatin from cells and treating it with DN-digesting enzymes, such as DNA ase-I, that destroy sections of the DNA that are not protected by bound transcription factors. Once the chromatin has been digested, the bound protein

is removed and the DNA sequences that had been protected are identified. This method is called DNA foot printing. This is also used to locate the binding sites of proteins on RNA.

Method

(1) A pure DNA fragment that is labeled at one end with ^{32}P is isolated.

(2) This molecule is then cleaved with a nuclease or a chemical that makes random single-stranded cuts in the DNA.

(3) The DNA molecule is then denatured to separate into two strands.

(4) The resultant fragments from the labeled strand are separated on a gel and detected by autoradiography.

The pattern of bands from DNA cut in the presence of a DNA-binding protein is then compared with that from DNA cut in its absence. The protein covers the nucleotides at its binding site and protects from DNA ase. The labeled fragments that shows no cleavage will show an area which is missing in the electrophoretic gel, is leaving a gap is called "foot print".

11.4 Gene mapping by human pedigree analysis (microsatellite mapping)

Recombination (CO) mapping is very difficult in human because

(1) It is impossible to preselect the genotypes of parents and set up crosses.

(2) The data for the calculation of recombination frequencies have to be obtained by examining the genotypes of the members of successive generations of existing families.

(3) The data obtained are very limited and their interpretation is often - difficult because in human test cross rarely occurs and the number of family members and offsprings are not many.

Therefore, gene mapping in human may be done by pedigree analysis.

Let us, suppose a family of two parents and six children were studied with a genetic disease. The diseased state is due to one allele and the healthy state is due to second allele. Diseased allele is dominant over healthy allele.

The pedigree showed that mother is affected because four of her children are affected by the disease (3♂ + 1♀).

The grand mother is affected. The grand father is dead but? We can assume that he was also affected. We can include them in the pedigree analysis.

The aim is to map the position of the gene for the genetic disease. For **that** purpose one is studying its linkage to a microsatellite marker M, four alleles of which *viz.*, M_1, M_2, M_3 and M_4 are present in the living family members. One has now to calculate the number of children who are recombinants.

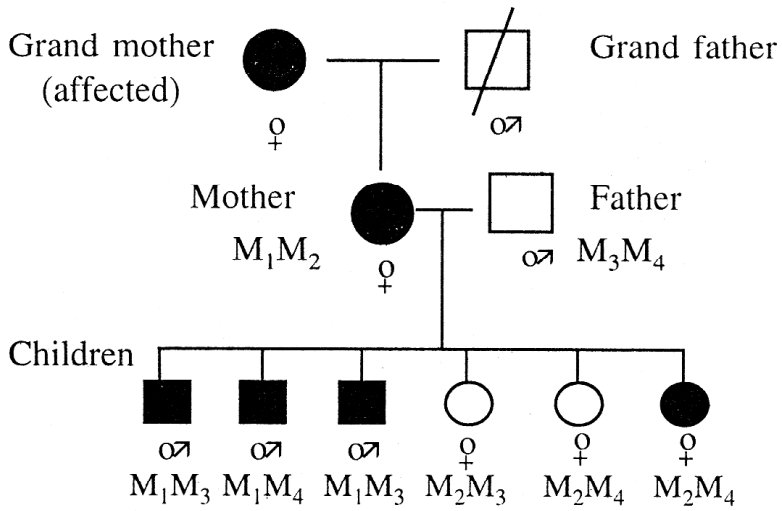


Fig. 11.3 Pedigree of a family with inheritance of genetie disease (solid means affected person)

The pedigree can be interpreted by two different hypotheses because child-1, 2 and 3 have the disease allele and microsatellite allele M], the child-4 and 5 have the healthy allele and microsatellite allele M_2.

In hypotheses-1, the mother would have the genotype = $\frac{\text{Disease allele} - M_1}{\text{Healthy allele - } M_2}$

The child no-1, 2, 3, 4 and 5 all would have parental genotype. Only the child no-6 would be a recombinant. The recombination percentage is 16.66 i.e., the disease gene and **the** microsatellite allele are relatively closely linked.

In hypothesis- II, the genotype of the mother would be = $\frac{\text{Healthy - } M_1}{\text{Disease - } M_2}$

Here the child **no-1,** 2, 3, 4, 5 are recombinants while child no-6 is with parental genotype. The recombination percentage is 88.33, which means that the disease gene and microsatellite gene are far apart on the chromosome.

		Possible genotypes of the mother	
		Hypothesis-I	Hypothesis-II
		Disease-M_1	**Healthy-M_1**
Child-1	**Disease-M_1**	**Healthy-M_2**	**Disease-M_2**
Child-2	**Disease-M_1**	Parental	Recombinant
Child-3	**Disease-M_1**	Parental	Recombinant
Child-4	**Healthy-M_2**	Parental	Recombinant
Child-5	**Healthy-M_2**	Parental	Recombinant
Child-6	**Disease-M_2**	Recombinant	Parental
	Recombination frequency	16.66%	83.33%

Fig. 11.4 Probable interpretation of the pedigree

Imperfect pedigrees are analysed statistically by using a measure called "lod score" (Morton, 1955). This stands for logarithms of the odds that genes are linked. This is used to determine whether the two markers lie on the same chromosome or not. If lod analysis establishes the linkage then the data will give confidence about their recombination frequencies.

If the number of the family members are larger the result would be more satisfactory. Atleast three generations are to be tested. Atleast four grand parents and atleast eight second generation children could be sampled.

Unit 12 Genetics of Cell Cycle

Structure

12.1 Genetic regulation of cell division in yeast and eukaryotes

12.1.1 Introduction

A cell reproduces by performing an orderly sequence of events in which it duplicates its contents and then divides into two. This cycle of duplication and division is known as the cell cycle. The basic organization of the cell cycle and its control system are essentially the same in all eukaryotic cells. Three **eukaryotic** systems in which cell-cycle is commonly studied are yeasts, frog embryo and cultured mammalian cells.

Yeasts are tiny, single-celled fungi. Two species are generally used in studies of cell cycle. The fission yeast viz. *Schizosaccharomyces Pombe,* is a rod shaped cell that grows by elongation at its ends. Division occurs by the formation of a septum or cell plate in the centre of the rod. It has a typical eukaryotic cell cycle with G,, S, G_2 and M phases. In contrast with that happening in higher eukaryotic cells, the nuclear envelope of the yeast cell does not break down during M-phase. The microtubules of the mitotic spindle are formed inside the nucleus and are attached to spindle pole bodies (SPB) at its periphery. The cell devides by forming a partition (cell plate) and splitting into two. The mitotic chromosomes are readily visible.

The budding yeast *Saccharomyces cerevisiae,* also called baker's yeast is a oval cell and divides by forming buds which first appears during G_1 and grows until it separates from the mother cell after mitosis. It has normal G_1 and S-phase but does not have a normal G_2 phase. The microtubule based spindle begins to form inside the nucleus early in the cycle during S-phase. Nuclear envelope remains intact during mitosis and the spindle forms within the nucleus.

12.1.2 Genetic regulation of cell cycle in *S.pombe* (Fig. 12.1)

(1) *ede 2* is identified as a crucial regulator by its involvement at both stages of cell cycle block, i.e. between G_2 and M-phase and in G_1 at start.

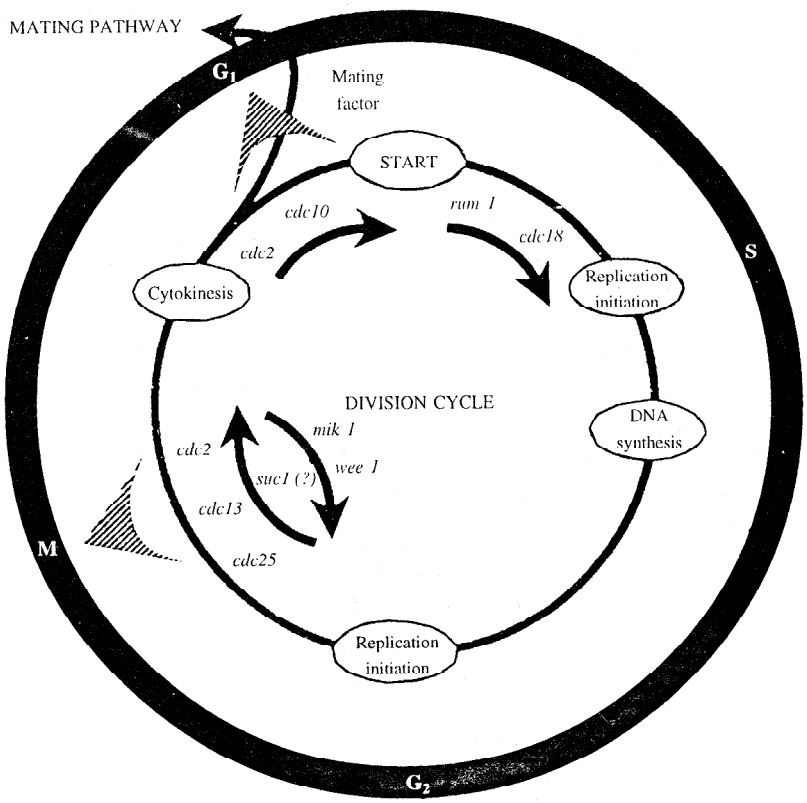


Fig. 12.1 The cell cycle in *S. pombe* requires *cdc* genes to pass specific stages, but may be retarded by genes that respond to cell size *(wee 1).* Cells may be diverted into the mating pathway early in G1

(2) During G_2 and M-phase cdc 2 has a partner which is the product of *Cdc 13* generating an M-phase kinase (resembling the p^{34}-B cyclin dimer of animal cells). The activity of the cdc 2 catalytic subunit in these dimers are controlled by phosphorylation (in the same way as p^{34} in animal cell).

But the difference that in yeast there is no Thr-14 so there are only two relevant sites Tyr-15 where phosphorylation is inhibitory and Thr-161 where phosphorylation is required.

(3) Under normal conditions, the cell division cycle is related to the size of the cell. In the poor growth condition the cells increase in size more slowly. The genes involved in control cell size are identified, *wee* 1 usually inhibits cells from initiating mitosis until their size is adequate. It has been suggested that *wee* 1 is part of a check point that prevents the activation of cdc 2 until an adequate mass has been attained. *wee* 1 codes for a kinase and can phosphorylate serine or

threonine and tyrosine. It inactivates cdc 2 by phyrophorylating Tyr-15. Another gene viz., *mik 1* has similar effects.

(4) The product of *cdc 25* is required for the dephosphorylation of cdc 2 in the cdc 2/cdc 13 dimer. It is probably responsible for the key dephosphorylating event in activating the M-phase kinase. The level of cdc 25 increases at mitosis and its accumulation over a threshold level could be important, cdc 25 executes the checkpoint that ensures s-phase to be completed before M-phase can be activated.

(5) The products of *wee 1* and *cdc 25* plays antagonistic roles. The kinase activity of *wee I* acts on Tyr-15 to inhibit cdc 2 function. The phosphatase activity of *cdc 25* acts on the same site to activate cdc 2.

(6) During mitosis the cdc 2 of cdc 2/cdc 13 dimer is in the active state that locks the phosphate at Tyr 15 and has the phosphate at Thr 161. At the end of mitosis kinase activity is lost and cdc 13 is degraded cdc 2 does not change at this point.

(7) During G, the active form of cdc 2 has a different partner, the B-like cyclin, cig 2 encoded by cig 2 gene. The dimer in converted from inactive state to the active state by dephosphorylation of Tyr-15 residue of cdc 2.

(8) Progression of Gj into *S* is controlled by activation of cdc 2-G| cyclin. In *S. pombe* it is cdc 2/ cig 2.

(9) Transcription of *cdc 18* is activated as a consequence of passing START and cdc 18 is required to enter S-phase. Over expression of cdc 18 allows multiple cycle of DNA replication without mitosis.

(10) For cdc 18 to be active, cdc 2/cdc 13 must be inactive. Again when M-phase kinase is active, it causes cdc 18 to be inactive possibly by phosphorylating it and prevents initiation of another S-phase.

(11) Activity of cdc 2/cdc 13 M phase kinase is influenced by a factor rurn-1, which controls entry into S-phase. When rum-1 is depleted, premature entrance into mitosis occurs and over expression causes cells to fail to enter mitosis. This suggests that M_{m-1} is an inhibitor of the M-phase kinase. It is expressed between G_1 and G_2 and keep M-phase kinase in an inactive state.

12.1.3 Genetic regulation of cell cycle in *S.cerevisiae* (Fig. 12.2)

Cell cycle in *S.cerevisiae* consists of three cycles that separate after START and join before cytokinesis. The cells may be diverted into the mating pathway early in G_1.

(1) Chromosome cycle: In this cycle, duplication and separation of chromosomes, completion of S-phase and nuclear division occurs. Mutation of *cdc 8* stops this cycle in S-phase. Mutation in the chromosome cycle do not stops the cytoplasmic cycle.

(2) **Cytoplasmic cycle** : It consists of bud emergence and nuclear migration into the buds. This cycle can be halted before bud emergence by cdc 24 mutation but the mutation does not prevent chromosome replication.

(3) **Centrosome cycle** : This cycle consists of duplication and separation of spindle polar body (SPB) and organizes microtubules to allow chromosome segregation within the nucleus. Blocking of the cycle by cdc 31 does not prevent S-phase or bud emergence.

Completion of entire cell cycle requires all three constituent cycles because nuclokinesis needs both chromosome and centrosome cycles but cytokinesis requires all the 3 cycles.

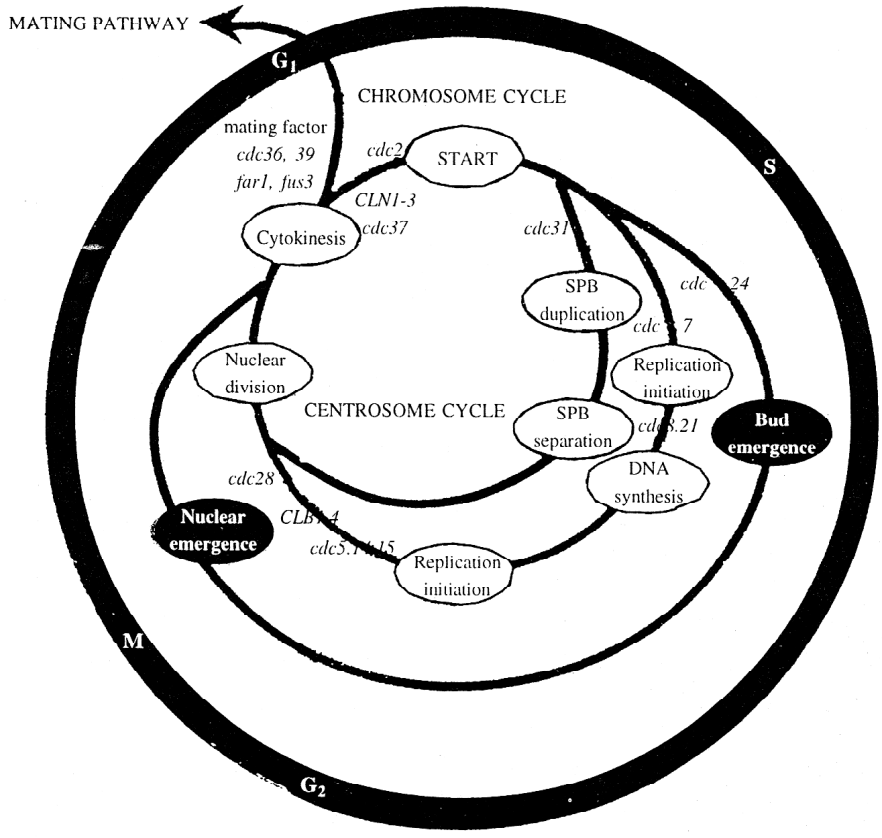
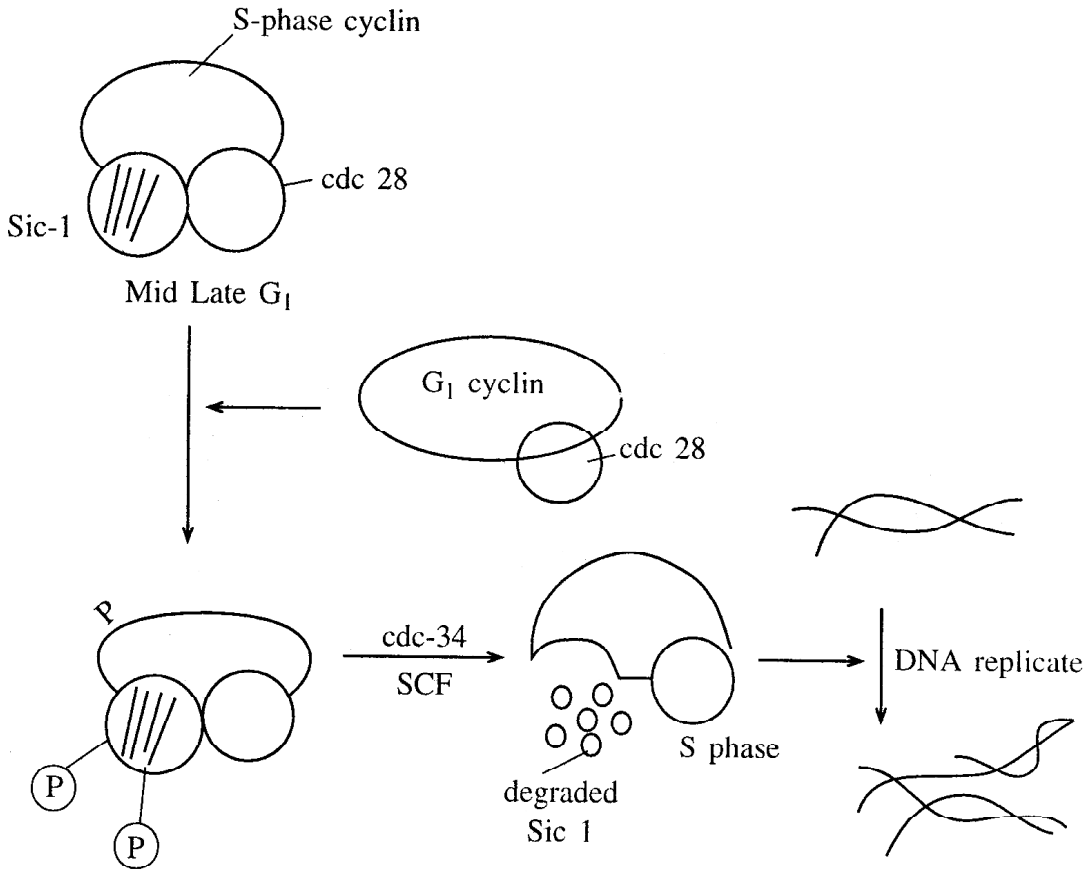


Fig. **12.2** The cell cycle in *S. cerevistae* consists of three cycles that separate after START and join before cytokinesis. Cells may be diverted into the mating pathway early in G1

(1) *S. cerevisiae* expresses a single cyclin dependent protein kinase (cdk) encoded by *cdc 28 gene* which interacts with several cyclins during different phases of the cell cycle.

(2) Just after cytokinesis, the decision on whether to initiate a division cycle is made before the START. The cells can diverted into mating type pathway by mating factors and *cdc 36* and *cdc 39* which appear to block the cell cycle before START and really function in the mating type pathway. Mutants block cell cycle by diverting cells into mating even in absence of the mating gene.

(3) After cytokinesis the mother ceil and the bud both the cell remain in the G_1 phase of the cell cycle. The ability to pass START is determined by environmental conditions. Prior to bud formation, spindle body duplication and DNA replication, the yeast cell exhausts its nutrients. When *S.cerevisiae* cells in G_1 have grown sufficiently in the growth medium, they begin a programme of



Fig, 12.3 Mechanism of Genetic regulation

gene expression that leads to entry into mitosis. Once G_1 cells reach the critical size, they become committed to completing cell cycle even if they are shifted to low nutrient medium.

(4) The crucial gene in pasing START is *cdc-28* which is homologns to *cdc-2* in *S.pombe.* Three cyclins are active in G,. CIn 1, Cln 2 and Cln 3 are encoded by *CLN-},* CLN-2, *CLN-3* respectively. Mutations in any one or two of these genes fail to block the cell cycle, thus the *CLN* genes are functionally redundant.

(5) The complexes formed between cdc 28 and the three G_1cyclins (Cln 1, Cln 2 and Cln 3) have protein kinase activity and constitute the hypothesized S phase promoting factor (SPF_S). In wild type yeast cells, Cln 3 is expressed at a nearly constant level throughout the cell cycle. Cln 1 and Cln 2 are expressed during the second half of G_1 and they increase rapidly and when their accumulation exceeded a critical threshold level, triggers the passage of *start* (START) into S-phase. After that its concentration declines gradually and are eliminated by the time of mitosis.

(6) cdc 28-Cln 3 phosphorylates and activates SBF and MBF. These induce transcription of *CLN 1* and *CLN 2* genes as well as several other genes required for DNA replication, including genes encoding DNA polymerase, RPA (ssDNA binding proteins), DNA ligase and certain enzymes acquired for deaxyribonucleoside triphosphate synthesis.

(7) cdc 28-Cln 1 and ede 28-Cln 2 phosphorylate APC in late G_1 and inactivate it.

(8) Two B-type cyclin genes *CLB 5* and *CLB 6* are also regulated by MBF and transcribed beginning in late Gj. The corresponding proteins Clb 5 and Clb 6 accumulate because of the inactivation of APC.

(9) At late G_1 ede 28-Clb 5 and ede 28-Clb 6 heterodimers accumulate and are inactivated by Sic-1 (an S-phase inhibitor), but it has no effect on ede 28-Cln complexes. Sic 1 is degraded falling polyubiquitination by E_2 associated with E_3. Once Sic is degraded, ede 28-Clb 5 and ede 28-Clb 6 kinases induce DNA replication.

(10) Initiation of DNA replication needs both assembly of pre replication complex and an active cdc 28-Cln complex. A second heterodimeric protein kinase ede 7-Dbf 4, which is expressed in G_1 is also required to trigger initiation. Once replication has initiated, Mem proteins and ede 45 move away from the origin along with DNA polymerases. Mem proteins are homologous to helicase, associated with replication fork movements.

(11) Later in S-phase, transcription of the genes *CLB 3* and *CLB 4* begins, encoding two additional B-type cyclins, Clb 3 and Clb 4, which also form

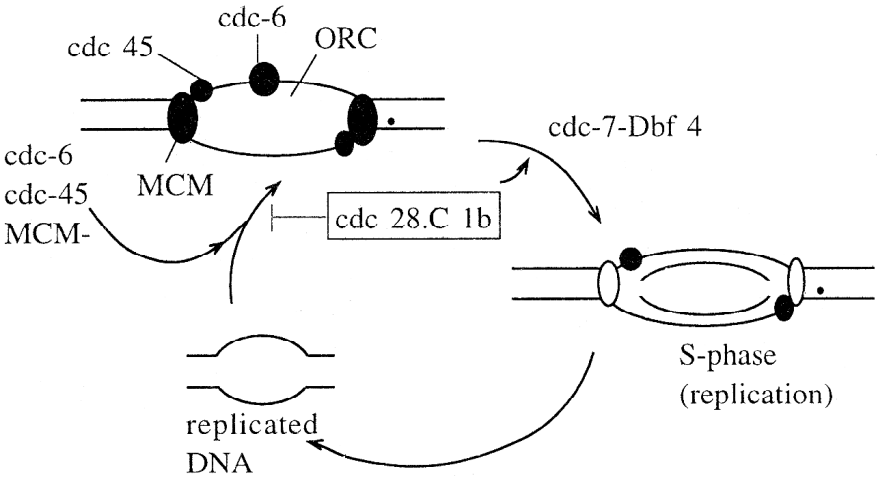


Fig. 12.4 Role of cdc in replication

heterodimeric protein kinases with cdc 28. These two cdc 28-Clb 3 and cdc 28-Clb 4 complexes also initiate the formation of mitotic spindle at the beginning of mitosis.

(12) As cells complete chromosome replication and enter G_2, two more B-cyclins are expressed. These are Clb 1 and Clb 2 encoded by *CLB 1* and *CLB 2* genes. These function as mitotic cyclins, associating with cdc 28 to form complexes that arc required for chromosome segregation and nuclear division.

12.2 Molecular basis of cellular checkpoints

12.2.1 Introduction

In most cells there are several points in the cell cycle, called *checkpoints* in which the cycle can be arrested if previous events have not been completed. Four checkpoint control can arrest the passage through cell cycle. These are (a) Gj-arrest due to DNA damage, (b) S-arrest due to unreplicated DNA, (c) C_2-arrest due to DNA damage, (d) M-arrest due to improper spindle formation.

In the checkpoints, the control system can be regulated by extra cellular signals from other cells. These signals either promote or inhibit cell proliferation. Checkpoints generally operate through negative intracellular signals.

12.2.2 The DNA replication checkpoint

Most of the cells by DNA replication checkpoint mechanism, avoided entry into cell division until the last nucleotide in the genome has been copied. The

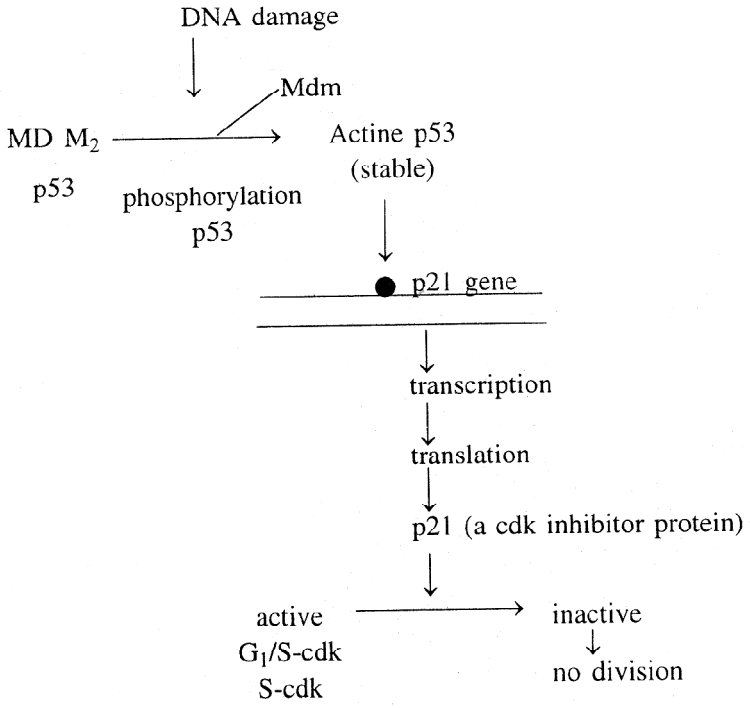


Fig. 12.5(a) Mechanism of G_1 arrest checkpoint

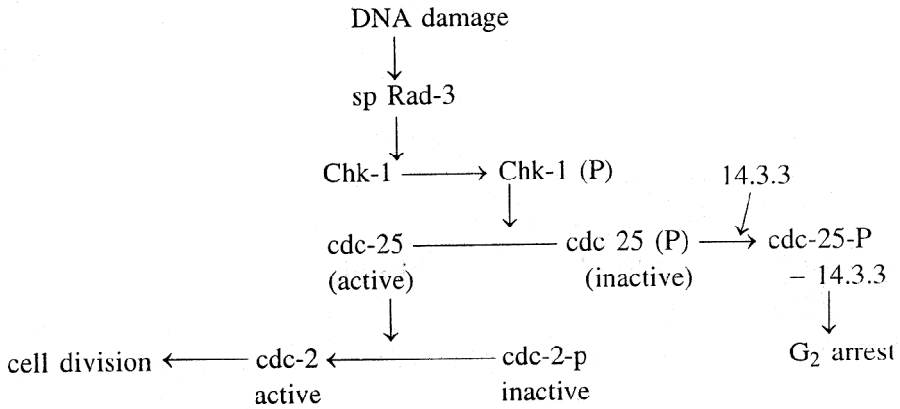


Fig. 12.S(b) Mechanism of G_2 arrest checkpoint

molecular mechanism has not been discovered but any signal from unreplicated DNA or unfinished replication forks send a negative signal to the cell cycle control system that blocks the activation of M-Cdk. Thus normal cells treated with chemical inhibitors of DNA synthesis viz., hydroxyurea, do not progress

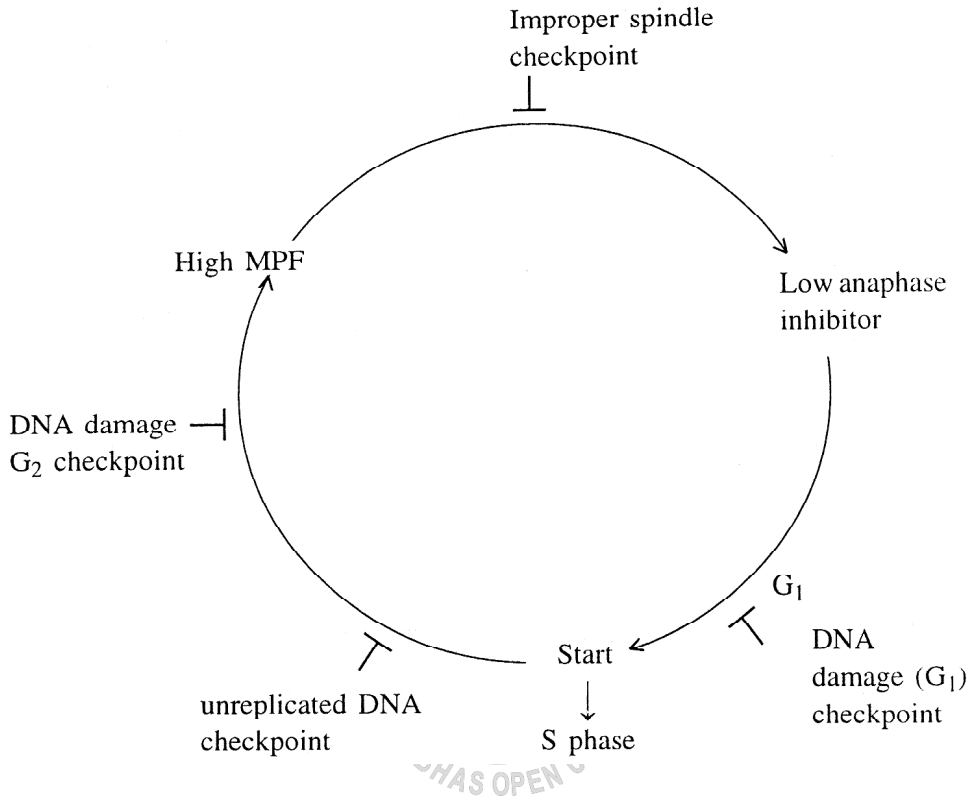


Fig. 12.5(c) DNA replication checkpoint

into mitosis. The block activates a checkpoint mechanism that arrests the cells in S-phase, delaying mitosis. But if caffeine is added along with hydroxyurea checkpoint mechanism fails and the cells proceed into mitosis according to their normal schedule with incompletely replicated DNA. As a result, the cells die.

12.2.3 The spindle attachment checkpoint

The effect of colchicin which inhibits spindle assembly, shows the presence of this checkpoint. In most cell types, a spindle attachment checkpoint mechanism operates to ensure that all chromosomes are properly attached to the spindle before sister chromatid separation occurs. During metaphase kinetochore regions of the chromosomes are attached to microtubules and any kinetochore that is not properly attached to the spindle, sends out a negative wait signal to the cell cycle control system that blocks ede 20-APC activation which is needed for sister chromatid separation. The nature of signal is not clear. But it has been seen that several proteins, including Mad 2 are recruited to unattached kinetochores which

are required for spindle attachment checkpoint to function. Even a single unattached kinetochore results inhibition of cdc 20-APC activation by binding with Mad 2.

In mice **MAD** 2 and BUB 1 and in humans **MAD** 2 genes have been recently identified. MAD 2 protein remains concentrated at the kinetochore until completion of microtubules attachment. This protein is continuously migrating into the cytoplasm and broadcasting signal throughout the cytoplasm. In mammals MAD 2 is associated with p^{55} CDC but in budding yeast MAD 2 is with cdc 20 and in fission yeast Mad 2-slp 1 (sip = sleepy) form a large complex with APC. Hct 1, another protein is associated with APC. APC is kept in check by MAD 2 and when APC is active the cell initiate anaphase by catalizing degradation of the two proteins with the help of Hct and cdc 20.

cdc 20/slp 1 promotes degradation of CUT2 but the activity of cdc 20/slp 1 is inhibited by MAD 2. Hct 1 promotes degradation of CLB-1, APC catalyzes this degradation by ubiquitination. This ubiquitination is inhibited by spindle attachment checkpoints.

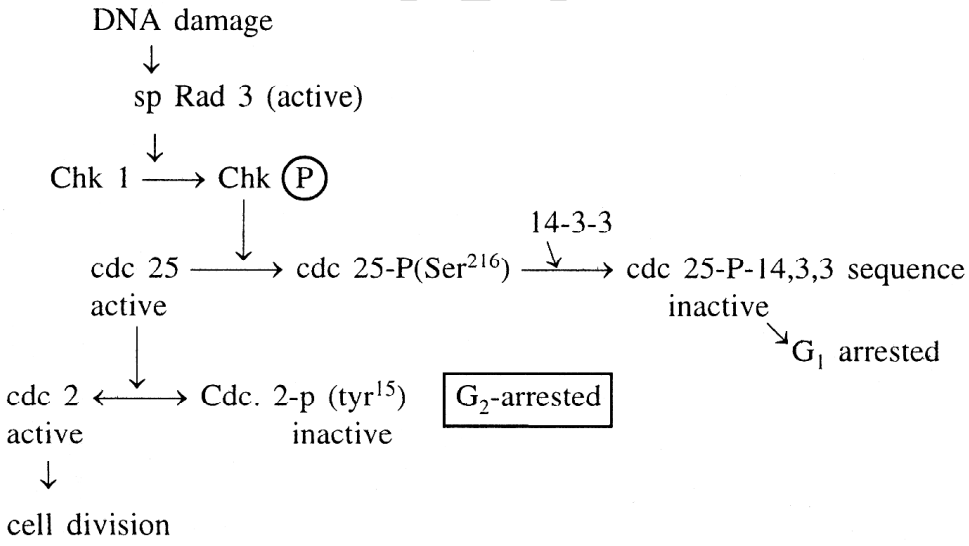


Fig. 12.6 DNA damage checkpoint

In mammals BUB 1 and BUB 3 proteins along with CNEP-E protein induce changes that lastly shut down the transmission of wait signal. MAD 2 dissociates from p^{55} cdc and loses its control on APC. APC is activated and catalyzes breakdown of cyclins and facilitates anaphase separation of chromatid.

12.2.4 DNA damage check points

The cell cycle control system can readily detect DNA *damage* and *arrest* the cell cycle at DNA damage checkpoints. These two checkpoints are one is in late G\ which prevents entry into S-phase and the other is in late G_2 which prevents entry into, mitosis.

G_2 checkpoint depends on a (*similar*) mechanism that delays entry into mitosis in response to incomplete DNA replication. The damaged DNA sends a signal to a series of protein kinases that phosphorylate and inactivate the phosphatase-cdc 25. This blocks the dephosphorylation and activation of M-cdk, thereby blocking entry into mitosis when the damaged DNA is repaired, the inhibitory signal is turned off and the cell division continues.

In the yeast *S.pombe,* the checkpoints sense the DNA damage and transduce inhibitory signal. Four genes including *RAD 9,* RAD 17, *RAD 24* and MEC 3 will sense this damage. The model proposed that DNA damage activates a protein sp Rad 3 (sp prefix means *S. pombe).* It brings about phosphorylation of Chk 1 and Chk 1-Ⓟ functions as a kinase, which brings about phosphorylation at Ser^{216} of Cdc 25. It then promotes binding of cdc 25 to a protein 14-3-3, coded by *rad 24* and *rad 25,* leading to sequestration of cdc 25. *Then cdc 25 is not available for activation of cdc 2 P (tyr^{15}) and the cell division is arrested at G_2-M transition.*

G_1 checkpoint blocks progression into S-phase by inhibition of G_1S-Cdk and S-Cdk complex.

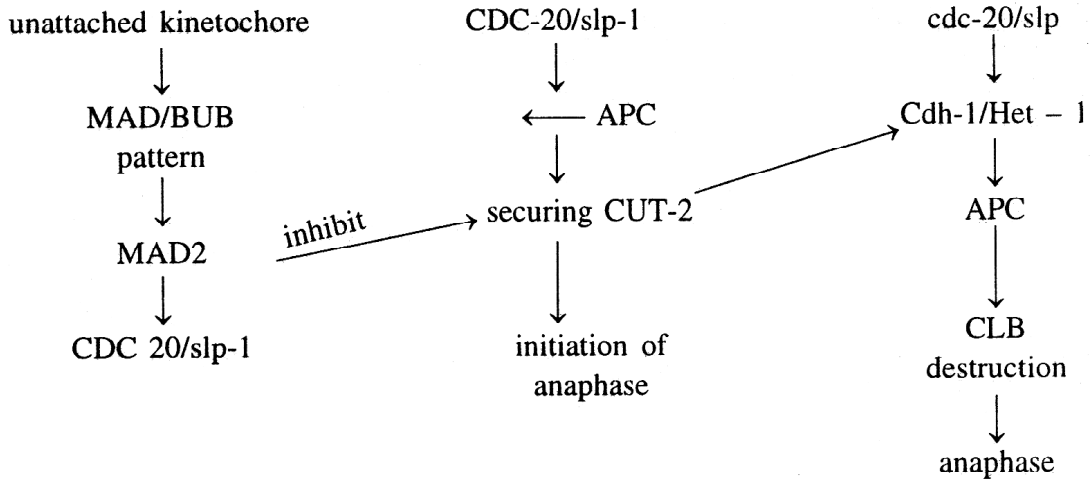


Fig. 12.7 Spindle attachment checkpoint

(1) In mammalian cells a gene regulatory protein p^{53} is being activated by DNA damage and it stimulates transcription of many genes.

(2) A CKI protein, called p^{21} is encoded by such activated genes which binds to G_1/S-Cdk and S-Cdk and inhibits their activity and thus blocks entry into mitosis.

(3) Actually in an undamaged cell, p^{53} is very unstable and is present at a low concentration. It interacts with another protein Mdm2, that causes destruction of p^{53} by ubiquitination mechanism. Damaged DNA activates protein kinase, phosphorylates p^{53} and reduce the binding of Mdm2. As a result p^{53} concentration rises and stimulates gene transcription of p^{21} which inactivates G_1/S-Cdk and S-Cdk activity.

(4) A rare genetic disease, Ataxia telangiectasia is caused by a defect in one of the protein kinases that phosphorylates and activates p^{53} in response to radiation and due to the loss of the DNA damage checkpoints, they suffer from increased rate of cancer.

GROUP
B (II)
Molecular Biology

Unit 1 History and Scope of Molecular Biology

Molecular Biology is one of the most rapidly growing domains of life sciences in the 21st century. Though the subject is relatively new in comparison to the specialized areas like—cytology, bacteriology, morphology and embryology; it has a history which is interesting and worth mentioning. Antony van Leeuwenhoek's invention of the microscope in around 1650 opened up the micro-world of biology. Advancement of knowledge in the fields of biochemistry, microbiology, virology and genetics in the early part of the 20th century gave rise to a new domain of activity which also attracted the attention of chemists and physicists. The remarkable development in physics and technological advancement opened up new frontiers in the biological arena one of which came to be known as **Molecular Biology** - the term was coined by **Warren Weaver** of the Rockefeller Foundation. Molecular biology attempts to explain the phenomena of life and its evolution, starting from the macromolecular properties that generate them. The two main macromolecules that contribute in the life process and remains the focus of the molecular biologist are the nucleic acids, (DNA & RNA)- the constituent of genes, and proteins, which are the active ingredients of life processes.

Charles Darwin's publication of **On *the Origin of Species*** (1859) had a profound impact on biological thinking. The rediscovery of the work of the Austrian monk Gregor Mendel probably might have been the first impetus to the development of molecular biology. August Weismann's description of reduction division in 1'887 followed by the description of meiosis, spermatogenesis by Walter Sutton (1903) lead to the proposal 'Chromosomal Theory of Inheritance". Acceptance of chromosomal theory of inheritance by 1935 became the stepping stone toward the journey for the search of chemical and physical nature of hereditary factors.

Several important developments took place during the early part of the 20th century that suggested the existence of genes within the chromosomes. Thomas Hunt Morgan (1909) showed that phenotypic change in *Drosophila* is linked to the events of crossing over in the chromosomes while Alfred Sturtevant (1910) moved a step ahead and mapped gene on chromosomes. Although existence of a new substance called nuclein was identified in the sperm cells by Friedrich Meischer in 1869, which later came to be known as deoxyribonucleic acid (DNA), it was Robert Fuelgen (1924) who first showed the coexistence of DNA along with proteins in

chromosomes by cytochemical staining. Appreciation of the nucleic acids quickly led to findings that there are two types of nucleic acids- RNA and DNA that differed in their sugar moieties. In 1929 Phoebus Levene showed that there are four types of DNA molecules - each of which he referred to as nucleotide. Each nucleotide had a deoxy-ribose sugar unit, a phosphate group and a nitrogenous base. The four nitrogenous bases were identified as Adenine, Guanine, Thiamine and Cytosine. He also suggested that the nucleotides are linked together through their phosphate sugar back bone but he made a mistake by considering the nucleotides to be present in short sequences and that the bases repeated in the same fixed order. However, the DNA molecule exists as a polymer was confirmed by Torbjorn Caspersson and Einar Hammersten (1934).

During this time Fred Griffith (1928) used pneumoccocus to describe gene transformation and George Beadle & Edward Tatum came up with "one gene - one enzyme" theory to demonstrated the existence of a precise relationship between genes and proteins. Following these discoveries, numerous research groups confirmed the importance of the gene in the life and development of organisms. It became apparently clear that genes are present in chromosomes and chromosomes are made of DNA and proteins. Initially the scientific community could not consider DNA as a hereditary material because of its utter simplicity. Rather, complex nature of the proteins was a preferred candidate to store hereditary information. Nevertheless, the chemical nature of genes and their mechanisms of action remained a mystery. Molecular biologists committed themselves to the determination of the structure of gene and the description of the complex relations between, genes and proteins. In the 50's, two important events took place almost simultaneously. Oswald Avery, Maclyn MacLeod and Colin McCarty in 1944 for the first time came out with the suggestion that genes are made up of DNA. Although their work was severely criticized by the scientific community then, Alfred Hershey and Martha Chase (1952) confirmed through their ingenious experiment that the genetic material of the bacteriophage is DNA.

Then came the discovery that revolutionized the world of science. Watson and Crick (1953) published their paper entitled "Molecular structure of Nucleic Acids" *Nature* 171, 737-738 (1953) where they gave a detail account of the double helix structure of the DNA molecule. The discovery of the double helical structure is history by itself. In brief there were three groups working to elucidate the structure of the DNA molecule. The first group worked at King's College, London and was led by Maurice Wilkins and was later joined by Rosalind Franklin. The second group working on DNA was Francis Crick and James D. Watson was at Cambridge

and the third group was at Caltech where the noble laureate Linus Pauling was leading the show. Inspiration and data from the works of Erwin ChargafPs work, published in 1947 and the X-ray diffraction patterns of DNA fibers produced by Maurice Wilkins and Rosalind Franklin at King's College facilitated Watson and Crick to design the double helix model using metal rods and balls in which they incorporated the known chemical structures of the nucleotides, as well as the known position of the linkages joining one nucleotide to the next along the polymer. For their pioneering work, Watson, Crick and Wilkins were awarded Nobel Prize in physiology in the year 1962. In their paper on the structure of DNA double helix, Watson and Crick wrote at the end of the paper, 'It has not escaped our notice that the specific pairing we have postulated immediately suggests a possible copying mechanism for the genetic material'—a path breaking thought that allows the preservation of hereditary information from generation to generation.

New thought processes began to evolve centering round the DNA molecule. Subsequently, Crick in 1957 proposed the 'Central Dogma' that explains the flow information of genes to proteins and the relationship between DNA, RNA and proteins. Controversies arose regarding the replication mechanism. However, ingenious experiment by Meselson-Stahl put an end to all controversies by showing the DNA replication is semi conservative in nature. Crick and coworkers showed that the genetic code was based on non-overlapping triplets of bases, called codons, and Har Gobind Khorana (1961) and others deciphered the genetic code not long afterward. These findings represent the birth of molecular biology which is still advancing at rapid pace everyday adding new information and creating history.

Following the deciphering of the genetic code, Arthur Kornberg described the action of DNA polymerase; Stanley Cohen (1968) discovered plasmids and antibiotic resistance gene. These findings gave birth to recombinant DNA technology which was immediately followed by the development of DNA sequencing technique by Walter Gilbert, allan Maxam, Fred Sanger in the year 1975. Simultaneously, Ceasar Milstein, Geo Kholer and Niles Jeme developed the art of monoclonal antibody production. Split gene concept was put forward by Richard Roberts and Philip Sharp in 1977 that marked the difference between the eukaryotic and prokaryotic genome. Another giant leap was made in the field of molecular biology when Kary Mullis (1985) successfully synthesized DNA polymers *in vitro* - a technique better known as Polymerase Chain reaction (PCR). Scientist then dared to sequence the 3 billion nucleotides present in human and launched the flamboyant project "Human Genome Project" in the year 1989 with a

target to complete the entire human sequence by 2010. Thanks to the technological development and discovery of efficient DNA polymerases that enabled to complete the project in the year 2003. The history does not end here. New branches like proteomics and genomics have come up to understand the functioning of cellular genes and to utilize the information for betterment of the man kind.

Unit 2 DNA Replication

Structure

2.1 Introduction

The fundamental biological process of reproduction requires the faithful transmission of genetic information from parent to offspring. Genetic information is stored in the form of an array of nucleotide sequences. The life process has evolved mechanisms to replicate the array of nucleotide sequences with great accuracy, that too at an astounding speed. The single, circular chromosome of *E. coli* contains about 4.7 million base pairs. Duplicating at a rate of more than 1000 nucleotides per minute, replication of the entire chromosome would require almost 3 days. Yet, these bacteria are capable of dividing every 20 minutes making minimum errors during the replication process. A huge amount of genetic information and an enormous number of cell divisions are required to produce a multicellular adult organism; even a low rate of error during copying would be catastrophic. Interestingly, mechanism to correct base pairing mistakes that occur during the process of replication has evolved in both prokaryotes and eukaryotes.

In 1953, Watson and Crick wrote at the end of their paper on the structure of DNA double helix, 'It has not escaped our notice that the specific pairing we have postulated immediately suggests a possible copying mechanism for the genetic material'. They recognized and explained that an inherent copying mechanism exists in the double helix DNA molecule. During replication, the two strands of the DNA helix unwind and unzip; each strand serves as a template for the new DNA molecule to be synthesized on it - a process termed as **semi conservative** replication.

2.2 Semiconservative replication

Semi conservative replication hypothesize that during replication of DNA double helix the strands unwinds and each strand serves as a template on which

new daughter strands are synthesized following the base paring rule. That is, an **'A'** at one position on the mother strand signals the addition of a **'T'** on the corresponding position on the newly forming strand. Similarly the presence of **'G'** on the mother strand will signal the addition of **'C'** on the daughter strand. This type of base pairing, which determine the nucieotide sequence of the new strand is known as **complementary base pairing.** Once the bases are aligned, DNA polymerase enzyme link the new incoming nucieotide with the previous aligned nucieotide of the daughter strand and eventually hydrogen bond is established between the base pairs. The process continues until the entire length of the mother DNA strand is copied. The newly formed DNA helix comprises of a new strand and an old original strand (conserved strand). Such pattern of DNA double helix duplication is called **semiconservative** replication. However, alternative mechanisms of replication were proposed which are known as **conservative replication** and **dispersive replication** (Fig. 2.1).

According to conservative replication, the original two strands serve as templates for the formation of new DNA strands. But, one of the double helix would consist entirely of original DNA strands, while the other helix would consists of two newly synthesized strands.

Dispersive replication suggests that the two DNA helix formed after replication comprises of interspersed blocks of new and old strands. However, evidences of such type of replication are still lacking.

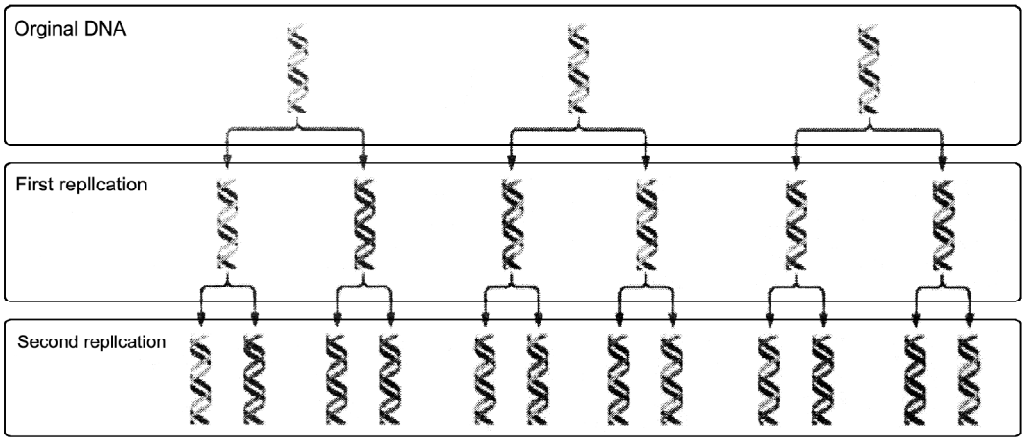


Fig. 2.1 : Proposed mechanisms of DNA replication: Semiconservative, Conservative and Dispersive

2.2.1 Experimental evidence of semiconservative replication

In 1958, M. Meselson and F. Stahl performed an imaginative experiment that confirmed the semiconservative nature of DNA replication. They grew *E. coti* in a medium containing a heavy isotope of nitrogen, ^{I5}N for several generations. After

growing for several generations in ^{15}N medium, practically all nitrogen atoms in the DNA of the bacterial cells were labeled with ^{15}N. They then transferred some of the cells from ^{15}N medium to new medium in which the nitrogen was ^{14}N. The bacteria were allowed to divide for one cycle only. Similarly, in another tube, the cells transferred to ^{14}N medium and were allowed to divide for two generation only. Any DNA synthesized after the transfer would contain a mixture of light & heavy isotopes. The newly cultured cells were then isolated, their DNA extracted and subjected to **equilibrium density gradient centrifugation** in Cesium chloride gradient to determine the density of the respective DNAs. (Cesium chloride [CsCl] centri-fuged at 50000 rpm, 250,000g for 2 days that produce the gradient).

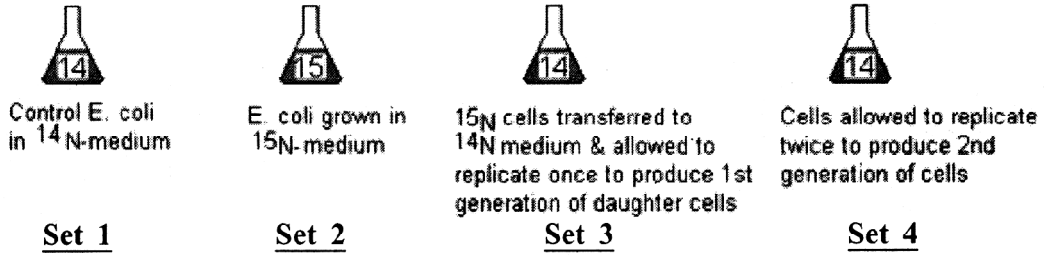


Fig. 2.2: *E. coli* grown in medium having either ^{15}N or ^{14}N as the only source for nitrogen for DNA replication

When DNA from each of these sets of cells was prepared and centrifuged in a cesium chloride gradient, the DNA formed bands depending on their respective density (**Fig. 2.3).**

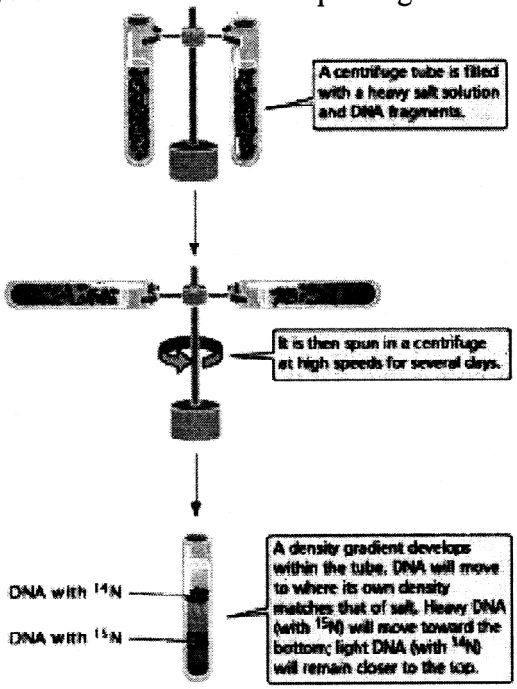


Fig. 2.3 Meselson and Stahl used equilibrium density gradient centrifugation

$^{I5}N^{I5}N$ -DNA being heavier, they formed band at bottom end of the tube **(Set 2).** $^{15}N^{14}N$-DNA produced after one round of cell division in ^{I4}N containing medium was lighter than the ^{I5}N type DNA and formed a band above the $^{15}N^{I5}N$ bands **(Set 3)**. In accordance with the semi conservative replication, the cells that were allowed to divide twice in ^{I4}N- medium should have a mixture of $^{I5}N^{14}N$ & $^{I4}N^{14}N$ DNA and should form two bands in CsCl gradient. As per prediction, the cells from the **Set 4** produced two bands. The lower band corresponded with the $^{15}N^{I4}N$ band observed in **Set 3** and the upper band corresponded to the $^{I4}N^{14}N$ band of Set 1 **(Fig. 2.4).**

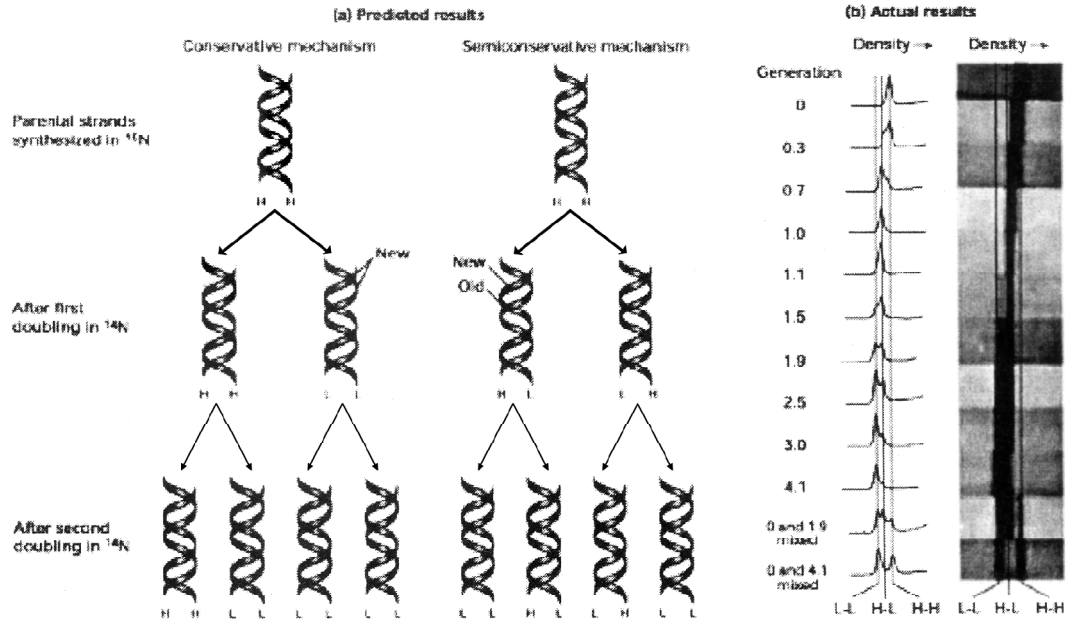


Fig. 2.4: Meselson and Stahl's experiment showing that DNA replication is semiconservati ve.

This classic experiment confirmed the prediction of the semi conservative mode of replication envisaged by Watson and Crick and disapproved all notions of conservative and dispersive models for DNA replication. Later, autoradiographic study by J. Cairns (1963) on bacterial DNA replication confirmed the observation of Meselson and Stahl (1958). The study also elucidated the circular nature of bacterial DNA and showed that DNA replication occurs simultaneously on both the strands at one or two moving 'Y' shaped forked junctions in the circular DNA. There are, however, several different ways that semiconservati ve replication can take place, differing principally in the nature of the template DNA—whether it is linear or circular— and in the number of replication forks. Individual units of

replication are called **replicons,** each of which contains a **replication origin.** Replication starts at the origin and continues until the entire replicon has been replicated. Bacterial chromosomes have a single replication origin, whereas eukaryotic chromosomes contain many.

2.2.2 Theta replication:

A common type of replication that takes place in circular DNA, such as that found in *E. coli* and other bacteria, is called **theta replication (Fig. 2.5),** because it generates a structure that resembles the Greek letter theta (e). In theta replication, double-stranded DNA begins to unwind at the replication origin, producing single-stranded nucleotide strands that then serve as templates on which new DNA can be synthesized. The unwinding of the double helix generates a loop, termed a **replication bubble.** Unwinding may be at one or both ends of the bubble, making it progressively larger. DNA replication on both of the template strands is simultaneous with unwinding. The point of unwinding, where the two single nucleotide strands separate from the double-stranded DNA helix, is called a **replication fork.** If there are two replication forks, one at each end of the replication bubble, the forks proceed outward in both directions in a process called **bidirectional replication,** simultaneously unwinding and replicating the DNA until they eventually meet. If a single replication fork is present, it proceeds around the entire circle to produce two complete circular DNA molecules, each consisting of one old and one new nucleotide strand.

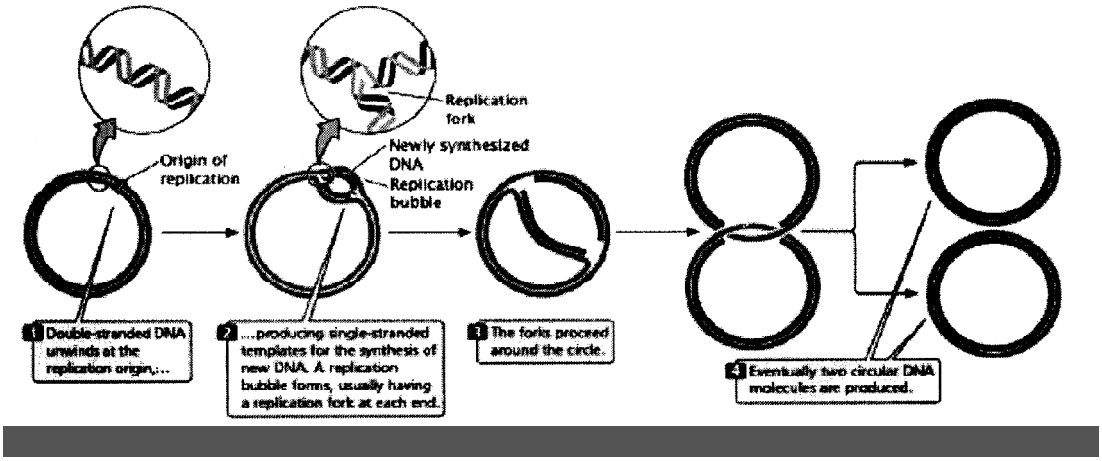


Fig. 2.5a. Theta replication in *E.coli* and other organisms possessing circular DNA

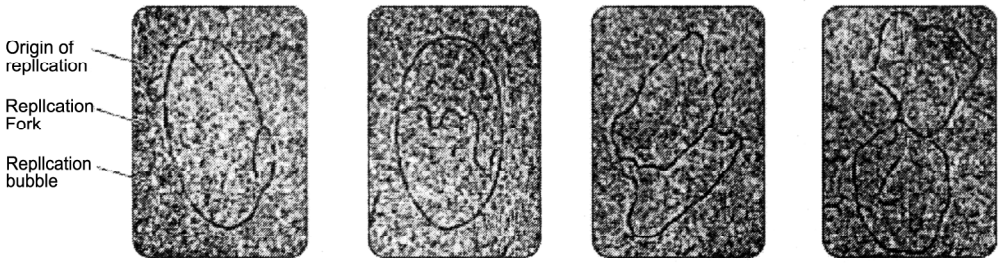


Fig. 2.5b: Experimental evidence produced by J. Cairns (1963) to show the Theta mode of replication in *E.coli*

Rolling-circle replication takes place in some viruses and in the F factors of *E. coli.* This form of replication is initiated by a break in one of the nucleotide strands that creates a 3'-OH group and a 5'-phosphate group. New nucleotides are added to the 3' end of the broken strand, with the inner (unbroken) strand used as a template. As new nucleotides are added to the 3' end, the 5' end of the broken strand is displaced from the template, rolling out like thread being pulled off a spool. The 3' end grows around the circle, giving rise to the name rolling-circle model. The replication fork may continue around the circle a number of times, producing several linked copies of the same sequence. With each revolution around the circle, the growing 3' end displaces the nucleotide strand synthesized in the preceding revolution. Eventually, the linear DNA molecule is cleaved from the circle, resulting in a doublestranded circular DNA molecule and a single-stranded linear DNA molecule. The linear molecule circularizes either before or after serving as a template for the synthesis of a complementary strand (Fig. 2.6).

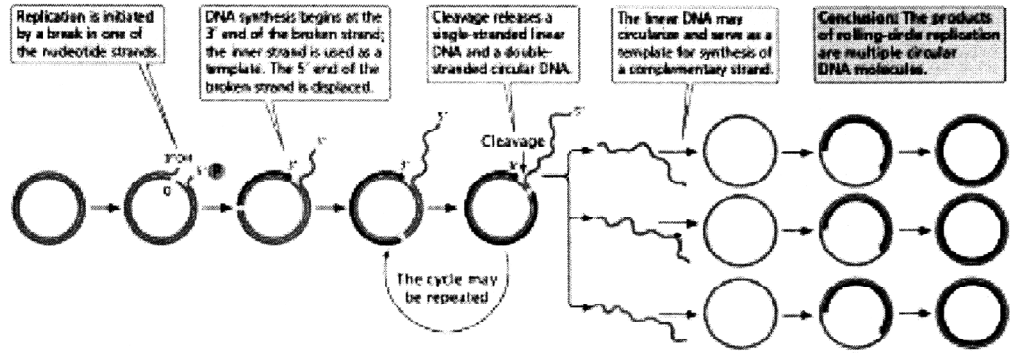
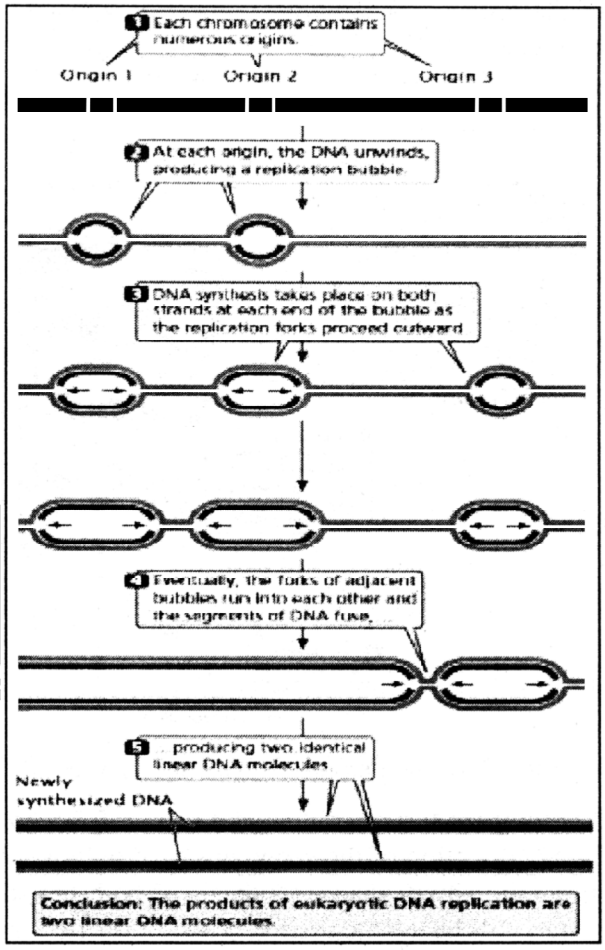


Fig. 2.6: The linear molecule serving as a template for the synthesis of a complementary strand.

2.2.3 Linear eukaryotic replication :

The large linear chromosomes in eukaryotic cells contain too much DNA and need to replicate speedily within a reasonable time frame and therefore cannot afford to have a single origin of replication fork as found in bacteria. Multiple

replication origin point exists in eukaryotic that proceed at a rate ranging from 500 to 5000 nucleotides per minute at each replication fork, considerably slower than bacterial replication but still replicate in a matter of minutes or hours, not days. This rate is possible because replication takes place simultaneously from thousands of origins. Typical eukaryotic replicons are-from 20,000 to 300,000 base pairs in length. At each replication origin, the DNA unwinds and produces a replication bubble. Replication takes place on both strands at each end of the bubble, with the two replication forks spreading outward. Eventually, replication forks of adjacent replicons run into each other, and the replicons fuse to form long stretches of newly synthesized DNA **(Fig. 2.7).** Replication and fusion of all the replicons leads to two identical DNA molecules.



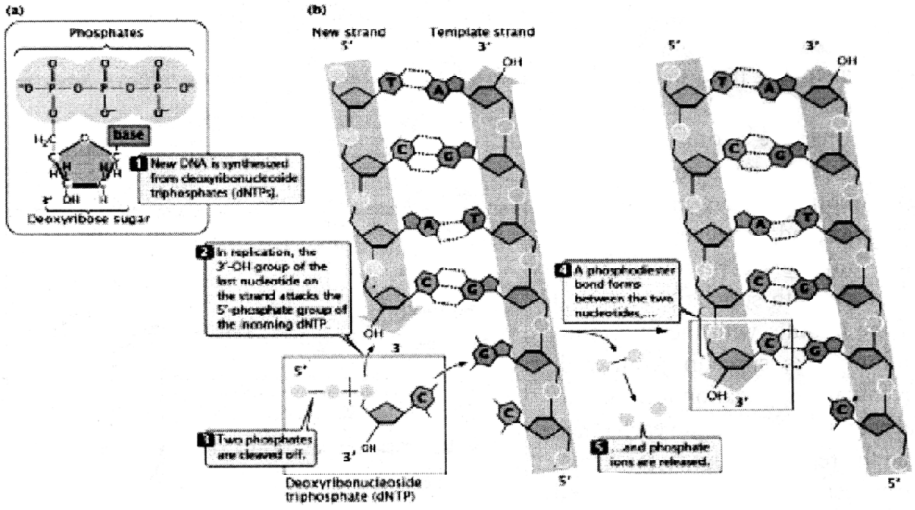
Fie. 2.7: DNA replication on linear chromosomes

Like all other metabolic processes, DNA replication is under the control of several proteins and enzymes, engaged in an intricate and coordinated interplay. Our understanding of DNA replication is primarily derived from physical, chemical and biochemical studies of enzymes and nucleic acids from *Es-cherichia coli,* their phages and their mutants. Prokaryotic and eukaryotic mechanism of DNA replication differs in many ways though the basic mechanisms are same.

2.3 DNA replication model

In the simplest model of DNA replication, the mother DNA strand is unzipped to produce a 'Y' shaped replication fork. At the replication fork, enzymes and protein factors facilitate the addition nascent nucleotides to the newly forming

DNA strand by way of complimentary base pairing and the event should occur simultaneously on both the strands. During DNA synthesis, nucleotides are added to the 3'-OH group of the growing nucleotide strand (Fig. 2.8). The 3'-OH group of the last nucleotide on the strand attacks the 5'-phosphate group of the incoming dNTP. Two phosphates are cleaved from the incoming dNTP, and a phosphodiester bond is created between the two nucleotides.



Fie: 2.8: Inclusion of a nascent nucleotide into a growing DNA molecule

But there lies a problem. All known DNA polymerases can add nucleotides in the 5'⟶3' direction of the growing strand only. As the DNA double helix is anti-parallel in nature, simultaneous synthesis of both the strands is difficult to conceive as one strand will be synthesized in 5'⟶3' direction and the other strand has to be synthesized in 3 '⟶5' if both the strands has to be synthesized simultaneously - which is not possible. Interestingly, mother-nature has evolved mechanism that allows both the strands of the DNA double helix to be synthesized simultaneously.

2.3.1 Continuous and discontinuous DNA replication

Auto radiographic studies confirmed that DNA synthesis occurs simultaneously on both the strands. The works of Okazaki and his colleagues enabled to explain the basic mechanism underlying the simultaneous synthesis of both the strands in DNA double helix. They studied the incorporation of radioactive thymidine at different phases of DNA synthesis and found that the radioactive materials were present only in short DNA fragments (100- 1000 nucleotides) extracted just a few moments after the radioactive pulse was inhibited. As the time

elapsed, the radioactive materials could be detected in high molecular weight DNA strands. Normally, this should not have happened as the feeding of radioactive material was stopped. They predicted that during DNA synthesis, continuous synthesis occurs on the 3'→5' template but on the 5'→3' template short DNA segments called Okazaki **fragments** are synthesized, which are subsequently linked together by the action of DNA polymerases. Thus, **continuous replication** occurs in one of the templates in the direction of the movement of the replication fork while **discontinuous replication** occurs in the other strand. The strand that allows continuous synthesis in known as **leading strand** and the strand on which discontinuous synthesis occurs is called the **lagging strand** (Fig. 2.9).

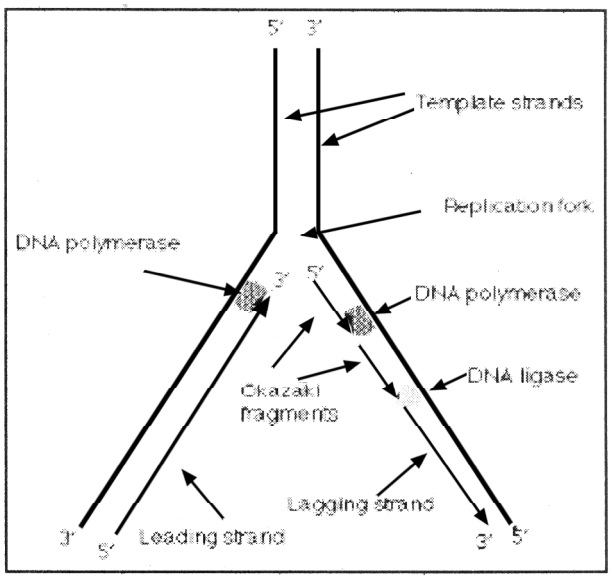


Fig: 2.9: DNA synthesis is continuous on one template strand of DNA and discontinuous on the other.

2.3.3 The fidelity of DNA replication

Overall, replication results in an error rate of less than one mistake per billion nucleotides. How is this incredible accuracy achieved? Answer lies in the activity of the DNA polymerase. These enzymes are very particular in pairing nucleotides with their complements on the template strand. Most of the errors that do arise in nucle-otide selection are corrected in a second process called **proofreading.** When a DNA polymerase inserts an incorrect nucleotide into the growing strand, the 3'-OH group of the mispaired nucleotide is not correctly positioned for accepting the next nucleotide. The incorrect positioning stalls the polymerization reaction, and the 3'→5' exonuclease activity of DNA polymerase removes the incorrectly paired nucleotide. DNA polymerase then inserts the correct nucleotide. Together, proofreading and nucleotide selection result in an error rate of only one in 10 million nucleotides. A third process, called **mismatch repair** corrects errors after replication is complete. Any incorrectly paired nucleotides produce a deformity in the secondary structure of the DNA; the deformity is recognized by specialized enzymes that excise an incorrectly paired nucleotide and replace it with the correct

nucleotide. Methylation on the old DNA strand allows the mismatch repair enzymes to distinguish between old and new strand.

2.3.3 Speed of replication :

The single molecule of DNA that is the *E. colt* genome contains 4.7 x 10^6 nucle-otide pairs. DNA replication begins at a single, fixed location in this molecule, the **replication origin,** proceeds at about 1000 nucleotides per second, and thus is done in no more than 40 minutes. And thanks to the precision of the process (which includes a "proof-reading" function), the job is done with only about one incorrect nucleotide for every 10^9 nucleotides inserted. In other words, more often than not, the *E. colt* genome (4.7 x 10^6) is copied without error.

2.4 Replication in eukaryotes

Our understanding of the DNA replication in eukaryotic cells is limited but development of new experimental techniques is rapidly changing the imbalance of knowledge. The replication machinery is similar to bacterial system but basic differences include: (1) replication on linear chromosomes associated with multiple proteins, (2) multiple replication origins in their chromosomes; (3) more types of DNA polymerases, with different functions; and (4) nucleosome assembly immediately following DNA replication. As the yeast *(Saccharomyces cerevisiae)* replication system is very similar to mammalian cells, isolation of various mutant yeasts, unable to produce specific gene products required for various aspects of replication has added to our understanding of eukaryotic replication. Further, the monkey virus **SV40** has single origin of replication where the viral encoded **large T antigen** binds along with several other proteins that are synthesized by the host DNA. Scientists utilize this knowledge to simulate DNA replication *in vitro* and understand the function of various cellular replication proteins.

2.4.1 Initiation of replication

In eukaryotes, cells have much more DNA and their polymerases synthesize DNA at a much lower rate. To compensate these difficulties, the eukaryotic cells replicate their genome in small portions, termed **replicons** (Fig. 2.10). From the radioactive studies it has been estimated that each replicon is approximately 15 to 100 urn in length (50 to 300 kb) and the replication fork proceeds in both direction. The heterochromatin regions tend to replicate late in the S phase as such the Barr body is the last to replicate while the active X chromosome is replicated at an early stage in females.

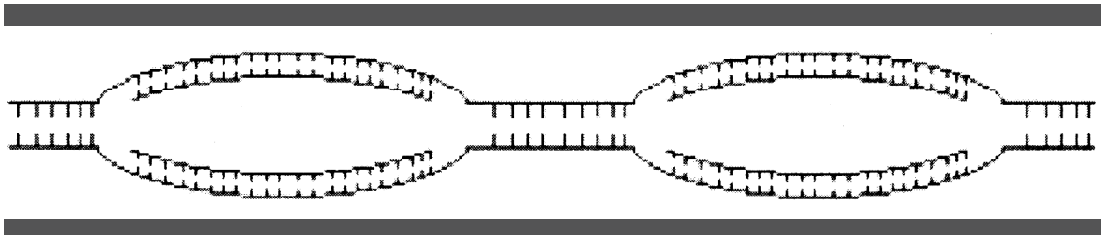


Fig: 2.10: Multiple site of origin of DNA replication and each is known is replicon

Initiation of replication in eukaryotes is much more complicated than in prokaryotes. In yeast cells, the site of origin of replication if removed and inserted in another DNA molecule, the hybrid DNA molecule acquires the ability to replicate *in vitro* or *in vivo.* As the sequences at the site of origin of replication promote replication of the DNA in which they are contained, they are referred to as **autonomous replicating sequences (ARSs).** There are about 400 ARSs scattered throughout the genome of yeast cells and each ARSs has a conserved 11 bp sequence that allow the binding of essential multiprotcin complex called the **origin recognition complex (ORC).** In lormal cells, ORC proteins remain bound to the ARSs all through the cell cycle. The binding of other proteins to the ORC-ARS complex allows initiation of replication. Mutated ARSs fail to bind ORC proteins and thus cannot initiate DNA replication at that site.

In mammals, virtually any type of purified naked DNA is suitable for initiation of replication **with** cellular extracts suggesting that, unlike yeast, mammalian DNA might not possess specific sites at which replication is initiated. However, in vivo studies on intact chromosomes indicate that replication does begin at specific sites along the DNA and initiation is not a random event. It appears that, mammalian DNA molecules have numerous sites where replication can be initiated, but because of the presence of nucleosome and higher order of organization of mammalian chromosome, most of the initiation sites remain suppressed while promoting initiation at specific sites that serve as replication origins. One such replication origin site has been located in the β-globin gene cluster.

As the Eukaryotic cells utilize thousands of origins, the cell needs to ensure that each segment of the DNA is replicated once during cell division. The precise replication of DNA is accomplished by the separation of the initiation of replication into two distinct steps. In the first step, the origins are licensed (see licensing factors), meaning that they are approved for replication. During replication, only the licensed sites can bind to replication initiation factors. The preliminary initiation factors at first displace the **replication licensing factors** and then induce the formation of replication bubble. The sites of origin of replication do not bind to any further replication licensing factors during the progression of the S phase until it enters

the mitotic phase thereby ensuring the replication of genome only once per cell cycle.

Two-replication fork are formed at each site of origin of replication and bidirectional DNA synthesis occur in a manner, which is similar in all organisms - whether it is virus, prokaryotes or eukaryotes. The replication forks are not randomly distributed in the nuclear matrix. There are 50-250 sites in the nucleus - called **replication foci,** where synthesis takes place. Each foci contain approximately active 40 replication forks. The clustering of replication forks may provide mechanisms for coordinating the replication of adjacent replicons over individual chromosomes. The nuclear matrix also seems to play an important during replication. The substances necessary for replication remain bound the nuclear matrix and are made available during the process of replication.

DNA replication in eukaryotic cells is limited to S phase of the cell cycle. Approximately 10^3 to 10^5 replication events occur in a coordinated manner, though not identically at all origins. This leads to great variation in the duration of S phase. Moreover, the associated histones and non histone proteins get synthesized either during.Gl or S phase.

Replication **'tool kit'** consists of helicase, single stranded DNA binding proteins, topoisomerases, primase, DNA polymerase, and DNA ligase. The DNA in eukaryotes is also synthesized in semi-discontinuous manner, although the Okazaki fragments of the lagging strand are much smaller (250 nucleotides in length).

2.4.2 Some proteins required for replication

The following table shows a glimple of different proteins required for replication.

DNA Polymerase Sub Units	Polymeras Activity	Exonuclease Activity	Cellular Function
α (alpha)	Yes	No	Initiation of nuclear DNA synthesis and DNA repair
β (beta)	Yes	No	DNA repair and recombination of nuclear DNA
γ (gamma)	Yes	Yes	Replication of mitochondrial DNA
δ (delta)	Yes	Yes	Leading & lagging-strand synthesis of nuclear DNA, DNA repair, and translesion DNA synthesis
ε (epsilon)	Yes	Yes	Unknown; probably repair and replication of nuclear DNA
ζ (zeta)	Yes	No	Translesion DNA synthesis
η (eta)	Yes	No	Translesion DNA synthesis

DNA Polymerase Sub Units	Polymerase Activity	Exonuclease Activity	Cellular Function
θ (theta)	Yes	No	DNA repair
ι (iota)	Yes	No	Translesion DNA synthesis
κ (kappa)	Yes	No	Translesion DNA synthesis
λ (lambda)	Yes	No	DNA repair
μ (mu)	Yes	No	DNA repair
σ (sigma)	Yes	No	Nuclear DNA replication (possibly), DNA repair, and sister-chromatid cohesion

Other DNA polynferases (ζ, η, θ, κ, λ, μ) allow replication to bypass damaged DNA (called translesion replication) or play a role in DNA repair. Many of the DNA polymerases have multiple roles in replication and DNA repair.

2.4.3 Eukaryotic DNA polymerase

Eukaryotic cells contain several DNA polymerases of which the most important are α,β,μ,γ,δ and ε (Table 1). **DNA polymerase** α, which contains primase activity, initiates nuclear DNA synthesis by synthesizing an RNA primer, followed by a short string of DNA nucleotides. After DNA polymerase α has laid down from 30 to 40 nucleotides, **DNA polymerase** δ completes replication on the leading and lagging strands. **DNA polymerase** β does not participate in replication but is associated with the repair and recombination of nuclear DNA. The γ polymerase is encoded by nuclear gene but located within the mitochondria and is responsible for the replication of the mitochondrial DNA;. γ polymerase like enzyme replicates chloroplast DNA in plants. Similar in structure and function to DNA polymerase δ , **DNA polymerase** e appears to take part in nuclear replication of both the leading and the lagging strands, but its precise role is not yet clear.

Inhibitors of polymerase activity : Aphidicolin (α, δ, & ε); N-ethylmalemide (α, β, δ, d); butylphenyl-dGTP (α,); dideoxynucleoside 5'triphosphate (γ).

2.4.4 The enzymes of DNA replication

1. **Topoisomerase** is responsible for initiation of the unwinding of the DNA. The tension holding the helix in its coiled and supercoiled structure can be broken by nicking a single strand of DNA. Try this with string. Twist two strings together, holding both the top and the bottom. If you cut only one of the two strings, the tension of the twisting is released and the strings untwist.

2. **Helicase** accomplishes unwinding of the original double strand, once supercoifing has been eliminated by the topoisomerase. The two strands very much want to bind together because of their hydrogen bonding affinity for each other, so the helicase activity requires energy (in the form of ATP) lo break the strands apart.
3. **DNA polymerase** proceeds along a single-stranded molecule of DNA, recruiting free dNTP's (deoxy-nucleotide-triphosphates) to hydrogen bond with their appropriate complementary dNTP on the single strand (A with T and G with C), and to form a covalent phosphodiester bond with the previous nucleotide of the same strand. The energy stored in the triphos-phate is used to covalently bind each new nucleotide to the growing second strand. There are different forms of DNA polymerase, but it is DNA polymerase III that is responsible for the processive synthesis of new DNA strands. DNA polymerase cannot start synthesizing de novo on a bare single strand. It needs a **primer** with a 3'OH group onto which it can attach a dNTP. DNA polymerase is actually an aggregate of several different protein subunits, so it is often called a holoenzyme. The holoenzyme also has proofreading activity, so that it can make sure that it inserted the right base, and nuclease (excision of nucleotides) activities so that it can cut away any mistakes it might have made.
4. **Primase** is actually part of an aggregate of proteins called the **primeosome.** This enzyme attaches a small RNA primer to the single-stranded DNA to act as a substitute 3'-OH for DNA polymerase to begin synthesizing from. This RNA primer is eventually removed by **RNase H** and the gap is filled in by **DNA polymerase I.**
5. **Ligase** can catalyze the formation of a phosphodiester bond given an unattached but adjacent 3'-OH and 5'phosphate. This can fill in the unattached gap left when the RNA primer is removed and filled in. The DNA polymerase can organize the bond on the 5' end of the primer, but ligase is needed to make the bond on the 3' end.
6. **Single-stranded binding proteins** (SSB) are important to maintain the stability of the replication fork. Single-stranded DNA is very labile, or unstable, so these proteins bind to it while it remains single stranded and keep it from being degraded.

2.4.5 The replication fork

Why can DNA polymerase only act from 5' to 3' The reason is the relative stability of each end of DNA. A triphosphate is required to provide energy for the bond between a newly attached nucleotide and the growing DNA strand. However,

this triphosphate is very unstable and can easily break into a monophosphate and an inorganic pyrophosphate, which floats away into cell. At the 5' end of the DNA, this triphosphate can easily break, so if a strand has been sitting in the cell for a while, it would not be able to attach new nucleotides to the 5' end once the phosphate had broken off. On the other hand, the 3' end only has a hydroxyl group, so as long as new nucleotide triphosphate are always brought by DNA polymerase, synthesis of a new strand can continue no matter how long the 3' end has remained free.

This presents a problem, since one strand of the double helix is 5' to 3', and the other one is 3' to 5'. How can DNA polymerase synthesize new copies of the 5' to 3^1 strand, if it can only travel in one direction? This strand is called the **lagging strand,** and DNA polymerase makes a second copy of this strand in spurts, called **Okazaki fragments,** as shown in the diagram. The other strand can proceed with synthesis directly, from 5' to 3', as the helix unwinds. This is the **leading strand.**

2.4.6 Nucleosome assembly

ukaryotic DNA is complexed to histone proteins in nucleosome structures that contribute to the stability and packing of the DNA molecule (see Fig. 2.11). The disassembly and reassembly of nucleosomes on newly synthesized DNA probably takes place during replication, but the precise mechanism for these processes has

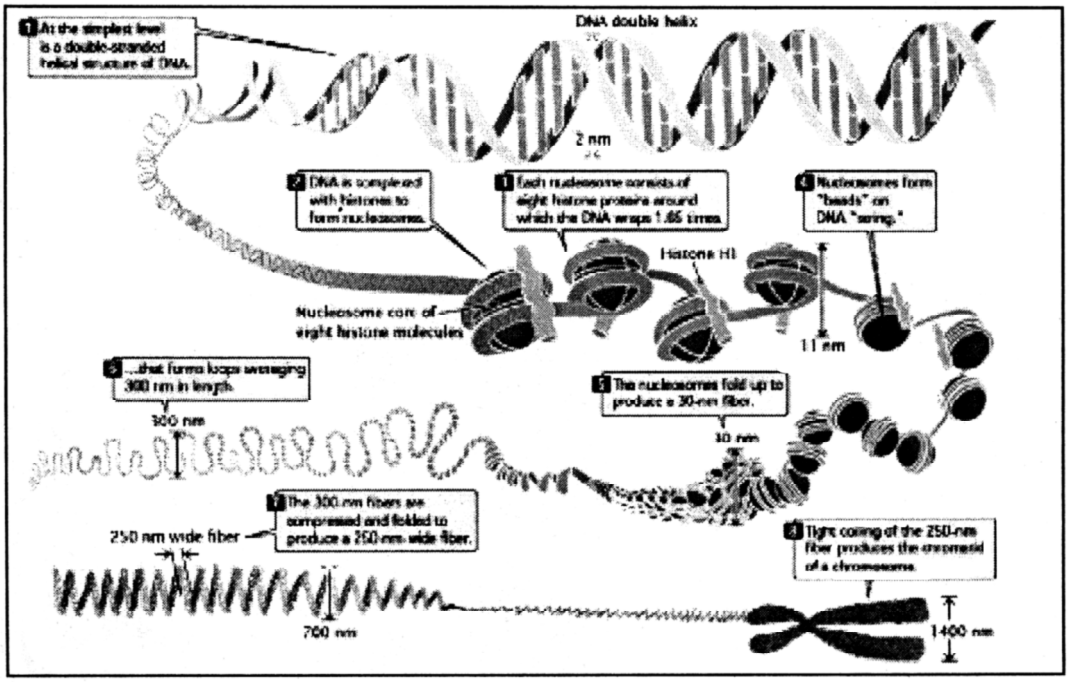


Fig. 2.11: Several levels of organization of eukaryotic chromosome.

not yet been determined. The unwinding of double stranded DNA and the assembly of the replication enzymes on the single-stranded templates probably require the disassembly of the nucleosome structure. Electron micrographs of eukaryotic DNA show recently replicated DNA already covered with nucleosomes indicating that nucleosome structure is reassembled quickly. Before replication, a single DNA molecule is associated with histone proteins. After replication and nucleosome assembly, two DNA molecules get associated with histone proteins. After replication and nucleosome assembly, two DNA molecules get associated with histone proteins. Whether the original histones remain together, attached to one of the new DNA molecules, or do they disassemble and mix with new histones on both DNA molecules is still not known. Experiments with radioactive labeled histones suggest that newly assembled octamers consist of a random mixture of old and new histones.

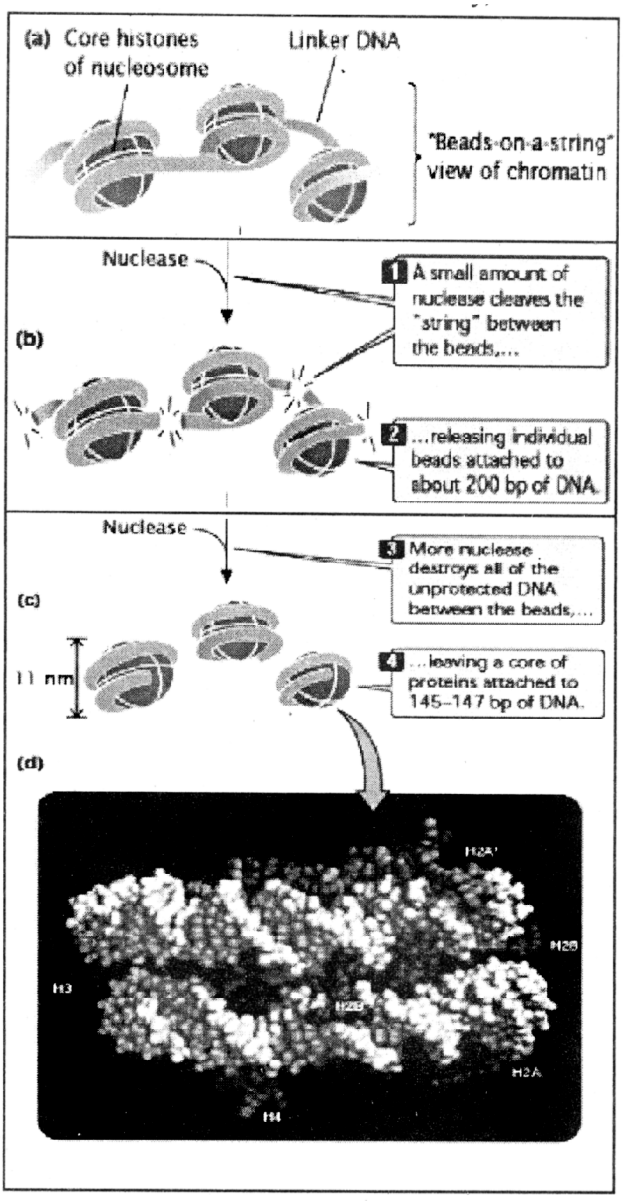


Fig. 2.12: The nucleosome model

2.4.7 The nucleosome

Chromatin has a highly complex structure with several levels of organization. The simplest level (Fig. 2.12) is the double helical structure as proposed by Watson and Crick (1953). At a more complex level, the DNA molecule is associated with proteins and is highly folded to produce a chromosome. Chromatins viewed under electron microscope, frequently looks like beads on a string. Partial digestion of chromatin with nuclease produces beads. Each individual bead has attached 200

bp of DNA. Further digestion with more nuclease chews up the entire DNA between the beads and leaves a core of proteins attached to a fragment of DNA (Fig. 2.12). These experiments demonstrated that chromatin is not a random association of proteins and DNA but has a fundamental repeating structure having the the simplest level of chromatin structure, the **nucleosome**

The nucleosome is a core particle consisting of DNA wrapped about two times around an octamer of eight histone proteins (two copies each of H2A, H2B, H3, and H4), much like thread wound around a spool (Fig. 2.12d). The DNA in direct contact with the histone octamer is between 145 and 147 bp in length, coils around the histones in a left-handed direction, and is supercoiled. It does not or kinks, in its helical structure as it winds around the histones.

The fifth type of histone, **H1,** is not a part of the core particle but plays an important role in the nucleosome structure. The precise location of H1 with respect to the core particle is still uncertain. The traditional view is that H1 sits outside the octamer and binds to the DNA where the DNA joins and leaves the octamer (Fig. 2.11). However, the results of recent experiments suggest that the H1 histone sits inside the coils of the nucleosome. Regardless of its position, H1 helps to lock the DNA into place, acting as a clamp around the nucleosome octamer. Together, the core particle and its associated H1 histone are called the **chromatosome,** the next level of chromatin organization. The H1 protein is attached to between 20 and 22 bp of DNA, and the nucleosome encompasses an additional 145 to 147 bp of DNA; so about 167 bp of DNA are held within the chromatosome. Chromatosomes are located at regular intervals along the DNA molecule and are separated from one another by wrap around the octamer smoothly; there are four bends, **linker DNA,** which varies in size among cell types—most cells have from about 30 bp to 40 bp of linker DNA. Nonhistone chromosomal proteins may be associated with this linker DNA, and a few also appear to bind directly to the core particle.

2.4.8 DNA synthesis at the ends of chromosomes

A fundamental difference between eukaryotic and bacterial replication arises because eukaryotic chromosomes are linear and thus have ends. As the 3'-OH group is needed by DNA polymerases to elongate, at the initiation of replication by RNA primers provide the 3 '-OH group synthesized by primase. RNA primers must be removed and replaced by DNA

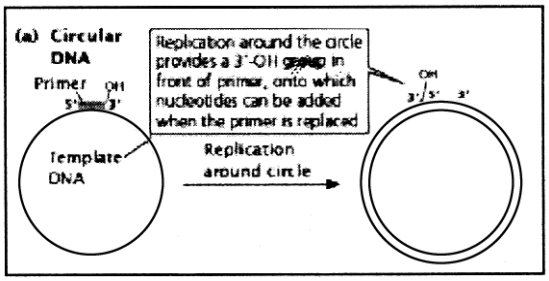


Fig. 13a: Replication at the ends of circular DNA where the 3'-OH group is available

Fig. 13a: Replication at the ends of circular DNA where the 3 '-OH group is available

mucleotides subsequently which is done by DNA polymerase I. In a circular DNA molecule, elongation around the circle eventually provides a 3'-OH group immediately in front of the primer (Fig. 2.13a). After the primer has been removed, the replacement DNA nucleotides can be added to this 3'-OH group.

In linear chromosomes with multiple origins, the elongation of DNA in adjacent replicons also provides a 3'-OH group preceding each primer (Fig. 2.13b.). At the very end of a linear chromosome, however, there is no adjacent stretch of replicated DNA to provide this crucial 3'-OH group. Once the primer at the end of the chromosome has been removed, it cannot be replaced with DNA nucleotides, which produces a gap at the end of the chromosome (Fig. 2.13c), suggesting that the-chromosome should become progressively shorter with each round of replication, leading to the eventual elimination of the entire telomere and destabilization of the chromosome, and cell death. But chromosomes don't become shorter each generation and destabilize. Interestingly, the ends of chromosomes that preserve the integrity of the chromosome structure and avoid being getting shorter with each cycle of cell division.

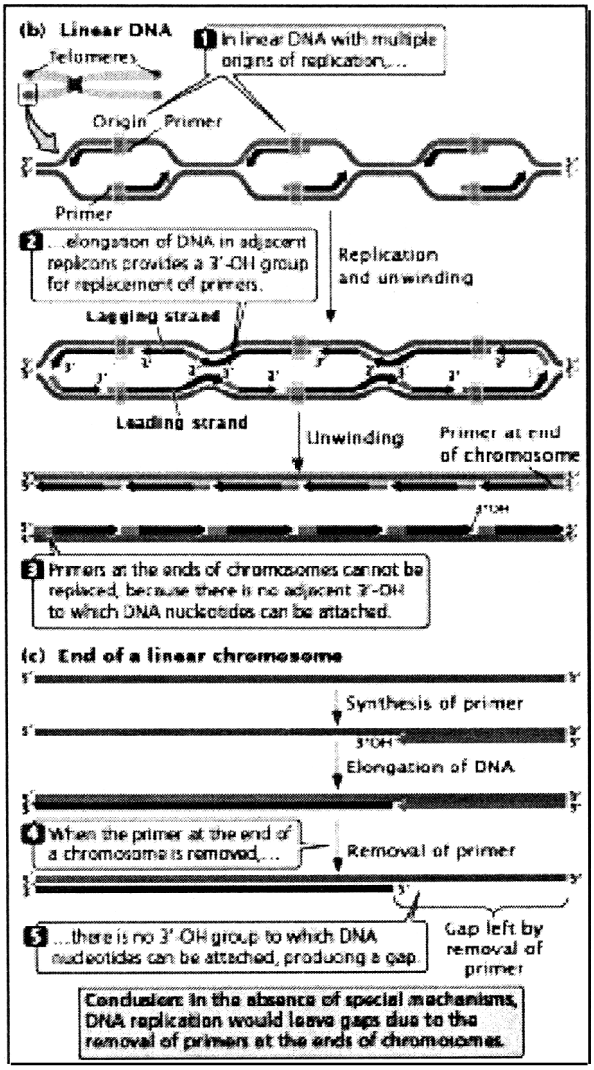


Fig. 2.13b : Replication at the ends of linear DNA

The telomeres possess several unique features, one of which is the presence of many copies of a short repeated sequence. In the protozoan *Tetrahymena,* this telomeric repeat is CCCCAA., with the G-rich strand typi-cally protruding beyond the C-rich strand (Fig. 2.14a): The single-stranded protruding end of the telomere can be extended by **telomerase,** an enzyme with both a protein and an RNA component (also known as a ribonucleoprotein). The RNA part of the enzyme contains from 15 to 22 nucleotides

that are complementary to the sequence on the G-rich strand. This sequence pairs with the overhanging 3' end of the DNA (Fig.2.14b) and provides a template for the synthesis of additional DNA copies of the repeats. DNA nucleotides are added to the the end of strand one at a time (Fig.2.14c) and, after several nucleotides have been added, the RNA template moves down the DNA and more nucleotides are added to the 3' end. Usually, from 14 to 16 nucleotides are added to the 3' end of the G-rich strand.

5' end of	3-CCCCAA	toward
Chromosome	3-GGGGTTGGGGTT	centromere

In this way, the telomerase can extend the 3' end of the chromosome without the use of a complementary DNA template. How the complementary Orich strand is synthesized is not yet clear. It may be synthesized by conventional replication, with priniase synthesizing an RNA primer on the 5' end of the extended (G rich) template. The removal of this primer once again leaves a gap at the 5' end of the chromosome, but this gap does not matter, because the end of the chromosome is extended at each replication by telomerase; no genetic information is lost, and the chromosome does not become shorter overall. The extended single-strand end may fold back on itself, forming a terminal loop by nonconventional pairing of bases, displacing a part of the original telomeric duplex. The loop structure is formed and stabilized by specific telomere-binding proteins (Fig. 2.15).

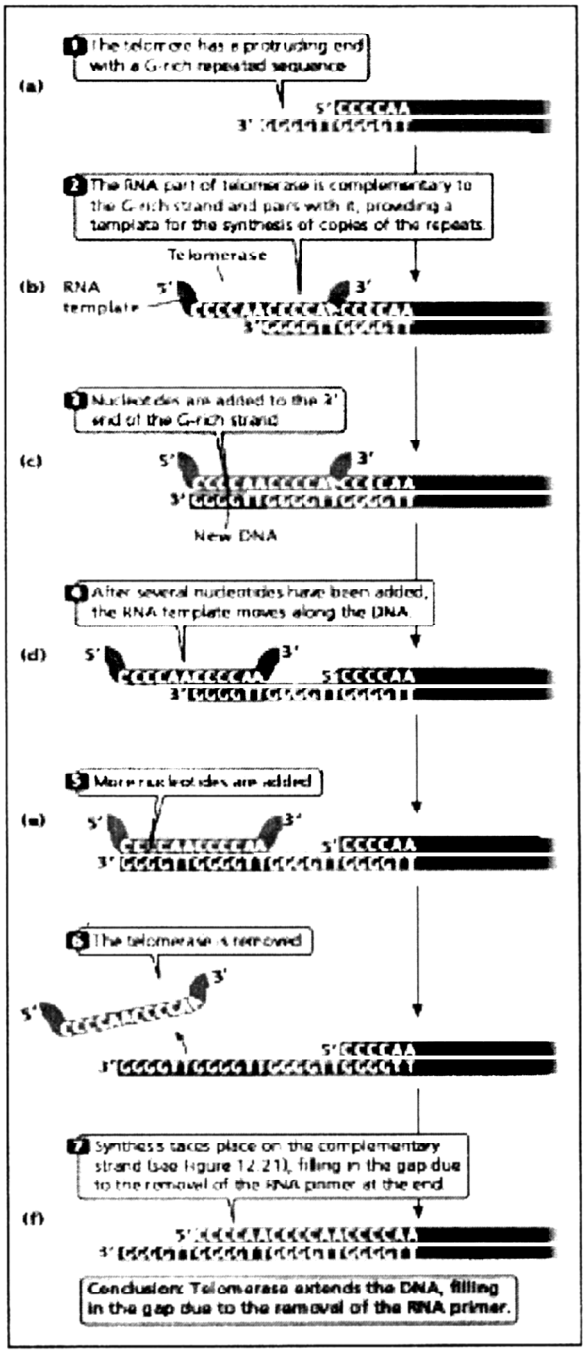


Fig. 2.14: DNA synthesis at the end point of linear chromosomes by telomerase enzyme

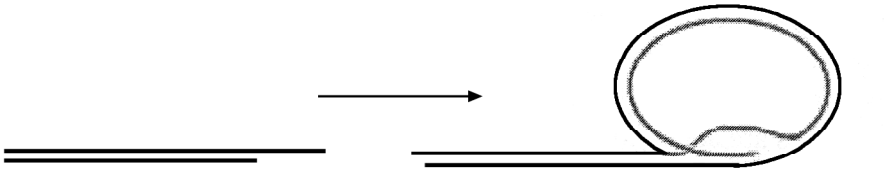


Fig. 2.15: Loop formation by the single stranded DNA strand at telomeric region by unconventional base pairing

2.4.9 Telomerase enzyme

Telomerase is a multi-subunit enzyme that is comprised of a RNA component - hTR, and a protein component - hTERT (Nakamura and Cech, 1998). hTR contains an 11 bp sequence that provides the template for the synthesis of telomeric repeats which are added to the chromosome, whereas hTERT, the reverse transcriptase component, catalyzes the synthesis reaction. Thus, addition of TTAGGG repeats to the 3' ends of chromosomes compensates for losses due to the end-replication problem. In humans, telomerase activity is absent in most normal cells but present in majority of tumors (Kim *et al,* 1994). However, activity has been detected in high levels in germ cells, early embryos (Xu and Yang, 2001), activated T and B cells and germinal centres of lymphoid organs.

Telomerase is present in single-celled organisms, germ cells, early embryonic cells, and certain proliferative somatic cells (such as bone-marrow cells and cells lining the intestine), all of which must undergo continuous cell division. Most somatic cells have little or no telomerase activity, and chromosomes in these cells progressively shorten with each cell division. These cells are capable of only a limited number of divisions; once the. telomeres shorten beyond a critical point, a chromosome becomes unstable, has a tendency to undergo rearrangements, and is degraded. These events lead to cell death. The shortening of telomeres may contribute to the process of aging. Genetically engineered mice that lack a functional telomerase gene do not express telomerase in somatic or germ cells and therefore experience progressive shortening of their telomeres in successive generations. After several generations, these mice show some signs of premature aging, such as graying, hair loss, and delayed wound healing. Through genetic engineering, it is also possible to create somatic cells that express telomerase. In these cells, telomeres do not shorten, cell aging is inhibited, and the cells will divide indefinitely. Telomerase also appears to play a role in cancer. Cancer tumor cells have the capacity to divide indefinitely, and many tumor cells express the telomerase enzyme. Telomerase activation alone does not lead to cancerous growth in most cells, but it does appear to be required along with other mutations for cancer to develop.

The length of the telomeric sequence varies from chromosome to chromosome and from cell to cell, suggesting that each telomere is a dynamic structure that actively grows and shrinks. The telomeres of *Drosophila* chromosomes are different in structure. They consist of multiple copies of the two different retrotransposons *Het-A* and *Tart,* arranged in tandem repeats. Apparently, in *Drosophila,* loss of telomere sequences during replication is balanced by transposition of additional copies of the *Het-A* and *Tart* elements. Farther away from the end of the chromosome, from several thousand to hundreds of thousands of base pairs form telomere-associated sequences. They, too, contain repeated sequences, but the repeats are longer, more varied, and more complex than those found in telomeric sequences.

2.4.10 Licensing: positive control of replication

The average human chromosome contains 150 x 10^6 nucleotide pairs which are copied at about 50 base pairs per second. The process would take a month (rather than the hour it actually does) but for the fact that there are many places on the eukaryotic chromosome where replication can begin. Replication begins at some replication origins earlier in S phase than at others, but the process is completed for all by the end of S phase. As replication nears completion, "bubbles" of newly replicated DNA meet and fuse, finally forming two new molecules.

In order to be replicated, each origin of replication must be bound by:

- An **O**rigin **R**ecognition **C**omplex of proteins (**ORC**). (These remain on the DNA throughout the process).
- Accessory proteins called **licensing factors**. (These accumulate in the nucleus during G_1 of the cell cycle. They include) :

  CDC-6 and CDT-1, which bind to the **ORC** and are essential for coating the DNA with

  MCM proteins. Only DNA coated with MCM proteins (there are 6 of them) can be replicated.
- Once replication begins in S phase,

  CDC-6 and CDT-1 leave the ORCs (the latter by ubiquination and destruction in proteasomes).

  The MCM proteins leave in front of the advancing replication fork.

2.4.11 Geminin: negative control of replication

G_2 nuclei also contain at least one protein — called **geminin** — that prevents assembly of MCM proteins on freshly-synthesized DNA (probably by sequestering Cdtl).

As the cell **completes mitosis,** geminin is degraded so the DNA of the two daughter cells will be able to respond to licensing factors and be able to replicate their DNA at the next S phase.

Some cells deliberately cut the cell cycle short allowing repeated S phases without completing mitosis and/or cytokinesis. This is called **endoreplication.** How these cells regulate the factors that normally prevent DNA replication if mitosis has not occurred is still being studied. Endoreplication is described on a separate page.

2.4.12 Post-replicative modification of DNA, methylation

One of the major post-replicative reactions that modifies the DNA is **methylation.** The sites of natural methylation (i.e. not chemically induced) of eukaryotic DNA is always on cytosine residues that are present in CpG dinucleotides. However, it should be noted that not all CpG dinucleotides are methylated at the C residue. The cytidine is methylated at the 5 position of the pyrimidine ring generating 5-methylcytidine.

Methylation of DNA in prokaryotic cells also occurs. The function of this methylation is to prevent degradation of host DNA in the presence of enzymatic activities synthesized by bacteria called restriction endonucleases. These enzymes recognize specific nucleotide sequences of DNA. The role of this system in prokaryotic cells (called the restriction-modification system) is to degrade invading viral DNAs. Since the viral DNAs are not modified by methylation they are degraded by the host restriction enzymes. The methylated host genome is resistant to the action of these enzymes.

The precise role of methylation in eukaryotic DNA is unclear. It was originally thought that methylated DNA would be less transcriptionally active. Indeed, experiments have been carried out to demonstrate that this is true for certain genes. For example, under-methylation of the MyoD gene (a master control gene regulating the differentiation of muscle cells through the control of the expression of muscle-specific genes) results in the conversion of fibroblasts to myoblasts. The experiments were carried out by allowing replicating fibroblasts to incorporate 5-azacytidine into their newly synthesized DNA. This analog of cytidine prevents methylation. The net result is that the maternal pattern of methylation is lost and numerous genes become under methylated. However, lack of methylation nor the presence of methylation is a clear indicator of whether a gene will be transcriptionally active or silent.

The pattern of methylation is copied post-replicatively by the **maintenance methy-lase system**. This activity recognizes the pattern of methylated C residues

in the maternal DNA strand following replication and methylates the C residue present in the corresponding CpG dinucleotide of the daughter strand.

The phenomenon of **genomic imprinting** refers to the fact that the expression of some genes depends on whether or not they are inherited maternally or paternally. **Insulin-like growth factor-2 (IGF2)** is a gene whose expression is required for normal fetal development and growth. Expression of IGF2 occurs exclusively from the paternal copy of the gene. Imprinted genes are "marked" by their state of methylation. In the case of IGF2 an element in the maternal locus, called an **insulator element,** is methylated blocking its function. The function of the un-methylated insulator is to bind a protein that when bound blocks activation of IGF2 expression. When methylated the protein cannot bind the insulator thus allowing a distant enhancer element to drive expression of the IGF2 gene. In the maternal genome, the insulator is not methylated, thus protein binds to it blocking the action of the distant enhancer element.

2.5 Mechanism of replication

2.5.1 Replication in prokaryotes

Replication in prokaryotes and viral DNA usually starts at specific sites on chromosome, referred to as **replication origin.** The origin in E. coli is specifically known as ***OriC***. Approximately 20-30 different proteins, some in multiple copies are required to initiate the DNA replication process. Some of the proteins characterized and their genes are given in Table 1.

Initiation and Unwinding of DNA: function of Helicases & Topoisomerase

In *E.coli,* ***OriC*** comprises of **245** base pairs that contains **four 9 bp** sites with similar sequences at which product of ***dnaA*** (homologous tetramer) binds and initiates the assembly of all other proteins and enzymes necessary for replication. In addition, the origin contains 11 methylation sites

Gene Product or/functions	Gene
Initiator protein; binds at OriC	*dnaA*
1HF protein-DNA binding protein; binds at OriC	*HimA*
F1S protein-DNA binding protein; binds at OriC	***fis***
Heiicase and activator of primase	*dnaB*
Proteins that complexes with dnaB protein and delivers to DNA	*dnaC*
Primase-synthesizes RNA primer	*dnaG*
Single stranded binding proteins (SSB-proteins)	*ssb*
DNA ligase	*Lig*
Gyrase (topoisomerase type. II)	*gyrA,, yrB*
TBP proteins (terminus (ter) binding proteins) stops replication	*Tus*
Topoisomerase I	*topA*
Topoisomerase IV	*parE*

Table 1: Different proteins that assimilate at the initiation complexes

recognized by **DNA methylase** and **three AT rich** direct tandem repeats consisting of 13 base pair each.

Initiation of replication begins with the binding of **dnaA** molecules at 4 sites consisting of 9 nucleotides provided that the 9-mers are fully methylated.

- Region of the DNA bound to dnaA coalesce and are joined by additional dnaA molecules to form a nucleosome like complex, which promotes nearby melting of the double helix at **AT rich** site (Fig. 2.16).

- Physical separation of the two strands requires untwisting of the DNA molecule. Unwinding of the DNA is facilitated by helicases, which is the product of the ***dnaB.*** dnaA, with the aid of dnaC binds to the helicases, which then makes contact with the DNA at the replication fork and separates the two strands to form a bubble by melting the hydrogen bonds. Energy required for unwinding is derived from the hydrolysis of ATP.

- During unwinding tension build up ahead of the replication fork because of the formation of super coils. DNA gyrase (topoisomerase

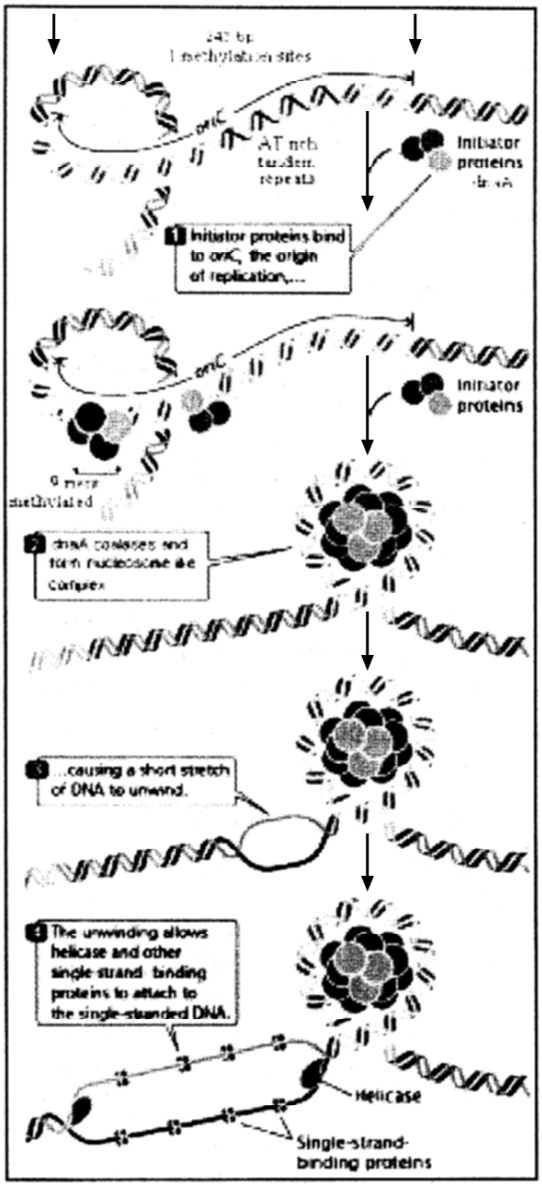


Fig. 2.16: Formation of initiation complex at the site of origin of DNA synthesis

I & **III** nicks single strand while **II** & IV nicks two strands to relieve tensions), reduce trosional strain (torque) that builds up ahead of the replication fork asa result of unwinding The topoiso-merase apparently nick one strand of the double stranded DNA ahead of the replication fork. The nicked DNA molecule then rotates, relieving the tension and avoid the formation of super coils. There are some indication that topoisomerase also induce negative super coils ahead of the replication fork thereby reliving the tensions and also assist the helicases in the process of unwinding **(Fig. 2.17).**

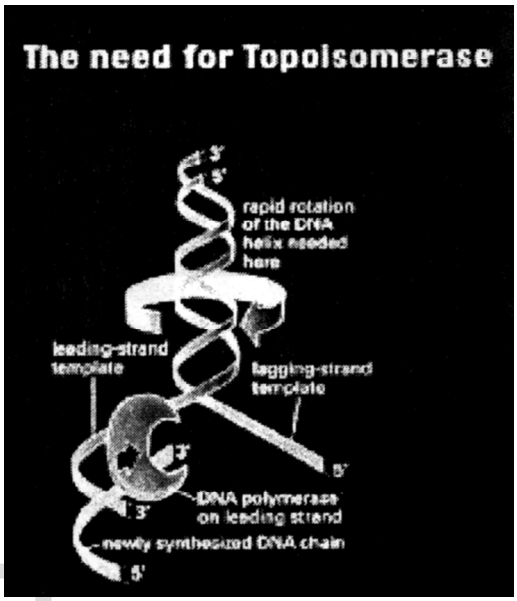


Fig. 2.17: Unwinding of the DNA helix ahead of the replication fork by topoisomerase releases tension and avoids DNA breaks.

2.5.2 Formation of Replisome:

The assembly of all protein factors and enzymes at the site of origin of replication is called a **replisome** (Fig. 2.18).

- Immediately after the bubble is formed **Single-strand binding proteins (SSBs)** binds to the single DNA strands and stabilizes them to avoid any unwanted reactions and- also to prevent the single strand DNA to reunite and form a duplex again, until replication at that region is complete. Formation of the bubble creates a '**Y**' shaped structure called a replication fork. A replication fork moves in one direction.
- Assembly of dnaB-dnaC complex is followed by the addition of four other poly-peptides - n, n', n" and i. This complex constitutes the prepriming complex.

The stage is now ready for the binding of the **Primase** the product of ***dnaG,*** Primase synthesize short RNA sequences complementary to DNA strands at the initiation site. This is because DNA polymerase **III** cannot initiate synthesis of a chain of DNA from free nucleotides and require a RNA primer to provide a free 3'-"OH end that can be extended by addition of nucleotides. Addition of Primase converts the priming complex to a **primosome.** The primase is much smaller that the usual RNA polymerase and is only 6000dalton. Primase is activated by dnaB, which then starts synthesizing short RNA primers on both the strands. The primers start with two purine nucleotides, most frequently pppAG The primers are usually

10 to 15 bases long. Assembly of the **replisome** is completed by the addition of **DNA polymerase III** (Fig. 2.18).

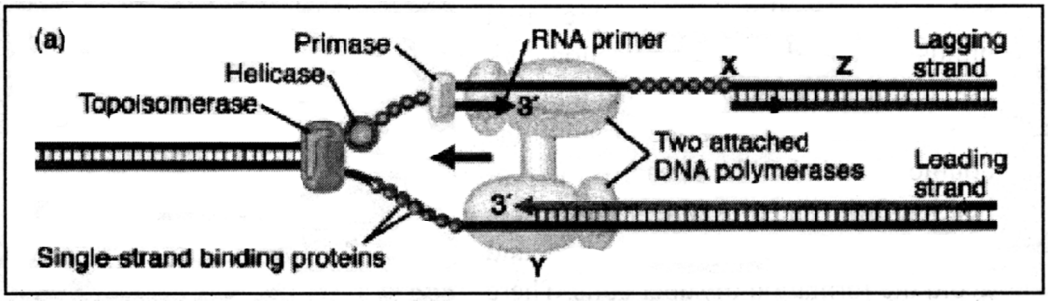


Fig. 2.18 : Replisome complex at the site of origin of DNA synthesis

2.5.3 Elongation

Once the single strands of DNA are stabilized, they serve as templates upon which new strands are synthesized by the enzyme referred to as DNA polymerase. DNA polymerases elongate the polynucleotide strand by catalyzing polymerization reaction. DNA polymerases add nucleotides to the 3-OH group on a nucleotide already paired with the template strand and consequently DNA synthesis can only proceed in the 5'–3' direction. Unlike the RNA polymerases, DNA polymerases require short primer sequences to initiate DNA synthesis on a single strand DNA template. Finally, the addition of new nucleotide is not random. DNA polymerase selects each deoxyribonucleotide that can form a complementary base pair with the nucleotide on the template strand DNA.

2.5.4 Termination of DNA synthesis

Termination occurs at *ter* or t locus, lying across from ***Ori C*** of the circular chromosome, between minutes 28 to 36. This region incorporates 6 **ter** sequences with sequence GTGTGTTGT that bind Tus protein **(terminator protein).** Three ter sequences arrest replication from the right and rest three sequences arrest replication coming from the left. One interesting aspect of E. coli DNA replication is that the cells are viable even if the whole terminator region is deleted and can terminate replication.

- Tus protein is a contra helicases, and functions by interfering with the ATP dependent function of dnaB helicases (rather than simply impending the propagation of this helicases along the double helix)
- Each Ter-Tus site has directional properties and it arrests only those replisomes that reach the Tus site from one specific direction.
- Replisome arriving from opposite direction apparently force dissociation of the TUS protein and thus can proceed unimpeded past the Ter-Tus site.

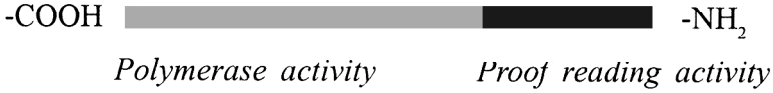
2.6 DNA Polymerases

The best-studied polymerases are those of *E. coli,* which has at least five different DNA polymerases. DNA polymerase I and DNA polymerase III play the major role during DNA replication, the pther three have specialized functions in DNA repair mechanism. All the DNA polymerases share the same fundamental property of adding new nucleotides only to the 3'-OH group on a previously existing paired nucleotide on a DNA template. Arthur Kornberg (1956) first identified and isolated DNA polymerase from the lysate of is. *coli,* which is now known as **DNA polymerase** I. Later, identification of DNA polymerase I deficient *E. coli* clones led to the isolation of two new polymerases - **DNA polymerase II** and **DNA polymerase III.** Characteristic properties of different polymerases isolated from *E. coli* are given in **Table 2.**

DNA Polymerase	5′→3′ Polymerization	3′→5′ Exonuclease	5′→3′ Exonuclease	Function
1	Yes	Yes	Yes	Removes and replaces primer
II	Yes	Yes	No	DNA repair: restarts replication after damaged DNA halts synthesis
III	Yes	Yes	No	Elongates DNA
IV	Yes	No	No	DNA repair
V	Yes	No	No	DNA repair: traitslesion DNA synthesis

Table 2: Properties of DNA polymerases found in *E. coli*

DNA polymerase I : DNA polymerase I participate in lagging strand synthesis by eliminating primer RNAs and also have a role in repair mechanism. This enzyme is coded by locus polA and is a single polypeptide chain. When treated with proteolytic enzymes, it is cleaved into two fragments- the larger fragment is know as **Klenow** fragment (used for *in vitro* synthesis). Two third of the protein chain, beginning from the C terminal end has polymerase activity while the rest one third on N-terminal end contains proofreading exonuclease activity, (no. of units per cell = 400)



DNA polymerase II: The biological function of DNA polymerase II is unclear, although this enzyme is induced when chromosomal DNA is damaged.

DNA polymerase III: DNA polymerase III is a huge multiprotein holoenzyme complex that plays the major role during DNA replication in *E. coli.* No. of units present per cell is approximately 10 to 12. It consists of 10 different polypeptide

chains. The catalytic core is composed of three subunits. The α subunit possess 5′→3′ DNA synthetic activity, ε subunit has the 3′ →5′ exonuclease activity. The θ subunit possibly participate in assembly of the enzyme. The remaining seven auxiliary subunits enhance the processivity of the core by clamping it onto the template. The addition of the τ causes the core to dimerize. The functions of γ and δ subunits are less well defined. The formation of holoenzyme is completed by the addition of the β ? subunit (Table 3). Structural analysis has revealed that DNA polymerase III is an asymmetric dimmer with twin polymerase activity sites, capable of synthesizing DNA strands simultaneously on both leading and lagging strand. The 5′ →3′ DNA synthetic activity and the 3′ →5′ exonuclease activity together allow DNA polymerase III to efficiently and accurately synthesize new DNA molecules.

DNA polymerase	***No. of Units***	***Gene***	***mol. wt.***	***Polymerization 5′→3′***	***Exonuclease activity***	***molecules. per cell***
I	One	*polA*	103.0 kd	Yes	3′ 5′ & 5′	3′ 400
II	One	*polB*	90.0 kd	Yes	3′ 5′	?
III	Ten	——		Yes	3′ 5′	10-12
α = dna*E*,130.0kd; ε = *dnaQ*, 27.5 kd; θ = *holE*, 10.0 kd; *T* (tau) = *dnaX*, 71.0 kd; γ = *dnaX*,47.0 kd						
β = *dnaN*,40.6 kd δ = *holA* 35.0 kd; δ = *holA*, 33.0kd; *x(chi)* = *holB*,15.0 kd: ψ(psi) =*hold;* 12.0 kd						

Table 3: Different subunits of the DNA polymerase III, their molecular weights and cellular function

All the *E. coli* DNA polymerases have 3′ →5′ exonuclease activity. This means that the DNA polymerases check the accuracy of the most recently assembled base pair. If a wrong base pair is included, exonuclease activity removes the erroneous

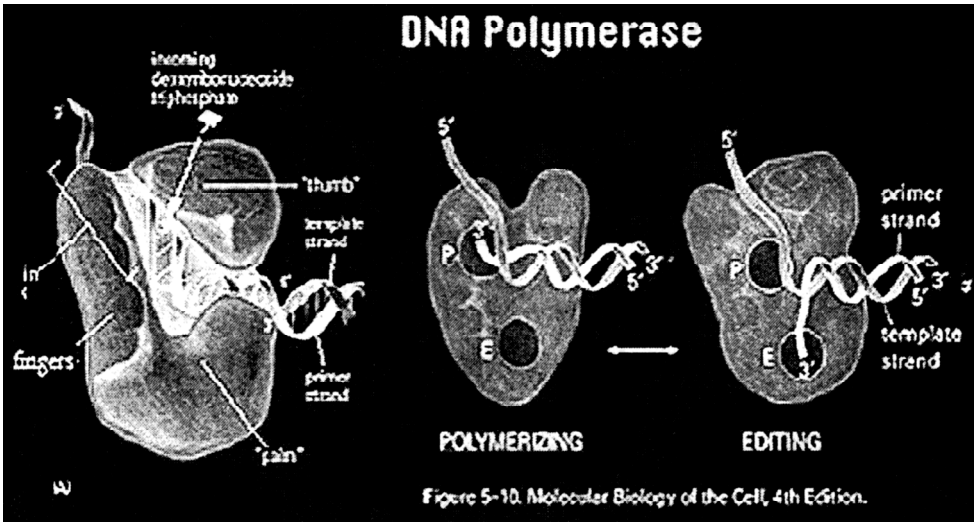


Fig. 2.19

nucleotide by excision and catalyzes the formation of the correct base pair. Thus in DNA replication 3′→5′ exonuclease activity is a proofreading mechanism that helps to keep the frequency of DNA replication errors at very low level (10^9/cycle). In addition, DNA polymerase I has 5′→3′ exonuclease activity and can remove nucfeotides from the DNA 5' end of a DNA strand or of an RNA primer strand. This activity is important for DNA repair.

2.6.1 DNA polymerase activity

- DNA synthesis begins immediately after the addition of DNA polymerase III by complementary base pairing at the 3'-OH of the primer.
- Because of the anti parallel nature of the DNA helix and the unidirectional movement of the DNA polymerase along the replication fork poses a problem, since DNA polymerase can only make new DNA strands in the 5' to 3' direction. Interestingly, both the strands of the DNA double helix are synthesized simultaneously.
- The 3′ →5′ template in the direction of fork movement is synthesized in a continuous manner and is called the **leading strand** (Fig. 2.20).
- The 5′ →3′ template in the direction of fork movement is synthesized in a discontinuous manner and is called the **lagging strand** (Fig. 2.20).
- DNA polymerase III complexes are endowed with the property to synthesize continuously on the leading strand and synthesize discontinuously on lagging strand.

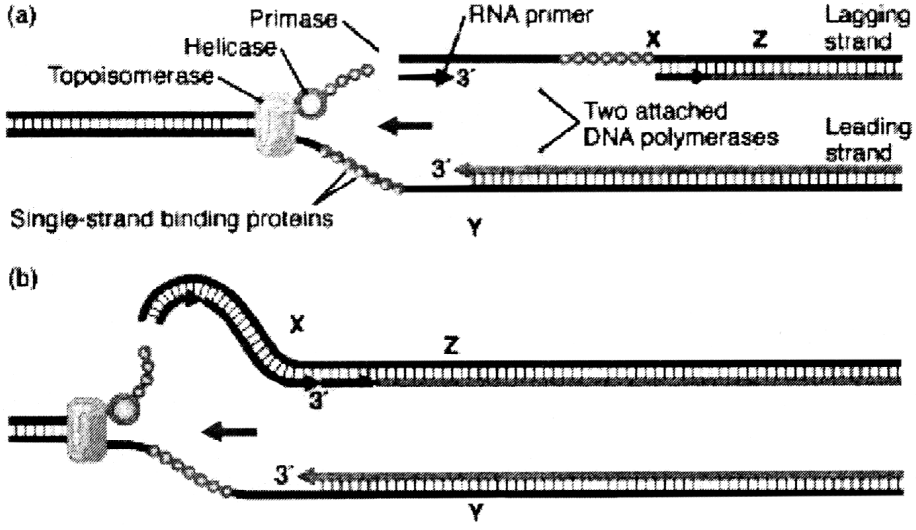


Fig. 2.20: (a) Different enzymes and factors at the site of replication fork,
(b) The replication on leading and lagging strand

- Lagging strand is synthesized discontinuously, in short stretches- known as **Okazaki fragments.** Formation of Okazaki fragments is also initiated by **Primase** at sites selected by pre-priming proteins. Each Okazaki fragments starts with a primer - a sequence of RNA, approximately 10 bases long that provides 3'-OH end for extension by DNA polymerase III.

- When a nascent Okazaki fragment reaches the 5 end of the previously synthesized Okazaki fragment, the lagging strand template is released and un-looped.
  The RNA primers of the Okazaki fragments in *E. coli* are removed by the combined activity of **RNase H** and **DNA polymerase I** that fills the gap. DNA polymerase I follows DNA polymerase III and, using its 5′ →3′ exo-nuclease activity removes the RNA primer. It then uses its 5′ →3′ polymerase activity to replace the RNA nucleotides with DNA nucleotides one at a time.

- Two Okazaki fragments are joined by DNA **ligase** producing intact DNA daughter strand.

Unit 3 Prokaryotic Transcription

Structure

3.1 Introduction

Genomic DNA contains a set of information that governs all cellular activities. These instructions are implemented by synthesis of RNA and proteins. Genetic character of a cell is determined by not only what genes are inherited, but also on how and when the genes are expressed.

Three years after Watson and Crick (1953) published the DNA model, Crick proposed the Central Dogma, describing the two-step process of protein synthesis. According the theme, flow of information is always unidirectional i.e. DNA to RNA and RNA to Protein. However after the discovery of reverse transcriptase in retro viruses by Temin and Baltimore in 1970, the dogma now stands as follows (Fig. 3.1)

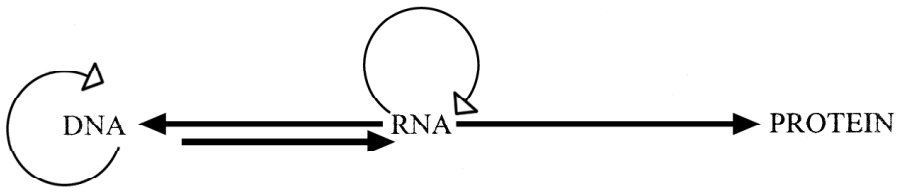


Fig. 3.1: Flow of genetic information from DNA to Protein

information never flows back from protein to RNA, in other words it can be said that acquired characters are never inherited.

The first level of gene expression involves the transfer of information stored in DNA to single stranded RNA molecule by way of a process called **transcription.** In the second step, the information scripted in the RNA molecule is translated into a linear sequence of amino acids and the process is called **translation.**

3.2 Similarities and difference between replication and transcription

Our understanding of the transcriptional process comes from the study of *E. coli.* There are several similarities between the transcription mechanism and DNA replication. Both the synthesis process utilize the similar nucleotide building blocks and : use the same chemical method of attack by a terminal -OH group of the growing chain on the triphosphate group of an incoming nucleotide. Both replication and transcription are fueled by the hydrolysis of the pyrophosphate group that is released upon attack. There are however, a number of important differences between these two distinct processes.

a) One major difference rests on the fact that while DNA replication copies an entire helix, DNA transcription only transcribes specific regions of one strand of the helix. During DNA transcription, only short stretches (about 60 base pairs) of the template DNA helix are unwound. As the RNA polymerase transcribes more of the DNA strand, this short stretch moves along with the transcription machinery. This process is different from that in DNA replication in which the parent helix remains separated until replication is done.

b) There are slight differences in the substrates that are used in DNA replication versus transcription. Transcription machinery utilize ribonucleotide instead of deoxyribonucleotide triphosphates. Additionally, in RNA the thym-ine base is replaced with the base uracil. Both of these differences can be seen in DNA transcription.

c) Another major difference is that DNA replication is a highly regulated proc- I ess that only occurs at specific times during a cell's life. DNA transcription is also regulated, but it is triggered by different signals from those used to control DNA replication.

d) One final difference lies in the capabilities of RNA polymerase versus DNA polymerase. RNA primers are needed to begin replication because DNA polymerase is unable to do it alone. DNA transcription does not require any primer synthesis as the RNA polymerase is capable of initiating RNA syn- I thesis. The structure of the RNA polymerase is necessary for understanding all of the processes that underlie initiation, elongation, and termination and I also explain some of its added capabilities.

3.3 General idea of transcription

Some 50 different protein **transcription factors** bind to **promoter** sites, usually on the 5' side of the gene to be transcribed

Transcription in both prokaryotes and eukaryotes is catalyzed by **RNA polymerase,** that synthesizes a complementary RNA molecule using one strand of the duplex DNA as template

The DNA strand that acts as template is called **template strand** and the DNA strand that is identical in sequence to the RNA strand is called **sense strand** or coding strand or nontemplate strand

5' ...ATGGCCTGGACTTCA... 3' Sense strand of DNA

3' ...TACCGGACCTGAAGT... 5' Antisense strand of DNA

↓ **Transcription of antisense strand**

5' ...HUGGCCUGGRCUUCH... 3' mRNA

↓ **Translation of mRNA**

Met— Ala— Trp— Thr — Ser — Peptide

Transcription of RNA proceeds in the 5′→3′ direction

Several types of RNA molecules are transcribed by RNA polymerases:

messenger RNA (mRNA). This will later be **translated** into apolypeptide.

ribosomal RNA (rRNA). This will be used in the building of ribosomes: machinery for synthesizing proteins by translating mRNA.

transfer RNA (tRNA). RNA molecules that carry amino acids to the growing polypeptide.

small nuclear RNA (snRNA). DNA transcription of the genes for mRNA, rRNA, and tRNA produces large precursor molecules **("primary transcripts")** that must be processed within the nucleus to produce the functional molecules for export to the cytosol. Some of these processing steps are mediated by snRNAs.

small nucleolar RNA (snoRNA). These RNAs within the nucleolus have several functions.

micro RNA (miRNA). These are tiny (-22 nts) RNA molecules that appear to regulate the expression of messenger RNA (mRNA) molecules.

XIST RNA. This inactivates one of the two X chromosomes in female vertebrates.

RNA primers formed during DNA synthesis

telomerase RNA,

ribozymes that act as enzymes

In prokaryotes, there is only a single type of RNA polymerase responsible for synthesizing all types of RNAs.

E. coli RNA polymerase is a holoenzyme comprised of subunits ââ′á(dimer) and ó^{70}.

The ó^{70} subunit is the subunit which binds to the promoter region, but is unable to initiate RNA synthesis.

After the ó^{70} subunit subunit binds, the other subunits bind forming a function RNA polymerase.

After approximately 10 base pairs have been transcribed the ó^{70} subunit leaves and the **core polymerase** continues on.

In eukaryotes, there are three major classes of RNA polymerase:

1. RNA polymerase I transcribes large rRNAs
2. RNA polymerase II transcribes mRNAs
3. RNA polymerase III tRNAs, small rRNAs and other small RNAs

In prokaryotes, the transcription and translation process are **coupled**

In eukaryotes, transcription and translation events are **compartmentalized.** RNA molecules are transcribed in the nucleus. All types of RNAs is then exported into the cytoplasm for translation

Messenger RNAs are processed in eukaryotes prior to translation. Both ends of the RNA are modified in the nucleus. The transcribed intron sequences are removed by splicing to produce a final mRNA ready for translation.

rRNAs and tRNAs are processed in both prokaryotes and eukaryotes. Most rRNAs are synthesized as a single large precursor RNA that is then cleaved into it final products. In tRNA, many individual nucleotides are chemically modified to produce the final tRNA

Transcription proceeds through the steps of **initiation, elongation** and **termination** in all cell types.

3.4 Transcription in prokaryotes

3.4.1 RNA polymerase

The structure of the RNA polymerase is necessary for understanding all of the processes that underlie initiation, elongation, and termination and also explain some of its added capabilities. In prokaryotes a single RNA polymerase transcribes all genes. There are two main segments of the RNA polymerase molecule: the core enzyme, and the sigma subunit. These two pieces are together referred to as the "holoenzyme". The *E. coli* RNA **core enzyme** is tetrameric, containing two â ? one **â'** ?and one â' type subunits. The core enzyme is sufficient for transcriptional elongation, but correct initiation requires the sigma subunit called ó^{70} factor. The ó^{70} subunit binds to the promoter region, but is unable to initiate RNA synthesis. After the ó^{70} subunit binds, the other subunits bind forming a functional RNA polymerase. Specific function of each polypeptide is given in the table 1.

Sub units	***MW***	***Location***	***Possible function***
α *chains*	*40 kd*	*core enzyme*	*Promoter binding*
β *chain*	*155 kd*	*core enzyme*	*Nucleotide binding substrate*
β′ *chain*	*160 kd*	*core enzyme*	*Template binding*
σ *factor*	*32-90 kd*	*Sigma factor*	*Initiation of transcription*

Table l: Different components of RNA polymerase that comprises the holoenzyme

The start site represents the location on the DNA that marks where the first nucleotide of an RNA chain will commence which is also designated as the "plus one position". All positions designated as upstream of the start site are labeled with negative numbers according to their position relative to the start site. Sequences located just upstream of the start site is called the **promoter** region. This region contains the information that signals the RNA polymerase to start transcription. In prokaryotic cells, free RNA polymerase molecules are constantly colliding with DNA helices. The collision leads to a weak association between the DNA and RNA polymerase, which is soon broken. However, when the RNA poJymerase binds to a specific sequence on the DNA, it binds tightly, forming a DNA/RNA polymerase complex. The **ó**? factor has two functions, it recognizes the promoter and it converts the closed promoter complex into an open promoter complex. After approximately 10 base pairs have been transcribed the ó^{70} subunit leaves and the **core polymerase** continues on. The core enzyme can bind to DNA in the absence of ó factor but with low efficiency and no specificity. The primary function of the ó factor is thus to increase the binding efficiency of RNA polymerase at the promoter and decrease (non-specific, binding. A single ó factor (ó70 in *E. coli)* initiates the transcription of

most genes, but other ó factors are present that functions only with specialized genes. For example (ó^{32}, ó^{415}, ó^{54}, ó^{28}) function under adverse conditions like high temperature, starvation, nitrogen deficiency and for chemotaxis. Some bacteriophages (e.g. T4) encode their own ó factor, which subvert the host enzyme into transcribing the phage genes, T3 and T7 phages encode their own RNA polymerases, which are single polypeptides with great affinity for the phage promoters.

3.4.2 Initiation

Initiation of transcription begins with the binding of the RNA polymerase to DNA strand. RNA polymerase binds to the DNA non-specifically and with low affinity. In the presence of the 6 factor, the holoenzyme associate with the DNA at specific sites called promoters.

Promoters: The DNA sequence to which RNA polymerase binds to initiate tran- I scription of gene is called the promoter. The DNA sequence at the promoter region is more or less conserved. Mutation or deletions at promoter region severely affects the transcriptional efficiency of that gene. The promoter is cis-acting and is always located upstream from the point of initiation of transcription. In *E. coli,* there are two promoter motifs situated at -10 and -35 sequence upstream from the start point of RNA synthesis. The consensus sequence at -10 position is **TATA AT** (Pribnow box) and at **-35** is **TTGACA** (Fig. 3.2).

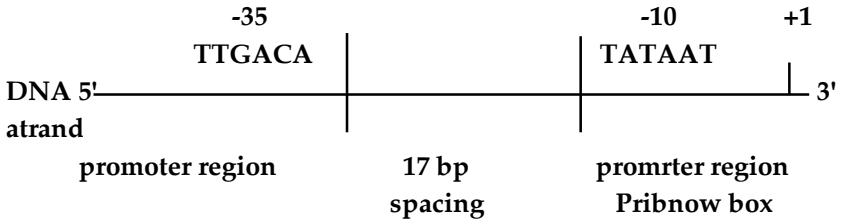


Fig. 3.2: Organization of the promoter region in *E. coli*

The -35 sequence (also called recognition sequence) is essential for binding of the RNA polymerase. At the -10 promoter region, the DNA strands' unwinds when associated with the RNA polymerase, preparing for initiation of transcription. The initial binding between the polymerase and promoter is referred to as a **closed promoter complex** because DNA is wounded. The unwinding of approximately 15 bases of DNA by the RNA polymerase around initiation site is called **open promoter complex.** Ideally, the gap between the two promoter regions is 17 base pairs long. Deviations from this spacing have significant effects on the strength of the promoter region. The closer a promoter region is to matching this canonical promoter sequence, the greater its strength. There is a third promoter element that

is sometimes seen in very strong promoters which is called the **UP element.** It usually is composed of alternating stretches of 5 adenine and thymine bases. It is located upstream of the -35 region.

Transcription usually begins with GTP or ATP and unlike subsequent nucleotides; the first nucleotide retains its triphosphate moiety. A cycle of abortive occurs generating 2-9 base short RNA sequence before actual transcription begins. Once initiation succeeds, 6 factor dissociates from RNA holoenzyme after 9-10 RNA nucleotide polymerizes and then core enzyme completes further elongation of the RNA.

3.4.3 Elongation

Elongation proceeds in the 5′ to 3′ end direction. As RNA polymerase travels down stream, it unwinds the double stranded DNA molecule ahead of it and rewinds the DNA molecule behind it, maintaining an unwounded region of about 17 bp in the region of transcription. Transcription proceeds at the rate of about 30 to 50 nucleotides per second. Energy required for polymerization of nucleotides is obtained from the cleavage of the triphosphate nucleotides (Fig. 3.3)

The RNA-DNA hybrid is very transient. The nascent RNA molecule rapidly separates from DNA and at any given time 2-12 nucleotides associate with each other. Certain proteins influence the rate of the synthesis. For example **NusA** - a protein, binds to the core enzyme that slows down elongation so that ribosome molecules are able to bind.to the nascent RNA molecule right behind the point of synthesis.

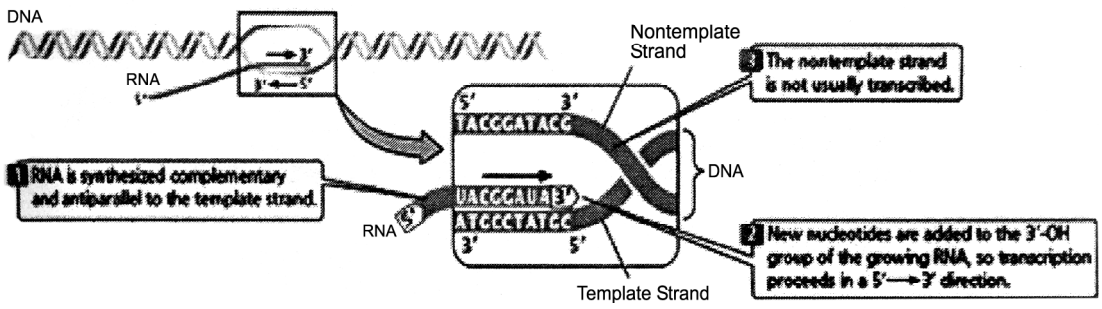


Fig. 3.3: The transcription bubble on the DNA template

3.4.4 Termination

TTermination of RNA synthesis occurs when RNA polymerase reach the end of the gene. Two mechanisms operate in bacteria and viruses: 1. Intrinsic termination or called ρ independent termination 2. ρ dependent termination.

Intrinsic termination (ρ independent termination): In this type, nucleotide sequence near the end of the transcribed RNA specifies the termination of RNA transcription. The sequences at the end of RNA are self-complementary that form a hairpin structure followed by a conserved sequence with the consensus sequence UUUUUUA. This region is called the **intrinsic terminator** (Fig. 3.4). For example:

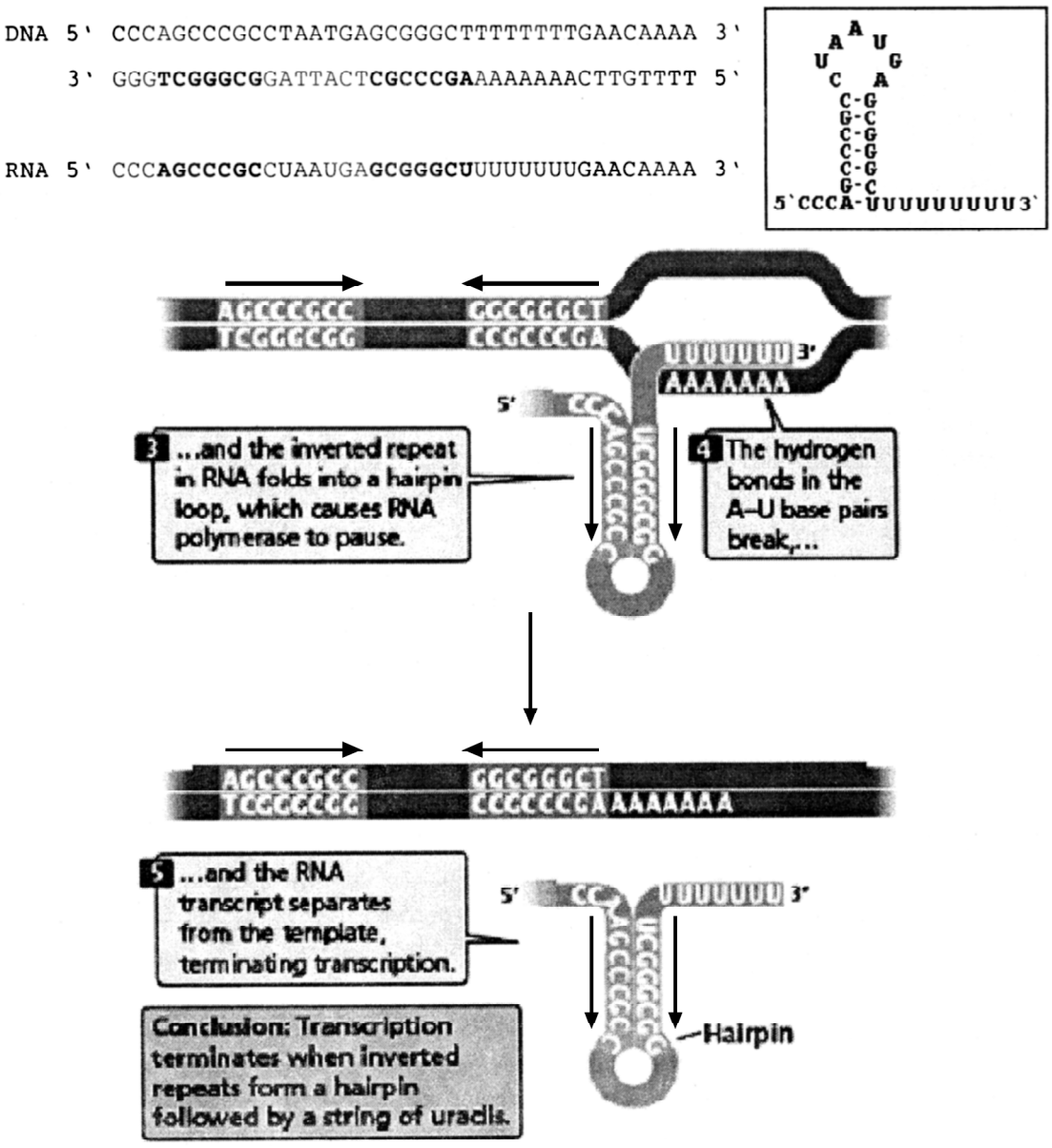


Fig. 3.4. Formation of hairpin loop on the transcribing RNA and ρ independent termination

RNA molecules having self-complementary bases at the termination point form a secondary structure called **hairpin loop** immediately after it is synthesized. The hairpin loop is followed by poly U sequence. Formation of hairpin loop and presence of poly U sequence probably halts the RNA polymerase and cause termination of RNA synthesis..The exact mechanism of action is not clear, probably the formation of secondary structure and the poly U sequence immediately before the termination signal on DNA destabilizes the RNA polymerase and releases the polymerase of the template strand. The weak bonding between the poly A DNA sequence and poly U strand on RNA may also contribute to the termination of RNA synthesis. However, hairpin loop alone not sufficient to terminate polymerase activity.

ρ dependent termination: In ρ dependent termination, the DNA sequences produce a pause in transcription towards the end of transcription and RNA does not produce any secondary structures and also lack poly U sequences. In this type, the p protein plays the principal role in termination. The ρ protein has two domains. One domain bind to RNA and the other domain bind to ATP molecule. Hydrolysis of the ATP molecule activates the ρ factor that then bind to specific site on RNA molecules in the termination region. When RNA polymerase encounters the terminator, it pauses, allowing rho to catch up. The rho protein has helicase activity, which it uses to unwind the RNA-DNA hybrid in the transcription bubble, bringing an end to transcription. Hereto, the exact mechanism of termination needs to be worked out, but most likely it destabilizes the association of the RNA-DNA-polymer association.

3.5 Transcription in eukaryotes

In eukaryotes, the genetic material remains enclosed within the nuclear membrane and is physically separated from the other organelles of the cell. The transcriptional process occurs within the nucleus. Although the transcription proceeds by the same fundamental mechanism as in prokaryotes, the regulatory mechanism is far more complex in eukaryotic cells. There is a significant difference between the transcription of eukaryotic and prokaryotic mRNAs in the initiation process. Eukaryotic promoter involves a large number of factors that bind to a variety of cw-acting elements. Eukaryotic cells possesses three different types of RNA polymerase, each specialized to transcribe different types of RNA. The promoter region on the DNA strand is defined by the nature of the RNA polymerase and transcription factors that will bind to specific sequences and support transcription at the normal efficiency and with the proper control. In fact, RNA polymerase does not make interaction with the upstream region of the promoter. The increased complexity of eukaryotic transcription presumably facilitates the

sophisticated regulation of gene expression needed to direct the activities of the many different cell- types of multicellular organisms.

3.5.1 Eukaryotic RNA polymerases and their promoters

In eukaryotic cells, there are three different types of RNA poiymerases, each located in different locations in the nucleus and is responsible for synthesizing different classes of genes as shown in the table below:

ENZYME	FUNCTION	SENSITIVITY
RNA polymerase **I** (nuckoli)	Transcription of the 45S rRNA precursor (later cleaved into 5.8S, 18S, 28S rRNA (class I genes) RNA processing	Insensitive to a-amanitin, sensitive to actinomycin D
RNA polymerase **II** (nucleoplasm)	Transcription of all protein encoding genes and most genes for small nuclear RNAs (class II genes)	Inhibited by α-amanitin
RNA polymerase **III** (nucleoplasm)	Transcription of tRNA genes, 5S rRNA genes and genes encoding U6 sn RNA and the various sn RNAs (class III genes)	Moderately sensitive to α-amanitin depending on species

Eukaryotic polymerases differ in template specificity, location in the nucleus and susceptibility to different inhibitors. Each RNA polymerase is a complex enzyme having approximately 12 to subunits and weighing about 500 kd. Five subunits of are common to all RNA polymerases. The largest subunits of each polymerase are homologous to each other and to the α, β and β′ subunits of *E. coli* RNA polymerase. There is no counterpart to the bacterial ct factor, and the eukaryotic RNA polymerases are consequentially unable to recognize or bind to their promoters. Weil et al (1979) discovered that RNA polymerases require the assistance of additional proteins not only to bind to promoter region but also to initiate transcription.

RNA polymerase I and its promoter: Apart from the basic subunits required for DNA transcription, **RNA polymerase** I requires specialized 4 Core promoter binding proteins - (It is called SL1, TIF-IB, Ribl in different species) and upstream binding factors called **UBF.** The core binding factor proteins ensures the positioning of the RNA polymerase I at the start point and can initiate transcription at a low basal frequency (Fig, 3.5). SL1 consists of four proteins. One of them, called **TBP** (TATA-binding protein), is a factor that is required also for initiation by RNA polymerases II and III. The UBF factors interact with the core proteins and greatly enhance the transcription frequency. RNA polymerase I most likely exists as a holoenzyme that contains most or all of the factors including the **TBP** required for initiation and is probably recruited directly to the promoter. The **RNA polymerase**

I promoter comprises of two separate regions. The first element surrounds the start point extending from -45 to +20, and is sufficient to initiate transcription. This element is unusually GC rich and includes a short AT rich conserved sequence called the **Inr.** However, presence of an upstream promoter element **(UPE)**, extending from -180 to -107 greatly enhances the efficiency of the primary promoter. The bipartite organization of RNA polymerase I promoter is seen in all organisms although the actual sequence may vary

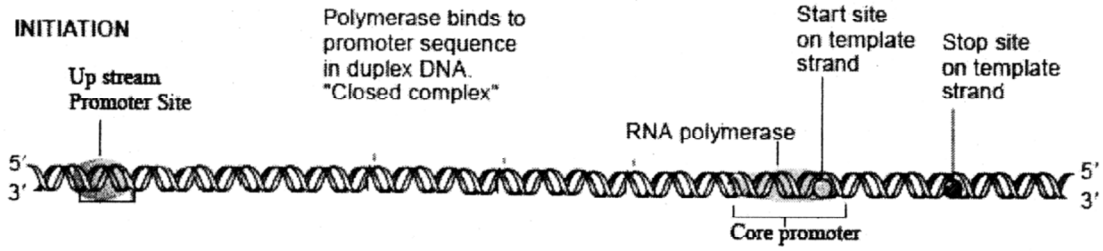


Fig 3.5: RNA polymerase I bipartite promoter

RNA polymerase II and its promoter: Several transcription factors and proteins are required by RNA polymerase II at all promoters for initiation of transcription. The subunits of RNA polymerase II and the general transcription factors are conserved among eukaryotes. Surrounding the startpoint of the core promoter region, the RNA polymerase II and the transcription factors assemble and bind with the DNA. However, the specificity and efficiency of binding to the promoter region depends on certain factors called the activators. The activators bind to target sequence ~ 100 bp upstream of startpoint or further away and influence the formation of initiation complex. Mutational studies led to the identification of three short consensus sequence centered around -30, -75 and -90 bp. Mutation of the TATA box, located at -30 does not prevent initiation but plays a crucial role in positioning the basal factors at precise location and positioning the RNA polymerase to start transcription from the right place.

The enhancing sequence elements located upstream of the start point influences the rate of transcription, probably, by interacting with the basal transcription factors. The CAAT box located at -75 or further away has a strong effect in determining the efficiency of the promoter. The consensus sequence at -90 is GGGCGG, which is very common and often exists in multiple copies.

RNA polymerase III and its promoter: RNA polymerase III uses both downstream and upstream promoters sequences. The promoters for 5S and tRNA genes lie downstream (identified in Xenopus laevis) of the startpoint between +55 and +80 bp. The promoters for snRNA (small nuclear RNA) genes lie upstream of the startpoint similar to other promoters. In both cases, the transcription factors

recognize individual promoter sequences, which in turn direct the binding of RNA poiymerase. Three types of RNA poiymerase III promoters are shown in Fig. 3.6.

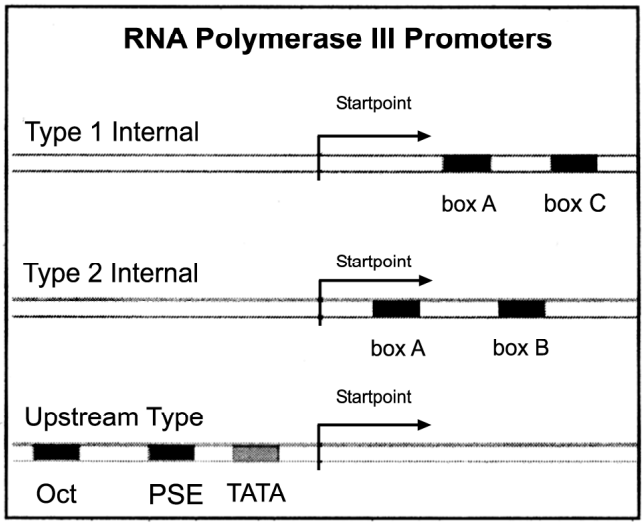


Fig 3.6 : RNA polymerase III Promoters

3.5.2 Internal promoters

The two internal type promoters have bipartite structure. Two short sequence elements remain separated by a variable sequence. The distance between boxA and boxB in a type 2 promoter vary extensively, but bringing the boxes too close inhibits function. Internal promoters bind to three different factors. TFIIIA (zinc finger proteins), TFIIIB (consists of TBP and two other proteins) and TFIIIC (large protein complex [>500 kD], has at least 5 subunits and the size is comparable to RNA poiymerase III itself). Most likely, at type 2 promoters, TFIIIC recognizes boxB, but binds to a more extensive region including both boxes A and B. At type 1 promoters, TFIIIA binds to a sequence that includes boxC, and this is required to enable TFIIIC to bind. In both cases, the binding of TFIIIC in turn enables TFIIIB to bind to a sequence surrounding the startpoint.

TFIIIA and TFIIIC removal of from the promoter by high salt concentration in vitro but allowing the presence of TFIIIB in the vicinity of the startpoint is sufficient to allow RNA poiymerase III to bind at the startpoint. So TFIIIB appears to be the only true initiation factor required by RNA poiymerase III and TFIIIA and TFIIIC are assembly factors, whose role is to assist the binding of TFIIIB at the right location. This sequence of events explains how the promoter boxes downstream can cause RNA poiymerase III to bind at the startpoint, farther upstream. TFIIIB includes the same protein, TBP that is present in SL1; this could be the subunit of TFIIIB that interacts directly with RNA poiymerase III. Any alteration in the upstream the internal promoter region can alter the efficiency of transcription.

Genes having upstream promoters has the TATA element which confer specificity for type III polymerases. Interestingly, some snRNAs are transcribed by poiymerase II while others are transcribed by poiymerase III. In both the cases,

the same type of promoters is present. Initiation commences at TATA site but the presence of PSE (Proximal Sequence Element) and Oct elements along with their factors increases the transcription efficiency. The PSE element may be essential at promoters used by RNA polymerase II, whereas it is stimulatory in promoters used by RNA polymerase IH

RNA polymerase III terminates transcription with U's immediately after a GC-rich region but there is no fonnation of stem-loop structure. Termination usually occurs at the second U within a run of four U's, but some molecules terminate with 3 or even 4 U's.

3.5.3 Eukaryotic promoters

Most eukaryotic genes have conserved DNA sequences at -25 to -30 from the start point of RNA transcription called TATA box or Hogness box that specify a particular start point during RNA transcription.. By convention, the sequence is given on the non-template strand. The TATA box has a consensus sequence 5' -TATAAAA- 3'. Mutation at this region affects the transcriptional process. Promoters that lack TATA box, there is no definite initiation point but appears to be controlled by a CT-rich area, called the initiator element (Inr) having a consensus sequence TCA, at +1 coupled with a down stream promoter element (DPE) at about +28 to +34.

Further upstream, many promoters have a CAAT box, found approximately at -75 positions of many genes but the position may vary. The consensus sequence is 5' -GGCCAATGT- 3'. Alteration of CAAT box markedly reduces transcriptional rate.

Additional sequences like GC box (5' -GGGCGG- 3') is found at -90 positions and often there are more than one copy of GC elements. This sequence may function in either orientation i.e. 5' -GGGCGG- 3' or 5'GGCGGG-3'. Interestingly, promoter of one gene may vary considerably from the other gene and no element is essential for all promoters.

3.5.4 Enhancer & Silencer

Enhancer sequences influences the transcriptional rate of a gene. Enhancer may function in either orientation and is usually located far away from the actual initiation point of transcription, sometime 1000 bp apart from the promoter sequence, usually upstream the start point. In animal cells, enhancers can be located down stream from the initiation point.

Similarly there are sequences that have the same properties like that of the enhancers but they repress rather that activate gene transcription. The elements

are called silencer elements. Silencers are less common than enhancers. There are no consensus sequences for eukaryotic enhancer or silencer element and their exact mechanism of action is still not clear. It is proposed that specific protein factors interact with the enhancer elements and folds DNA in a way so that enhancer elements interact with the transcriptional factors and regulators in the promoter region and subsequently activate (enhances) or repress (silencer) RNA transcription.

3.5.5 Initiation of transcription by RNA polymerse II

Robert Roder (1979) discovered that eukaryotic RNA polymerase II could not initiate transcription unless some other protein factors are added to the reaction mixture *in vitro.* Biochemical analysis revealed the existence of specific proteins called **Transcription** factors that are required by RNA polymerase II to start RNA synthesis. Some of these factors bind directly with the DNA while others appear to bind to the RNA polymerase. Transcription factors are usually designated by letters TF (for transcription factor) followed by Roman numeral - I or II or III to indicate the type of polymerase they bind and finally followed by a letter characterizing the type of factor, e.g. TFIIB, TFIIE etc.

Two general types of transcription factors are involved:

1. Those proteins that are required by all polymerase II and are called **Basal Transcription factors.** At least five basal transcription factors are required for initiation of transcription ob RNA polymerase II *in vitro* system.

2. Additional transcription factors that bind to DNA sequences and control the expression of specific genes and thus responsible for regulation of gene expression.

The first step in formation of a initiation complex is the binding of a complex called **TFIID.** One of the subunit of TFIID recognizes and binds to TATA box and this subunit is called **TATA-binding protein (TBP).** In essence, TFIID appears to be similar to the sigma factor in RNA polymerase (Fig. 3).

TBP is a 30 kd protein that binds to the minor groove of the DNA at the TATA promoter sequence. Binding is 10^5 times more tightly to TATA than with other sequences. TBP is saddle shaped protein with two similar domains. The TATA box binds to the concave surface of TBP. This binding induces large conformational changes. The minor

groove widens from 5Å to 9Å bat does not break the hydrogen bonds. This substantial unwinding of the minor groove enables extensive contact with the anti-parallel â strands on the concave side of the TBP. Immediately outside the TATA box, classical B-DNA resumes. This complex is distinctly asymmetric. The asymmetry is crucial for specifying a unique start site and ensuring that transcription proceeds unidirectionally.

The surface of the TBP saddle provides docking sites for the bindings of other transcription factors.

DNA bound TBP of TFIID first recruits **TFIIA** which further enhances the binding of the TBP.

Then **TFIIB** is recruited forming a complex **TBP-TFIIB** at the promoter region.

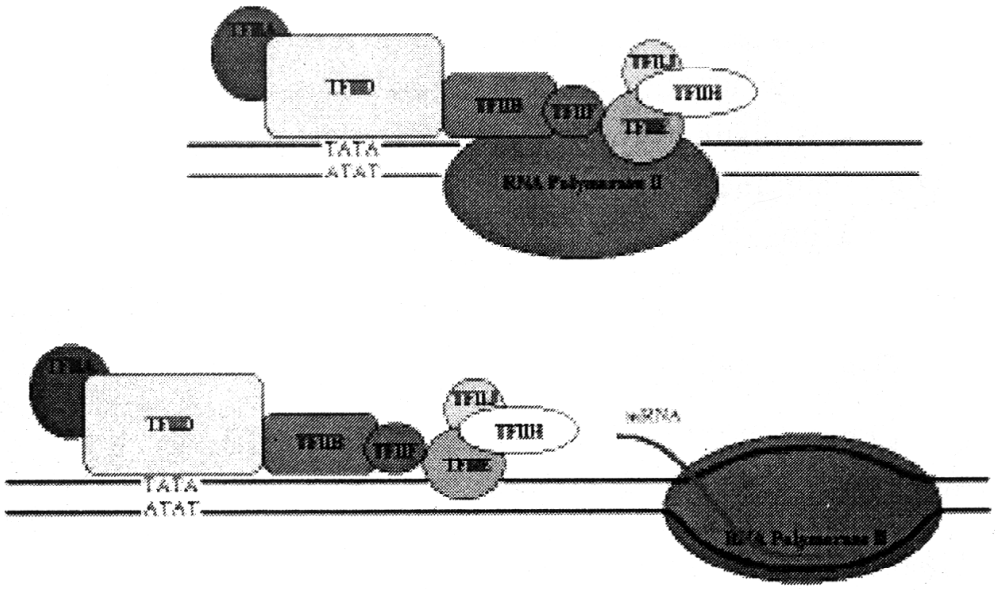
Binding of TFIIB sets the stage for the binding of RNA polymerase, which binds to the TBP-TFIIB complex in association with a third factor, **TFIIF.**

Two additional factors - **TFIIE** and **TFIIH** bind to the initiation complex and appear to be necessary for initiation of transcription. TFIIH is multi-subunit factor. First two subunits has hehcases activity, which my unwind DNA around the initiation site. Another subunit of TFIIH is a protein kinase that phosphorylates repeated sequences present in the C-terminal domain of the largest subunit of RNA polymerase II.

Phosphorylation of the C-terminal domain of the RNA polymerase II releases the enzyme from its association with the initiation complex, allowing it to proceed along the template as it elongates the growing RNA chain.

The various transcription factors described here represents the minimal system required for transcription *in vitro;* additional factors may be needed within the cell. Furthermore, RNA polymerase II appears to remain associated with some transcription factors *in vivo* prior to the assembly of a transcription complex on DNA. Such preformed holoenzyme complexes probably are recruited to a promoter via direct interaction with TFIID. What actually occurs within the eukaryotic cell

during RNA transcription still needs to be worked out. Moreover, the functions of many of the basal transcription factors are still unknown and unanswered.



3.5.6 Elongation

Placement of the first ribonucleotide with its corresponding deoxyribonucleotide in the DNA each new ribonucleotide attaches to the 3′ -OH group of the previous ribonucleotide. RNA synthesis is continuous and proceeds in the 5′ to 3′ direction. Energy is derived from the cleaving ofihe two phosphate group of the new incoming ribonucleotide that pairs with the DNA and bjnds to the 3′ -OH of the previous ribonucleotide residing within the RNApolymerase. The double stranded RNA-DNA hybrid is very transient. As RNA polymerase moves forward, the nascent RNA separates from the DNA. At any given time, the number of nucleotides of RNA that remain paired with the DNA template may be as many as 2 to 12. The unfolded DNA rapidly rewinds after RNA synthesis is over at that point and therefore nicking of the DNA is not required to release the tension of unwinding. Moreover, only 18 bases are unwounded at any time. The rate of synthesis is not always consistent as other proteins present in the cell often influence the rate of RNA synthesis.

Occasionally, RNApolymerase II stalls synthesis, may be due to some configuration change in the polymerase or my encounter a nucleotide sequence that causes the polymerase to stall on the DNA. When this happens, **TFIIS** causes the RNA polymerase to move backward, and then **TFIIS** removes the 3′ end of the RNA, permitting the RNA polymerase to attempt elongation again over the point

where the stall occurred. It may be noted that eukaryotic RNA polymerases have 3′ — 5′ exonuclease activity.

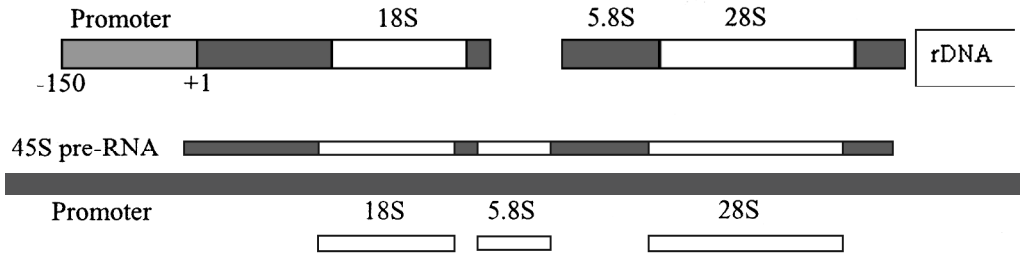
3.5.7 Termination

The termination of transcription in eukaryotic genes is less well understood than in bacterial genes. The three eukaryotic RNA polymerases use different mechanisms for termination. RNA polymerase I requires a termination factor, like the rho factor utilized in termination of some bacterial genes. Unlike rho, which binds to the newly transcribed RNA molecule, the termination factor for RNA polymerase I binds to a DNA sequence downstream of the termination site. RNA polymerase III ends transcription after transcribing a terminator sequence that produces a string of Us in the RNA molecule, like that produced by the rho-independent terminators of bacteria. Unlike rho-independent terminators in bacterial cells, RNA polymerase III does not requre that a hairpin structure precede the string of Us. In many of the genes transcribed by RNA polymerase II, transcription can end at multiple sites located within a span of hundreds or thousands of base pairs.

3.5.8 Transcription by RNA polymerse I

RNA polymerase I & III apply the same basic mechanism of transcription and only differ in the recruitment of specialized transcription factors that recognizes and associate with appropriate promoter sequences.

RNA polymerase I is solely responsible for the transcription of ribosomal RNA genes, which are present in tandem repeats. Transcription of these genes yields a large **45S** pre-RNA, which is then processed to yield the 28S, **18S,** and **5.8S** rRNAs.



The promoter sequence of ribosomal RNA gene is recognized by two transcription factors, **UBF** (upstream binding factor) and **SLl** (selectivity factor 1), which together bind to the promoter site and then recruit polymerase I to form the initiation complex. The SLl is composed of four protein subunits one of which is TBR The promoter here lacks the TATA box, and TBP therefore do not bind to

specific promoter sequences. Instead, the association of TBP with ribosomal RNA genes is mediated by the binding of other proteins in the SL1 complex to the promoter, situation similar to the association of TBP with the Inr sequences of polymerase II genes that lack TATA boxes.

The genes for tRNAs, 5S rRNA and some of the small RNAs involved in splicing and protein transport are transcribed by polymerase III. Interestingly, the promoters of these genes lie within, rather than upstream of the transcribed sequence. The well studied 5S RNA of *Xenopus* revealed that at first TFIIIA binds to the promoter followed by TFIIIC, TFIIIB and then the polymerase. In case of tRNA, the promoter sequence is recognized by TFIIIC and not TFIIIA. The multimeric TFIIIB protein appears to be the most common factor for all the polymerases to initiate transcription.

Ribosomal RNA (rRNA)

There are 4 kinds. In eukaryotes, these are

18S rRNA. One of these molecules, along with some 30 different protein molecules, is used to make the **small subunit** of the ribosome.

28S, 5.8S, and 5S rRNA. One each of these molecules, along with some 45 different proteins, are used to make the **large subunit** of the ribosome.

The S number given for each type of rRNA reflects the rate at which the molecules sediment in the ultracentrifuge. The greater the number, the larger the molecule (but not proportionally).

The 28S, 18S, and 5.8S molecules are produced by the processing of a single primary transcript from a cluster of identical copies of a single gene. The 5S molecules are produced from a different cluster of identical genes.

Transfer RNA (tRNA)

There are some 32 different kinds of tRNA in a typical eukaryotic cell.

Each is the product of a separate gene.

They are small (~4S), containing 73-93 nucleotides.

Many of the bases in the chain pair with each other forming sections of double helix.

The unpaired regions form 3 loops.

Each kind of tRNA carries (at its 3' end) one of the 20 **amino acids** (thus most amino acids have more than one tRNA responsible for them).

At one loop, 3 unpaired bases form an **anticodon.**

Base pairing between the anticodon and the complementary codon on a mRNA molecule brings the correct amino acid into the growing poiypeptide chain. Further details of this process are described in the discussion of translation.

Messenger RNA (mRNA)

Messenger RNA comes in a wide range of sizes reflecting the size of the poiypeptide it encodes. Most cells produce small amounts of thousands of different mRNA molecules, each to be translated into a peptide needed by the cell.

Many mRNAs are common to most cells, encoding "housekeeping" proteins needed by all cells (e.g. the enzymes of glycolysis). Other mRNAs are specific for only certain types of cells. These encode proteins needed for the function of that particular cell (e.g., the mRNA for hemoglobin in the precursors of red blood cells).

Small Nuclear RNA (snRNA)

Approximately a dozen different genes for snRNAs, each present in multiple copies, have been identified. The snRNAs have various roles in the processing of the other classes of RNA. For example, several snRNAs are part of the spliceosome that participates in converting pre-mRNA into mRNA by excising the intrbns and splicing the exons.

Small Nucleolar RNA (snoRNA)

As the name suggests, these RNAs (there are probably over 100 of them) are found in the nucleolus where they are responsible for several functions:

Some participate in making ribosomes by helping to cut up the large RNA precursor of the 28S, 18S, and 5.8S molecules.

Others chemically modify many of the nucleotides in these molecules, e.g., by adding methyl groups to ribose.

Still others serve as the template for the synthesize of telomeres.

In vertebrates, the snoRNAs are made from **introns** removed during RNA processing.

3.6 Transcription factors

Transcription factor is, a protein that works in concert with other proteins to either promote or'suppress the transcription of genes. More specifically, transcription factors regulate gene expression. They bind to specific sequences of DNA upstream or downstream to the gene they regulate and then either enhance or repress transcription of these genes by assisting or blocking RNA polymerase binding respectively.

A defining characteristic of transcription factors is that they contain a **DNA binding domain** (DBD) which bind to gene specific regulatory sites *{e.g.,* promoter sequences). In addition, transcription factors often contain a second domain that sense external signals and in response transmit these signals to the rest of the transcription complex resulting in up or down regulation of gene expression. In some cases the DBD and signal sensing domains reside on separate proteins that associate within the transcription complex to regulate gene expression.

Other proteins such as coactivators, chromatin remodelers, histone acetylases, deacetylases, kinases, and methylases, also playing crucial roles in gene regulation but they lack DNA binding domains, and therefore are not classified as transcription factors.

3.6.1 Regulation

Gene regulation is a highly complex process as it is dependent upon a number of factors. *In vitro* experiments suggested that the assembly of transcription factors dictated by the DNA sequence. However, epigenetic information present on DNA appears to play an important role in transcriptional activation.

3.6.2 Classes of transcription factors

There are three classes of transcription factors: a) Upstream transcription factors are proteins that bind somewhere upstream of the initiation site to stimulate or repress transcription, b) Inducible transcription factors are similar to upstream transcription factors but require activation or inhibition, c) General transcription factors

General transcription factors are involved in the formation of a preinitiation complex that participate in the transcription of class II genes to mRNA templates.Tata binding protein, (TBP) is a general transcription factor, that binds to the TATAAbox, the motif that resides upstream from the coding region in all genes. TBP is responsible for the recruitment of the RNA Pol II holoenzyme, the final event in transcription initiation. The most common general transcription factors are TFIIA, TFIIB, TFIID, TFIIE, TFIIF, TFIIH. They are ubiquitous and interact with tire core promoter region surrounding the transcvription start site(s) of all class II genes.

TFIIA

TFIIA consists of two subunits in yeast and three in humans and drosophila (two subunits are derived from a precursor protein). TFIIA binds directly to TBP and stabilizes its binding to DNA, perhaps through direct contact with the DNA. TFIIA binding does not preclude TFIIB binding or other components of the transcription complex. However, binding of TFIIA to TBP is mutually exclusive with binding of some negative regulatory proteins. TFIIA acts as an anti-repressor, stabilizing TFIID binding by blocking repressors of transcription that inhibit binding of other transcription factors or that remove TBP from the DNA. Activation of transcription may be dependent on this TFIIA function.

TFIIB :

It is **a** single subunit measuring 35 to 40 kd. **The protein possesses a** zinc finger domain at the N-terminus and a direct repeat in a proteolytically stable C-terminal domain. TFIIB binds directly to TBP, recruits RNA polymerase II, in part through an interaction with the small subunit of TFIIF. Several acidic activators can bind TFIIB in vitro. **The protein probably** stabilizes TBP binding to TATA element and is required for association of RNA polymerase II to the initiation complex. TATA-Binding Protein is shown in green, TFIIB in red (Fig. 3.7).

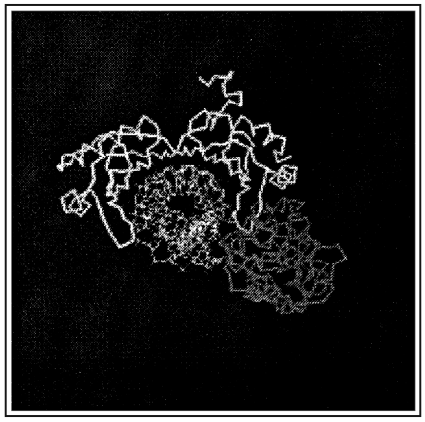


Fig. 3.7: 3D structure of TFIIB

TTFIID is a multi-component (>5 subunits) transcription factor that recognizes and binds to the promoter DNA. TFIID consists of a DNA binding subunit that recognizes the TATA element and is therefore designated TATA-binding protein (or TBP), as well as several TBP-associated factors (or TAFs). TFIID helps in recruiting the rest of the factors through a direct interaction with TFIIB. The TBP subunit of TFIID is sufficient for TATA element binding and TFIIB interaction, and can support basal transcription. However, this basal transcription reaction does not respond to upstream transcription activators. Many of these regulatory factors interact with TBP or TAFs in various in vitro assays. TBP also interacts directly with TFIIA.

Features: TBP consists of a 180 ammo acid domain that is sufficient for activity. This domain is made up of an imperfectly repeated sequence, and the repeats are reflected in the symmetry of the molecule (see picture below). The protein resembles a saddle, with the inner surface contacting DNA and the outer surface presumable making protein-protein contacts. TFIID binding is thought to be the first step in transcription initiation. Some of the TAFs also bind to initiator elements. TBP is also a component of the RNA polymerase I and RNA polymerase III transcription complexes.

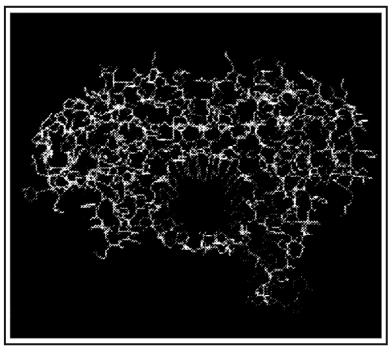
TFIIE: It has two subunits, probably a tetramer consisting of two molecules of each subunit. The large subunit has a zinc finger domain. TFIIE modulates the helicase and kinase activities of TFIIH and the two factors show species-specific interactions. It recruits TFIIH to the initiation complex and modulates TFIIH kinase and helicase activities. Appears to be required for escape of the RNA polymerase into elongation mode (promoter clearance).

TFIIF: The molecule has two subunits. TFIIF binds directly to RNA polymerase II. TFIIF is necessary for RNA polymerase II to stably associate with the TFIIF-TFIIB-promoter complex. There is a protein interaction between the small subunit and TFIIB in vitro and a genetic interaction between the large subunit and TFIIB. It Helps recruit RNA polymerase II to the initiation complex in collaboration with TFIIB. TFIIF is a component of the yeast holoenzyme and mediator complexes. Promotes transcription elongation, may remain associated with the elongating polymerase.

TFIIH: Mammalian and yeast TFIIHs have at least six subunits. Most subunits are now cloned, although not all are published.Yeast subunits: SSL2(RAD25), RAD3, SSL1, TFB1, TFB2, TFB3, TFB4. In addition to the TFIIH subunits, there is an associated complex consisting of a CDC-like kinase and cyclin-like subunit. This kinase complex is sometimes referred to as TFIIK. Yeast subunits: KIN28 and CCL1. The two largest TFIIH subunits are ATP-dependent helicases of opposite polarity. Two of the smaller subunits have possible zinc finger domains. TFIIH appears to be dependent upon TFIIE for incorporation into the initiation complex. The associated kinase (TFIIK) complex can phosphorylate the C-terminal domain of the pol II largest subunit. TFIIH is essential for promoter melting (separation of the two DNA strands) and/or promoter clearance (i.e. for pol II to break free of the initiation complex into elongation mode). Surprisingly, TFIIH also is essential for Nucleotide Excision Repair (NER) of damaged DNA. The relationship between TFIIH's transcription and repair functions is not understood yet.

SPT16 **:** Subunit of the heterodimeric FACT complex (Sptl6p-Pob3p), facilitates RNA Polymerase II transcription elongation through nucleosomes by destabilizing and then reassembling nucleosome structure of the DNA. *Sptl6p* has been found to physically interact with *PobSp* (which has sequence similarity to some HMG chro-matin-associated proteins) and the catalytic subunit of DNA polymerase alpha, *PolJp.* Some of the *Sptl6p/Pob3p* complex in the cell is chromatin associated, and some copurifies with the DNA polymerase alpha-primase complex. The N-terminal third of *Sptl6p* is necessary for the maintenance of chromatin repression, but not for activation of genes. Homologs of *SPT16* have been found in *K. lactis* and human

TBP (TATA binding protein))is a DNA binding protein that binds sequence specifically to the TATA box. It is vital for all eukaryote transcription, and will in some cases be forced to bind non sequence specifically. It is involved in DNA melting (double strand separation) and bends the DNA by 80° (the AT-rich sequence to which it binds facilitates easy melting). The TBP is an unusual protein in that it binds the minor groove using a â sheet. TBP is â subunit of the eukaryotic transcription factor TFIID. TFIID is the first protein to bind to DNA during the formation of the pre-initiation transcription complex of RNA polymerase II (RNA Pol II). Binding of TFIID to the TATA box in the promoter region of the gene initiates the recruitment of other factors required for RNA Pol II to begin transcription. Each of these transcription factors are formed from the interaction of many protein subunits, indicating that transcription is a heavily regulated process. TBP is also a necessary component of RNA polymerase I and RNA polymerase III, and is perhaps the only common subunit required by all three of the RNA polymerases.



Yeast TBP bound to DNA.

When TBP binds to a TATA box within the DNA, it distorts the DNA by creating a nearly 90 degree bend. The distortion is accomplished through a great amount of surface contact between the protein and DNA. TBP binds with the negatively charged phosphates in the DNA backbone through positively charged lysine and arginine amino acid residues. The sharp bend in the DNA is produced through projection of four bulky phenylalanine residues into the minor groove. As the DNA bends, its contact with TBP increases, thus enhancing the DNA-protein interaction.

The strain imposed on the DNA through this interaction initiates melting, or separation, of the strands. Because this region of DNA is rich in adenine and thymine residues, which base pair through only two hydrogen bonds, the DNA strands are more easily separated. Separation of the two strands exposes the bases and allows RNA polymerase II to begin transcription of the gene.

FACT complex (Facilitates chromatin transcription complex): An abundant nuclear complex, which was originally identified in mammalian systems as a factor required for transcription elongation on chromatin templates. The FACT complex has been shown to destablilize the interaction between the H2A/H2B dimer and the H3/ H4 tetramer of the nucleosome, thus reorganizing the structure of the nucleosome. In this way, the FACT complex may play a role in DNA replication and other processes that traverse the chromatin, as well as in transcription elongation. FACT is composed of two proteins that are evolutionarily conserved in all eukaryotes and homologous to mammalian Sptl6 and SSRP1. In metazoans, the SSRP1 homolog contains an HMG domain; however in fungi and protists, it does not. For example, in S. cerevisiae the Pob3 protein is homologous to SSRP1, but lacks the HMG chromatin binding domain. Instead, the yFACT complex of Sptl6p and Pob3p, binds to nucleosomes where multiple copies of the HMG-domain con^{+}aining protein Nhp6p have already bound, but Nhp6p does not form a stable complex with the Sptl6p/Pob3p heterodimer.

3.6.3 Structural binding motifs

DNA-Binding motifs

In general, transcription activators have two domains: one domain binds specifically to DNA and the second domain interact with other transcription factors. Four different types of transcription activators have been identified and their DNA binding motif characterized.

1. **Zinc finger domains** : have repeats of cystein and histidine residues, which interact with Zn++ ions to fold in a fingerlike fashion to grasp the DNA. E.g. TFIIIA, Steroid hormone receptors also have Zinc finger domains, which regulate gene transcription in response to hormones like estrogen & testosterone.

2. **Helix-turn helix :** In this type of activators, one helix makes most of the contact with DNA while the other helix lies across the complex to stabilize the interaction. E.g. Homeodomain protein, catabolic activator protein (CAP) in *Drosophila*
   Note: Homeodomain proteins play critical roles in the regulation of gene expression during embryonic development.

1. **Leucine Zipper:** contain DNA binding domains formed by dimerization of two polypeptide chains. The leucine zipper contains four or five leucine residues spaced at intervals of seven residues resulting in the exposure of hydrophobic side chains at one side of the helical region. This hydrophobic region of the domains serves as the point of dimerization of the two domains of the transcription activator, thereby interlocking the DNA strand within it by the interaction of positively charged lysine and arginine with the DNA.

2. **Helix loop helix:** In the type, the amino acid sequence is similar to leucine zipper domains, except that their dimerization domains are formed by two other domains separated by a loop.
   An interesting feature is that both leucine zipper and helix loop helix transcription factors is that different members of these families can dimerize with each other. Such dimerization gives rise to formation of an array of transcription activators that differ in DNA specificity and also binding to other transcription factors. These two transcription activators play an important role in regulation tissue specific gene expression.
   The activation domains are not well characterized as their DNA binding domains. Some of these domains are acidic, some basic. But they somehow interact with basal transcription factors like TFIID, TFIIB and facilitate the initiation of transcription of specific genes.

3. **MADS box** is a conserved sequence element found in a family of transcription factor encoding genes. The length of the MADS-box is defined differently by various authors, but typical lengths suggested are 168 base pairs or 180 base pairs. The element encodes the MADS-domain that have DNA-binding properties. In plants, MADS-box genes have undergone a substantial radiation. In Arabidopsis the MADS box genes SOC and FLC have been shown to have an important role in the integration of molecular flowering time pathways. These genes are essential for the correct timing of flowering, and help to ensure that fertilization occurs at the time of maximal reproductive potential.

3.6.4 Protein-binding motifs

STAT

The Signal Transducers and Activator of Transcription (STAT) proteins regulate many aspects of cell growth, survival and differentiation. The transcription,factors of this family are activated by the Janus Kinase JAK and dysregulation of this pathway is frequently observed in primary tumors and leads

to increased angiogenesis and enhanced survival of tumors. Knockout studies have provided evidence that STAT proteins are involved in the development and function of the immune system and play a role in maintaining immune tolerance and tumor surveillance.

Function of STAT proteins

STAT proteins were originally described as latent cytoplasmic transcription factors that require phosphorylation for nuclear retention. The unphosphorylated STAT proteins shuttle between the cytosol and the nucleus waiting for its activation signal. Once the activated transcription factors reach the nucleus, they bind to a consensus DNA-recognition motif called gamma activated sites (GAS) in the promoter region of cytokine-inducible genes and activate transcription of these genes.

Activation of STAT proteins

Extracellular binding of Cytokines induces activation of the intracellular Janus kinase that phosphorylates a specific tyrosine residue in the STAT protein which promotes the dimerization of STAT monomers via their SH2 domain. The phosphorylated dimer is then actively transported in the nucleus via importin a/b and RanGDP complex. Once inside the nucleus the active STAT dimer binds to cytokine inducible promoter regions of genes containing gamma activated site (GAS) motif and activate transcription of this proteins. The STAT protein can be dephosphorylated by nuclear phosphatases which leads to inactivation of STAT and the transcription factor becomes transported out of the nucleus by exportin crml/RanGTP.

Unit 4 Post Transcriptional Modification of RNA

Structure

4.1 Introduction

In prokaryotes, tRNA and rRNA undergo modification after being transcribed but mRNA do not get the opportunity to undergo modification as the transcription and translation processes are coupled. In Eukaryotes, all the three types of RNAs (mRNA, tRNA and rRNA) undergo post transcriptional modifications.

4.2 Post transcription modification of r RNA

Prokaryotes have three ribosomal rRNA (23S, 16S and 5S) equivalent to the eukaryotic 28S, 18S and 5S rRNAs of eukaryotic cell. In both the cell types, the processing of a single pre-RNA transcript produces different rRNAs. In eukaryotes, only the 5S RNA does not undergo much modification as they are synthesized from a separate gene. The steps of cleavage to remove the introns and obtaining the final product are shown in the figure 4.1.

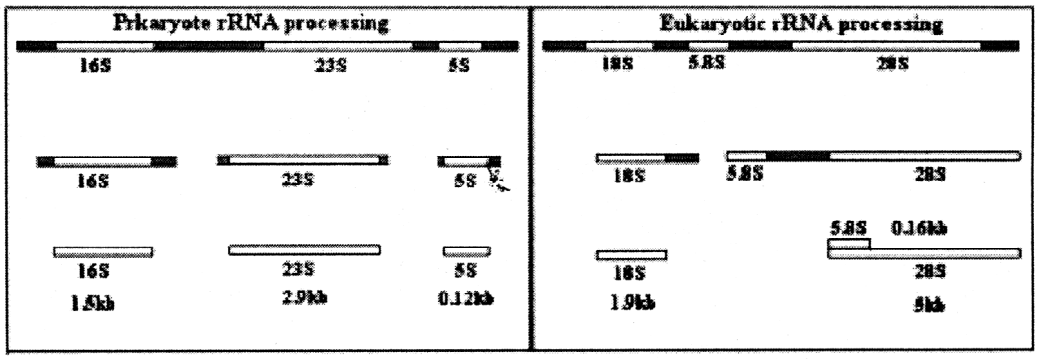


Fig. 4.1: Processing of newly synthesized rRNA

In eukaryotes, after the 5.8S RNA is produced, it is hydrogen bounded to 28S RNA. Further, rRNA processing involves the addition of methyl groups and sugar moieties to specific nucleotides, but the function of these modifications is unknown.

4.3 Post transcriptionaJ modification of tRNA

Prokaryotes and eukaryotes synthesize tRNA as precursor molecules. Some pre t-RNA transcripts have several tRNA sequences, which are cleaved and modified to obtain mature functional tRNA. Some tRNA sequences are present within the pre-RNA transcripts.

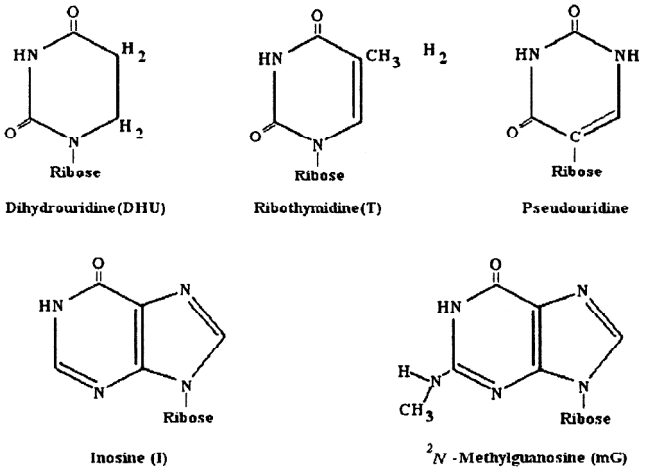
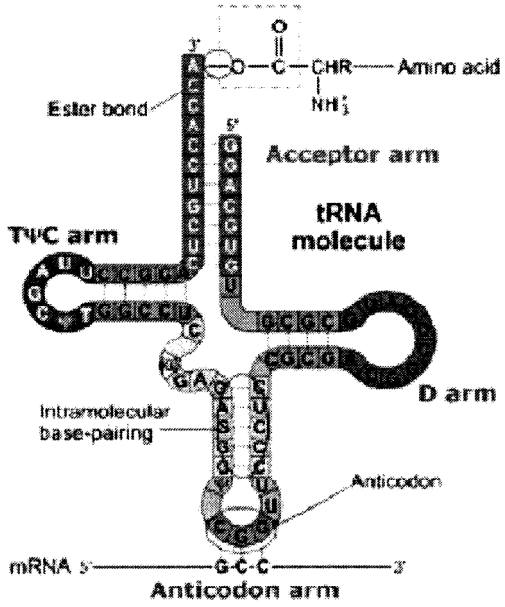


Fig. 4.2a : Modified bases



The 5′ end of all tRNAs is modified by RNase P (ribozyme). Conventional RNase modify the 3′ of end of tRNA which is the by followed by addition of -CCA nucleotides. Some tRNA has the information of-CCA already encoded in the DNA. All tRNAs have -CCA sequence at their 3' end. Moreover, 10% of the bases in tRNAs are altered to yield a variety of modified nucleotides (Fig. 4.2a & 4.2b) at specific positions in tRNA molecules but their exact function is still not clear.

4.4 Transcriptional modification of mRNA

mRNA in all organisms can be distinguished into three primary regions (Fig. 4.3).

a) The 5′ untranslated region (5′ UTR) also called the leader sequence do not code for any amino acids but carry vital information for subsequent mRNA modifications and translation. In bacterial mRNA, this region contains a consensus sequence called the Shine-Dalgarno sequence, which serves as the ribosomebinding site during translation. It is found approximately seven nucleotides upstream of the first codon translated into an amino acid (called the start codon). Eukaryotic mRNA has no equivalent consensus sequence in its 5′ untranslated region, ribosomes bind to a modified 5′ end of mRNA.

b) The next section of mRNA is the **protein-coding region,** which comprises the codons that specify the amino acid sequence of the protein. The protein-coding region begins with a start codon and ends with a stop codon.

c) The last region of mRNA is the **3′untranslated region** (3′ UTR), a sequence of nucleotides at the 3′ end of mRNA that is not translated into protein. The 3′ untranslated region affects the stability of mRNA and the translation of the mRNA protein-coding sequence.

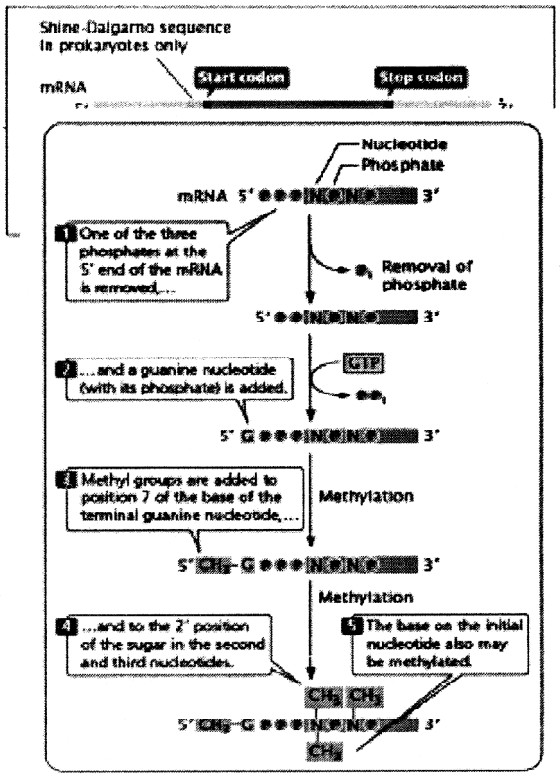


Fig. 4.3: Structure of mRNA molecule

Processing of mRNA

In eukaryotes, transcription and translation are separated in both time and space and this separation provides an opportunity for eukaryotic RNA to be modified before it is translated: The initial transcript of protein-encoding genes of eukaryotic cells is called pre-mRNA Eukaryotic mRNA undergoes extensive alteration after transcription. Changes are made at both the 5' end, the 3' end, and also in the protein-coding section of the RNA molecule

The addition of the 5′Cap

The 5′ end of all eukaryotic pre-mRNAs are modified by the addition of a extra nucleotide, followed by methylation (addition of CH_3 group) to the 2′-OH group of the sugar of one or more nucleotides at the 5′ end - a process termed as capping and the structure is called 5′ **cap** (Fig. 4.4).

Capping occurs a few moments after the commencement of transcription. Transcription starts with a nucleoside triphosphate which is usually a purine (A or G). The capping process involves the addition of GTP molecule in the reverse direction to the 5′ terminal residue. Addition of the 5′ terminal G is catalyzed by a nuclear enzyme, guanylyl transferase. Subsequently, methyl groups are added to the N7of GTP molecule.

The next step is to add another methyl group, to the 2′ -O position of the penultimate base (which was actually the original first base of the transcript before any modifications were made). This reaction is catalyzed by another enzyme (2′ -O-methyl-transferase). A cap that possesses this single methyl group is known as a cap 0 and is found in unicellular animals.

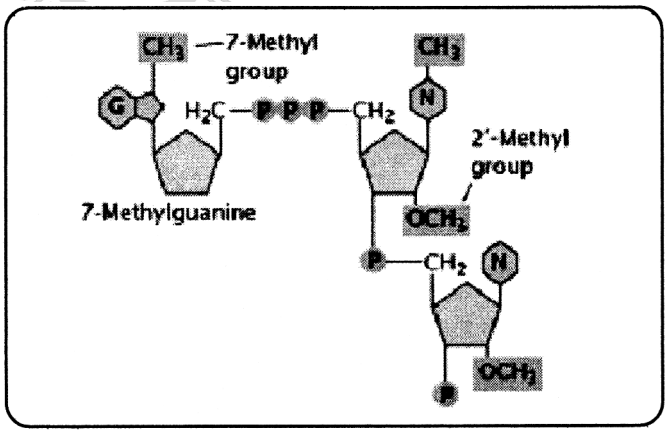


Fig 4.4 : Structure of 5' cap

A cap with the two methyi groups is called cap 1. This is the predominant type of cap in all Fig. 4,4: Structure of 5' cap eukaryotes except unicellular organisms. Methylation may also occur at second and third positions and is called cap 2. If the second nucleotide is adenine, then the methylation occurs at the N^6 position since it already has a methyl group at 2′ position. The third-base modification is always a 2′-O ribose methylation. This cap usually represents less than 10-15% of the total capped population.

In addition to the methylation involved in capping, a low frequency of internal methylation occurs in the mRNA only of higher eukaryotes. This is accomplished by the generation of N^6 methyladenine residues at a frequency of about one modification per 1000 bases. There are 1-2 methyladenines in a typical higher eukaryotic mRNA, although their presence is not obligatory, since some mRNAs do not have any.

The cap blocks the 5′ end of mRNA. The 5' cap helps to align mRNAs on the ribosome during translation initiation. Cap binding proteins recognize the cap and attach to it; a ribosome then binds to these proteins and moves downstream along the mRNA until the start codon is reached and translation begins. The presence of a 5′ cap also increases the stability of mRNA and influences the removal of introns.

4.5 The addition of the Poly(A) tail on mRNA

Most mature eukaryotic mRNAs have from 50 to 250 adenine nucleotides at the 3′ end [a **poiy (A) tail].** These nucleotides are not encoded in the DNA but are added after transcription **(Fig.** 4.5) in a process termed polyadenylation. Polymerase II encoded genes transcribes well beyond the end of the coding sequence (more than 1000) at the 3′ end which is then cleaved and the poly(A) tail is added.

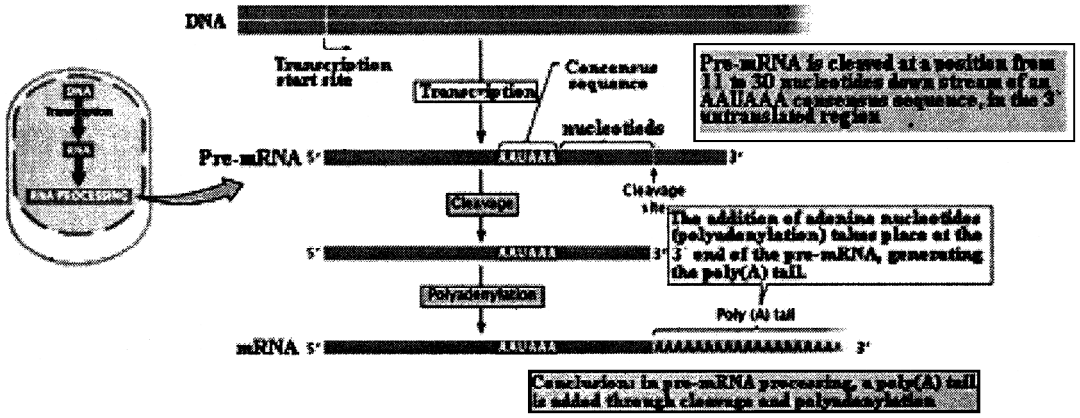


Fig. 4.5: Addition of poly(A) tail at 3′ end of pre-mRNA

The site of cleavage at the 3′ end of pre-mRNA is determined by specific upstream and downstream sequences **(Fig.** 6). Upstream consensus is usually AAUAAA that determines the site of the cleavage and resides 11 to 30 nucleotides

upstream of the cleavage site. A sequence rich in Us (or Gs and Us) is typically present down-stream of the cleavage site. In mammals, 3′cleavage and the addition of the poly(A) tail requires a complex consisting of several proteins:

a) Cleavage and polyadenylation specificity factor (CPSF);

b) Cleavage stimulation factor (CstF);

c) At least two cleavage factors (CFI and CFII) ;

d) Polyadenylate polymerase (PAP).

CPSF binds to the upstream AAUAAA consensus sequence, whereas CstF binds to the downstream sequence (Fig. 4.6). CstF after cleaving the pre-mRNA leave the complex. The cleaved 3′ end of the pre-mRNA is then degraded. CFSF and PAP remain bound to the pre-mRNA and carry out polyadenylation. After the addition of approximately 10 adenine nucleotides, a poly(A)-bindirig protein (PABH) attaches to the poly(A) tail and increases the rate of polyadenylation. As more of the tail is synthesized, additional molecules of PABII attach to it.

The poly(A) tail confers stability to many mRNAs, increasing its half life, making it available for longer time for the translational process, before it is degraded by cellular enzymes. The stability conferred by the poly(A) tail is dependent on the proteins that attached to the tail.

Eukaryotic mRNAs that lack a poly (A) tail depend on a different mechanism for 3′ cleavage. It requires the formation of a hairpin structure with the aid of a small ribonucleoprotein particle (snRNP)

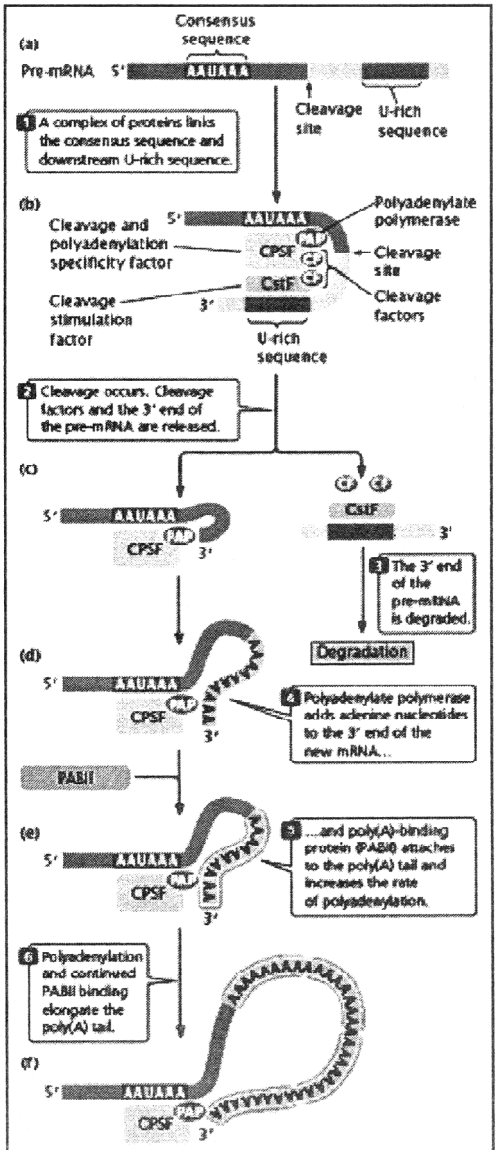


Fig. 4.6 : Process of binding of CPSF and Cstf

called U7 (Fig. 4.7). U7 contains an snRNA with nucleotides that are complementary to a sequence on the pre-rnRNA just downstream of the cleavage site. U7 most likely binds to this complementary sequence. A hairpin-binding protein binds to the hairpin structure and stabilizes the binding of U7 to the complementary sequence on the pre-mRNA and cleave the 3' non coding sequences

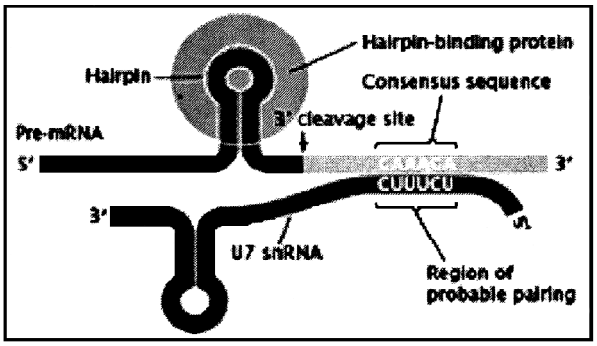


Fig. 4.7 : Ribonucleoprotein particle

4.6 RNA Splicing

4.6.1 Intron Removal

Most eukaryotic genes are interrupted by non-coding sequences called introns. The intron consists of GU at 5′ end and AG at 3′ end, while a branch site (A) in the middle and a (py)n, meaning a stretch of primidine near the 3′ end. During mRNA processing, the introns are precisely excised from the mature mRNA. At first, a protein complex called spliceosomes cleaves the 5′ end of the intron. In the second step, the 5′ end of the intron is joined to an adenine residue within the intron at the 2′-OH group of the adenine nucleotide to form a 2′, 5′ -phosphodiester linkage, which is quite unusual bond. The resulting intermediate is a lariat like structure. Next the 3 end of the intron is spliced followed by Hgation of two exon units (Fig. 4.8).

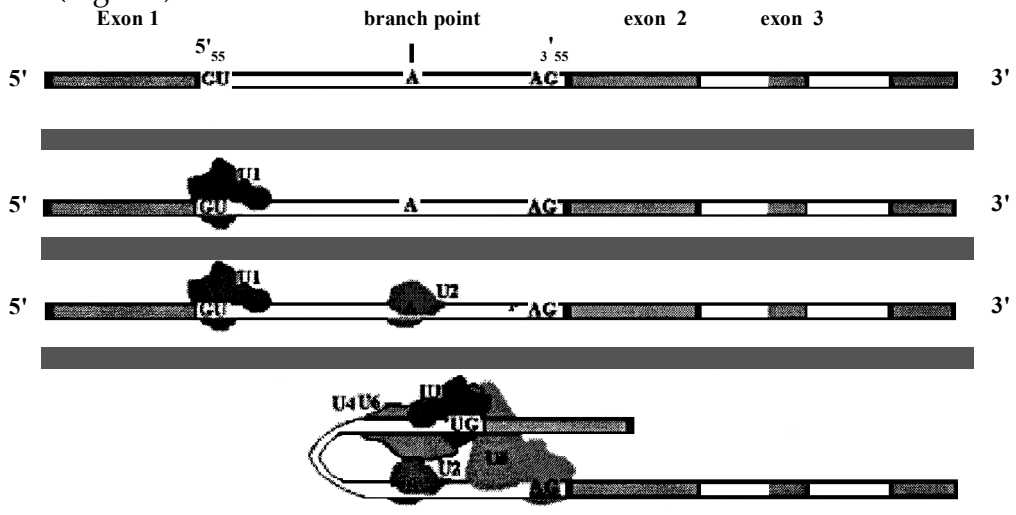


Fig. 4.8: Schematic diagram of splicing mechanism

The entire process is defined by three critical elements in the mRNA:

a. Sequences at the 5′ splice site of the intron (GU)
b. Sequence at the 3′ splice site of the intron (AG)
c. Sequences within the intron at the branch point where the 5' end joins.

4.6.2 Splicesosome

Spliesome are protein and snRNA complexes.

The RNAs are snRNAs called U_1 U_3 U_4 U_5 U_6 Their size varies from 50-200 nucleotides.

U_1 U_2 and U. along with specific proteins exist as independent units.

U_4 and U_6h along with their proteins are grouped together as a single unit.

U, snRNP first recognizes the 5′ splice site sequence and bind to it by complementary base pairing

U_2 then binds at the branch site with 'A' within the intron

A preformed, complex consisting of U_4U_6 and U_5 snRNPs is then incorporated into the forming splisosome

The U_5 component binds with both the 5′ and 3′ splice sites.

The snRNPs then catalyses the reaction by first cleaving the 5 splice site, then joining the 5′ terminal nucleotide with a specific adenine residue within the intron followed by cleavage at 3′ splice site and joining of the exons.

The snRNAs in the snRNPs actually catalyzes the reaction.

Note: RNAs are capable of self splicing- i.e. remove their own introns. Eg. 28S rRNA in *Tetrahymena.* Self-splicing RNAs are also present in mitochondria, chloroplast and bacteria. On the basis of the catalytic activity of self-splicing RNAs, they have been grouped into two classes.

Class I: In this type, rRNA first cleaves itself at the 5′ end of the intron. The 3′ end of the exon then catalyses the reaction at the 3′ end of the intron followed by joining of the two exons.

Class II: In this type, self-splicing rRNAs exhibit characteristics of the reaction as observed with mRNA described above.

Alternative splicing: Most pre-mRNAs have multiple introns and exones which can be arranged in alternative ways by splicing of the same mRNA can produce different mRNA- a novel means of controlling gene expression. This process is known as **alternative splicing** and occurs frequently in genes of eukaryotes that provide an important mechanism for tissue-specific and developmental regulation of gene expression. The regulation and selection of splice sites is done by Serine/Argi nine-residue proteins, or **SR proteins** (Fig. 4.9). The use of alternative splicing factors leads to a modification of the definition of a "gene".

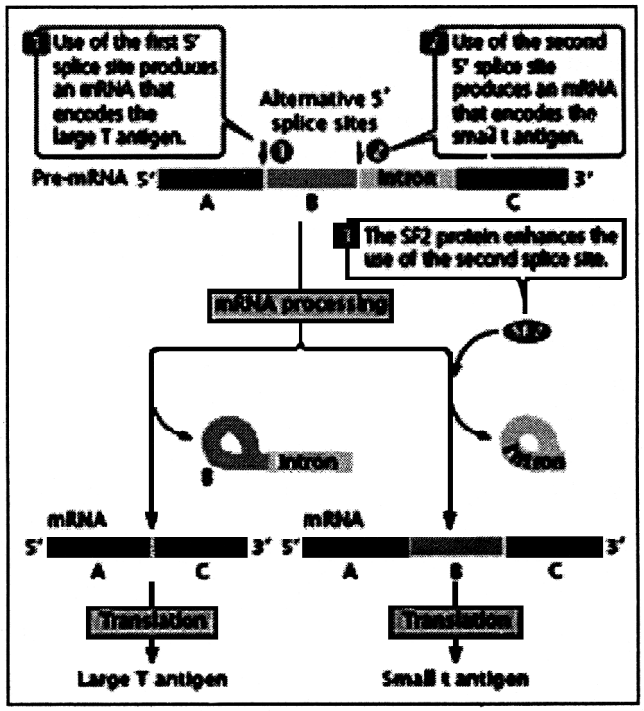


Fig. 4.9: SR Proteins

There are four known modes of alternative splicing:

Alternative selection of promoters: this is the only method of splicing which can produce an alternative N-terminus domain in proteins. In this case, different sets of promoters can be spliced with certain sets of other exons.

Alternative selection of cleavage/polyadenylation sites: this is the only method of splicing which can produce an alternative C-terminus domain in proteins. In this case, different sets of polyadenylation sites can be spliced with the other exons.

Intron retaining mode: in this case, instead of splicing out an intron, the intron is retained in the mRNA transcript. However, the intron must be properly encoding for amino acids. The intron's code must be properly expressible, otherwise a stop codon or a shift in the reading frame will cause the protein to be non-functional.

Exon cassette mode: in this case, certain exons are spliced out to alter the sequence of amino acids in the expressed protein.

E.g. 1. Transcriptional activators consist of two distinct domains: a DNA binding domain and an activation domain. These domains are generally encoded

in separate exons, so alternative splicing allows them to be reassorted into different combinations, thereby enabling the production of activators and repressors from the same gene.

2. In *Drosophila,* alternative splicing of the same pre-mRNA determines whether a fly will be a male or female.

Patterns of alternative splicing can vary in different tissues. Several protein factors have been isolated but the mechanism by which the correct splice sites are selected in pre-mRNA is not known. Variations in the expression of such splicing factors in different cell types may result in tissue specific patterns of alternative splicing, there by contributing to the regulation of gene expression during development and differentiation.

4.6.3 Significance of alternative splicing

Alternative splicing is of great importance to genetics - it invalidates the old theory of one DNA sequence coding for one polypeptide. External information provide the clue of alternative splicing. Since the methods of regulation are inherited, the interpretation of a mutation may be changed.

Alternative splicing allows more information to be stored much more economically in a limited space. It has been noted that it is unnecessary to change the DNA of a gene for the evolution of a new protein. Instead, a new way of regulation could lead to the same effect, but leaving the code for the established proteins unharmed. Another speculation is that new proteins could be allowed to evolve much faster than in prokaryotes. This machanism may allow for a higher probability for a functional new protein. Therefore the adaptation to new environments can be much faster - with fewer generations - than in prokaryotes. This might have been one very important step for multicellular organisms with a longer life cycle.

Trans-splicing :

In this type, exons from two different pre-mRNAs are joined to form a single mRNA. For example, in trypanosomes, all mRNAs have an identical spliced leader sequence of 35 nucleotides. This leader sequence is present the 5' end of a 137 nucleotide RNA chain. This leader sequence is then spliced and added to all the 5' end of mRNAs by trans-splicing reactions. Trans-splicing machinery also exists in mammals as mammalian cells are capable of carrying out trans-splicing reactions with the nematode spliced leader RNAs (Fig. 4.10).

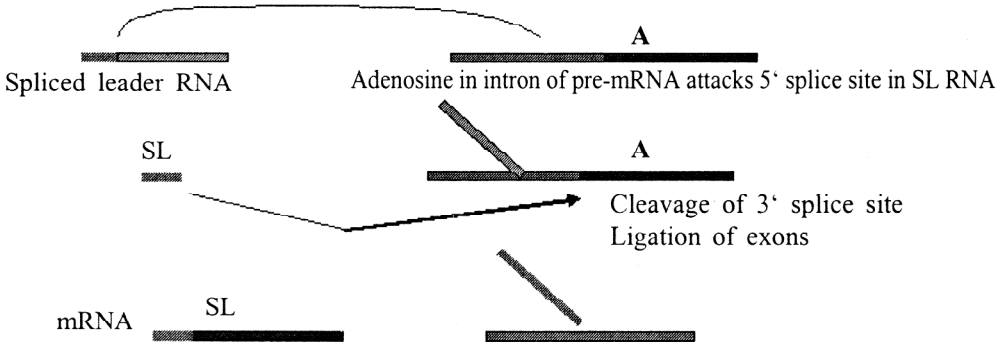


Fig. 4.10: Trans-splicing reactions with the nematode spliced leader RNAs.

RNA Editing

RNA editing refers to RNA processing events (other than splicing events) that alter the protein coding sequences of some mRNAs. In trypanosomes and related protozoans, addition of'U' and deletions of some nucleotides occur in some mRNAs. The information required for editing is encoded in "guide" RNAs, which are complementary to edited portions of the mature mRNA. The guide RNA contains poly-U tail, which donate the 'Us' during editing. Sometimes the editing is so extensive that half the nucleotide sequences are altered.

In mammalian cells similar events occur in mitochondria and also in the somatic cells. For example, in human body two types of apolipoproteins are found, Aop-BlOO (4536 aa: synthesized in liver is unedited form and transports lipids in the circulation) & Apo-B48 (2152 aa: synthesized in the intestine, is edited form and helps in the absorption of dietary lipids in the intestine). In the intestine, CAA codon is changed to a stop codon by enzymatic conversion of C to U by removal of the cytosine amino group at specific site in the mRNA.

4.7 Nuclear export of mRNA

4.7.1 Introduction

Compartmentahzation of the eukaryotic genome by the nuclear membrane was probably a necessity to have a greater control over the functioning of the genome and also to avoid unnecessary alterations in the genomic constituents by exposing it to the bustling biochemical activity that occur in the cytoplasm. This compartmentahzation further ensures the presence of specialized environments for different stages of gene expression, such as transcription and protein production.

Trafficking of materials between the nucleus and cytoplasm primarily rely

on transport receptors in the **importin-**β superfamily. However, export of mRNA uti- I lizes distinct soluble machinery. In yeast, it has been observed that protein-protein interaction is required for the export of mRNA from the nucleus to the cytoplasm. For example, in yeast Mex67p interact with Mtr2p and facilitates the export of I po!y(A)$^+$ RNA. In metazoans and in humans, TAP was confirmed to be the orthologue of Mex67p, redesignated as **NXFl** (nuclear export factor 1) which interacts with pl5/ NXT1 in the nucleus for transportation of mRNA. Mtr2p and pl5 share no sequence similarity but the Mex67p-Mtr2p complex displays similar structural architecture to I the NXFl-p 15 heterodimer. However, the distinction between mRNA export and the importin-β family/Ran network is not absolute, as an importin-13 family member has recently been implicated in mRNA export as well.

4.7.2 mRNA export is coupled to splicing

NXFl does not bind directly to cellular mRNA. Experiments with *Xenopus* oocytes demonstrated that the process of splicing can contribute to the efficiency of mRNA export; the spliced product from adenovirus major late (Adml) mRNA was shown to export more efficiently than an identical mRNA engineered to lack an intron. On the mRNAs, exon junction complexes (EJC) are formed after splicing. EJC complex include several components like REF (nuclear export factor), SRml60, RNPS1, DEK, Y14, and later its protein partner Magoh. Recruitment of a unique set of proteins to the spliced mRNA may promote export competency of mRNA.

The notion that EJC deposition leads to recruitment of NXFl is an attractive model to explain the stimulatory effect of splicing on export. Direct interactions between REF and NXFl have been observed in both human and yeast systems. REF also shuttles between nucleus and cytoplasm and enhances mRNA export when injected into *Xenopus* oocyte nuclei as a recombinant protein. The enhanced placement of REF onto mRNA in a splicing-dependent fashion, as well as its association with NXFl, made REF a prime candidate for recruiting NXF1 onto mRNA.

A connection between splicing and mRNA export was further solidified with the characterization of a novel role for the putative RNA helicase UAP56 [56-kDa U2AF(65)-associated protein]. Recruitment of REF to spliced mRNA is dependent upon its interaction with UAP56. From these data, a very simple yet elegant mode of coupling splicing with mRNA export became evident. Namely, REF is recruited to splice mRNA through direct interactions with UAP56, and consequently, REF (and the EJC in general) recruits the export factor NXFl to promote exit from the nucleus by mediating docking and presumably movement through the pore.

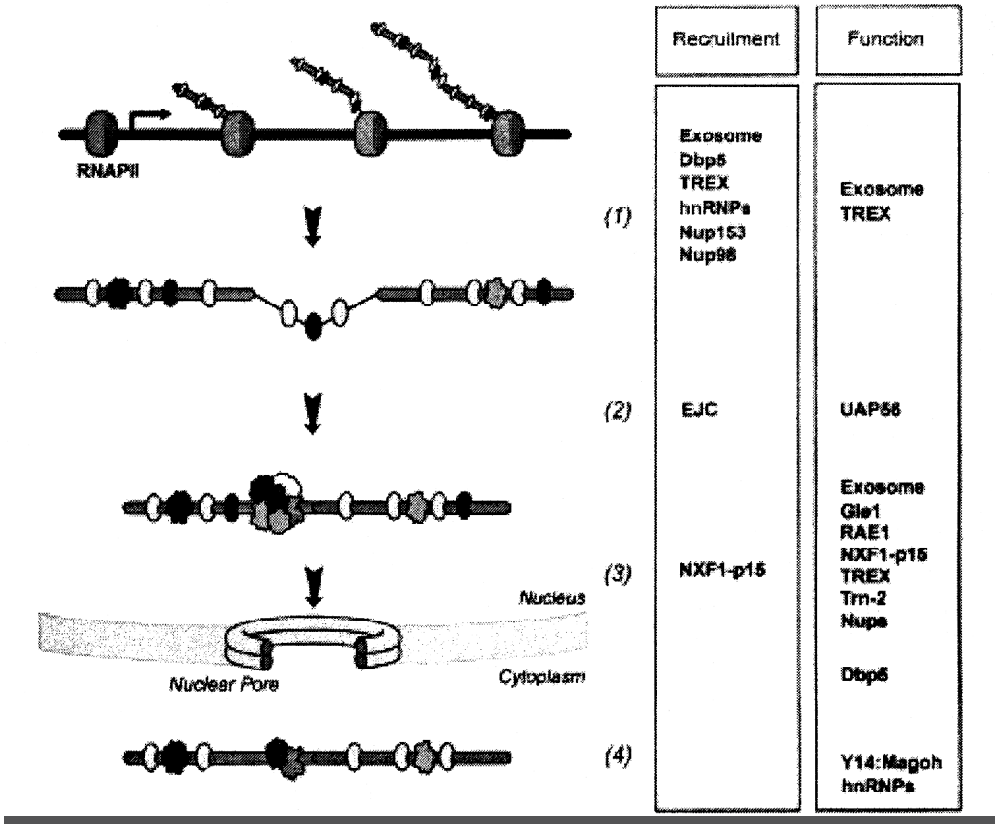


Fig. 4.11: A schematic illustration of mRNA biogenesis is depicted, with the proposed times of recruitment and functions for specific proteins indicated in the boxes.

(1) Transcription: Much evidence points to cotranscriptional loading of factors involved in RNA processing, export, and quality control to the nascent transcript. The mobile pore proteins, Nupl53 and Nup98, are candidates (indicated in red text) for loading onto mRNA early in its biogenesis, although this is yet to be demonstrated.

(2) Splicing. The splicing factor UAP56 interacts with REF, a component of EJC that is deposited on mRNA during splicing. Loading of certain transport factors, such as REF, can also occur independent of splicing as a part of TREX or if the RNA is sufficiently long.

(3) Remodeling and export. NXFl-pl5 is recruited to the mRNA via protein-protein interactions, readying export-competent mRNA for mobilization out of the nucleus. At this step, other proteins, such as Glel, RAE1, Trn-2 (transportin-2),

and TREX components, are also thought to function. Certain hnRNPs and EJC components are shed from the mRNP prior to export, and proper mRNP formation appears to be monitored at this stage by the exosome. Specific pore proteins, or Nups, are implicated in moving mRNA cargo through the pore (Fig. 4.12). Although loaded onto the transcript early in biogenesis, Dbp5 may play a late role in remodeling and/ or moving the mRNP complex. (4) Cytoplasmic function. Factors remaining on the mRNA, such as Y14 and Magoh, influence translation and localization of mature mRNPs.

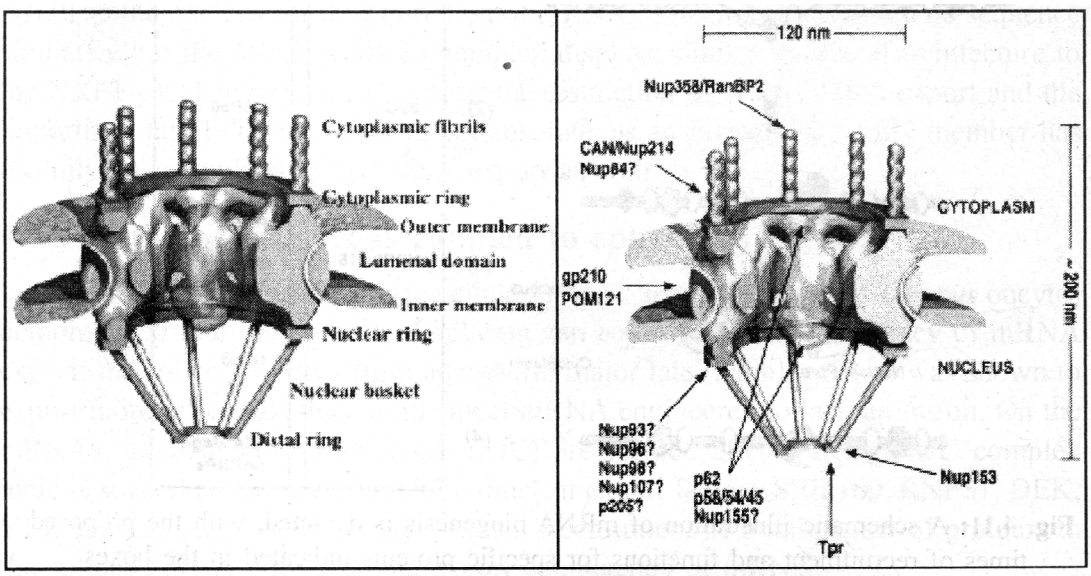


Fig. 4.12. Distribution and dynamics of pore proteins and associated factors implicated in mRNA export. Nupl53 and Nup98 are both dynamically associated with the nuclear pore in a manner dependent on ongoing transcription. However, in the case of Nupl 53, there is also evidence for a stable population, which is schematically illustrated here proximal to the inner nuclear membrane. The presence of distinct populations of"Nupl 53 is consistent with epitope exposure of this protein: different regions are exposed at the distal and proximal ends of the pore basket. Regardless of exactly how Nupl53 is arranged at the pore basket, there is evidence that the C-terminal region of this pore protein can extend into the cytoplasm, although Nupl53 does not appear to be released from this face of the pore. In contrast, Nup98 exists in equilibrium with a cytoplasmic pool and is known to interact with components of both sides of the pore. RAE1/Gle2 is a partner protein of Nup98. Both Nupl53 and Nup98 associate with the Nupl07-160 complex, a stable component of the pore. Tpr is also a component of the nuclear pore basket and relies on interaction with Nupl53 for correct localization. CAN/Nup214 is localized to fibrils extending from the cytoplasmic ring of the pore and is a docking site for the dynamic DEAD box helicase, Dbp5.

Despite the observation that splicing promotes export of mRNA, such processing is not the only or even the major route for export factor recruitment. Recent functional analysis of mammalian cells also suggested that splicing does not always have a major effect on mRNA export per se. Splicing and export factor recruitment that has been documented may represent only one way that export factors load onto mRNA and indeed may make a significant contribution at this step only when the RNA is particularly short

4.7.3 Coordinating mRNA export with transcription and turnover

The addition of a 5' cap, splicing, polyadenylation, and cleavage events occur in close connection to transcription. Concurrent with the processing events, mRNAs are also packaged with a number of proteins specific to this class of RNA (Fig. 1). A significant subset of such proteins was originally classified as hnRNPsfor their ability to associate with heterogeneous nuclear RNA. hnRNP Al a component of hnRNPs was initially implicated in mRNA transport, but it's exact role in transport mechanism remains to be elucidated.

Other proteins, not necessarily classified originally in the hnRNP category but loaded onto mRNA, have been functionally connected to both export and transcription. Yralp and Sub2p display both genetic and physical interactions with all members of the yeast THO complex, a protein complex identified originally for a role in transcription elongation. REF and UAP56 along with the vertebrate counterparts of THO and a new protein of unknown function, Texl, make up the TREX (transcription and export) complex. In yeast, specific TREX components associate with genes during transcription and, individually, their deletion results in nuclear poly(A)$^+$ accumulation. Together, this suggests that TREX proteins may be important in mediating cotranscriptional recruitment of factors important in export. For example, one protein in the TREX complex, Hprlp, is required for efficient targeting of Yralp and Sub2p to genes undergoing active transcription. Therefore, cotranscriptional recruitment and splicing-dependent recruitment represent two broad mechanisms by which mRNA export factors can associate with RNA cargo. For instance, REF maybe efficiently loaded onto specific RNA cargos via cotranscriptional targeting of the TREX complex and/or through splicing-dependent deposition of the EJC.

Another example of the connection between transcription and export is found in the DEAD box helicase Dbp5. Dbp5 is localized at steady state to the cytoplasmic fibrils of the nuclear pore complex (NPC) and has been hypothesized to be involved in a terminal step of mRNA release from the NPC, possibly acting in a remodeling

step to unwind mRNPs entering the cytoplasm. Interaction of Dbp5p with TFIIH during mRNA transcription and its shuttling between nucleoplasm and cytoplasm in S. cerevisiae is an indication that Dbp5 may load onto mRNA cargo very early in biogenesis that enables transport and remodeling of mRNA. Overall, much evidence is arising to support a link between mRNA synthesis and the effective recruitment of export factors to the nascent transcript.

4.7.4 Moving on to the nuclear pore

Nuclear protein complexes (NPC) span the nuclear envelope and serve as gateways of communication individual nuclear pore proteins or nucleoporins (Nups) present several times, creating octagonal symmetry. The pore also has asymmetric features on its nuclear and cytoplasmic faces. Although much of the process of mRNA export is being deciphered, there is still little known about how mRNPs interface with pore machinery. Some recent studies have focused on the roles of proteins that are closely associated with the pore, such as Glel and RAE1/ Gle2. Glel is essential for mRNA export in both yeast and human cells, and hGlel is a dynamic factor that shuttles between nucleus and cytoplasm. The shuttling domain of hGlel acts as a dominant-negative export inhibitor of both buIkpoly(A)$^+$ RNA and specific mRNA transcripts. Docking of hGlel at the NPC was recently shown to depend on an interaction with the pore protein Nupl55.

Murine RAE1 is essential; however, cells from mice bearing targeted disruption of RAE1 do not have a detectable defect on bulk mRNA export. In contrast, RAE1 deletion in yeast results in nuclear accumulation of po!y(A)$^+$ RNA. Although there appear to be redundant factors in vertebrates, hRAEl interacts with NXF1 and the nucleoporin Nup98, as well as with mRNA itself, and has been speculated to be involved in delivering mRNA cargo-receptor complexes to Nup98. Nup98, in turn, has been implicated through antibody inhibition studies in the export of mRNA as well as other classes of RNA. Nup98 shares similarity with yeast nuclear pore proteins Nupl45, Nupll6, and NuplOO. Deletion of yNup 145 causes the nuclear accumulation of poly(A)$^+$ RNA.

Vertebrate pore proteins have not been exhaustively screened and individually tested for roles in mRNA export. However, along with Nup98, five other vertebrate pore proteins, Nupl53, Nupl60, Nupl33, Tpr, and CAN/Nup214, have so far been implicated in the export of mRNA (Fig. 4.13). Mouse embryos deficient in CAN/ Nup214not only show arrest in the G_2 phase of the cell cycle but also demonstrate nuclear accumulation of poly(A)$^+$ RNA. Nupl59/Rat7,the

yeast orthologue of CAN/ Nup214, is similarly implicated in mRNA export, with a temperature-sensitive mutation causing very rapid onset of accumulation of poly(A)$^+$ RNA in the nucleus. CAN/Nup214 associates with the mRNA export factor Dbp5, an interaction conserved from yeast to vertebrates. In addition, CAN/ Nup214 is the only vertebrate nucleoporin with a steady-state localization exclusively on the cytoplasmic side of the pore that has been implicated in mRNA export thus far. Although much remains to be elucidated,Nup98 and Nupl53 are prime candidates for coupling the production of mRNA to its transport into the cytoplasm.

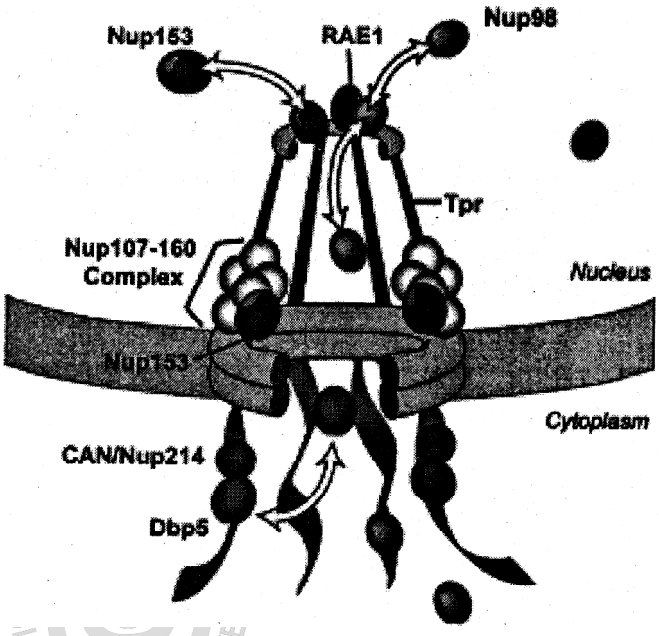


Fig. 4.13 : Different Pore Proteins

4.7.5 Transport through to pore : putting individual components into context

In recent years, several models to explain the mechanism of movement through the pore have been proposed. In each, the phenyalanine-glycine repeat motif (FG repeat) regions found in several nucleoporins play a prominent role, both in contributing to an exclusion barrier as well as in serving as binding sites for cargo-receptor complexes. Consistent with this, NXF1 directly interacts with the FG repeat domains of several nucleoporins in vitro. It is presumed that during translocation of mRNAs, associated large heterogenous mRNPs undergo remodeling in conjunction with transport. Elegant immunoelectron microscopy studies have gone on to illustrate that certain proteins are shed from the mRNP, while others accompany the RNA through the pore. Complicating things further, the nuclearpore basket itself has been observed to adopt different conformations when the Balbiani ring mRNA is traversing the pore. The dynamic nature of specific pore basket components themselves (Nup98 and Nupl53),as well as the sensitivity of such mobility to the transcriptional status of the cell, suggests that basket

remodeling may normally be ongoing in a manner linked to RNA trafficking. Recent information suggests that the native NPCs distal ring of the pore basket is not an open hole but rather a dense structure and thus has little scope of remodeling. An alternative point of entry into the pore, in between the fibers of the pore basket, has also been proposed. Much work is still needed to understand how mRNA enters and translocates through the pore. Future approaches that provide high-resolution real-time imaging as well as more precise functional assays are sure to yield a very interesting story about how the complicated network of mRNA biogenesis connects with translocation through the nuclear pore and the downstream fate of the mRNA.

4.7.6 Exportins and importins

The traffic through the nuclear envelope is mediated by a protein family which can be divided into **exportins** and **importins.** Binding of a molecule (a "cargo") to exportins facilitates its export to the cytoplasm. Importins facilitate import into the nucleus.

The function of exportins and importins is regulated by a G protein called **"Ran".** There are two types of G proteins: heterotrimeric G proteins and monomeric G proteins (or small G proteins). The latter includes Ras, Ran, Rho, Rab, etc. Like other G proteins, Ran can switch between GTP-bound and GDP-bound states. Transition from the GTP-bound to the GDP-bound state is catalyzed by a **GTPase-activating protein (GAP)** which induces hydrolysis of the bound GTR The reverse transition is catalyzed by **guanine nucleotide exchange factor (GEF)** which induces exchange between the bound GDP and the cellular GTP.

The GEF of Ran (denoted by RanGEF) is located predominantly in the nucleus while RanGAP is located almost exclusively in the cytoplasm. Therefore, **in the nucleus Ran will be mainly in the GTP-bound state** due to the action of RanGEF while **cytoplasmic Ran will be mainly loaded with GDP.** This asymmetric distribution has led to the following model for the function of exportins and importins.

It is thought that binding between an exportin or importin and its cargo depends on their interaction with Ran: **RanGTP enhances binding between an exportin and its cargo but stimulates release of importing cargo; RanGDT has the opposite effect,** namely, it stimulates the release of exportin's cargo, but enhances the binding between an importin and its cargo. Therefore, the exportin and its cargo may move together with RanGTP inside the nucleus, but the cargo will be released as soon as the complex moves into the cytoplasm (through nuclear pores), since RanGTP will be converted to RanGDP in the cytoplasm. By contrast,

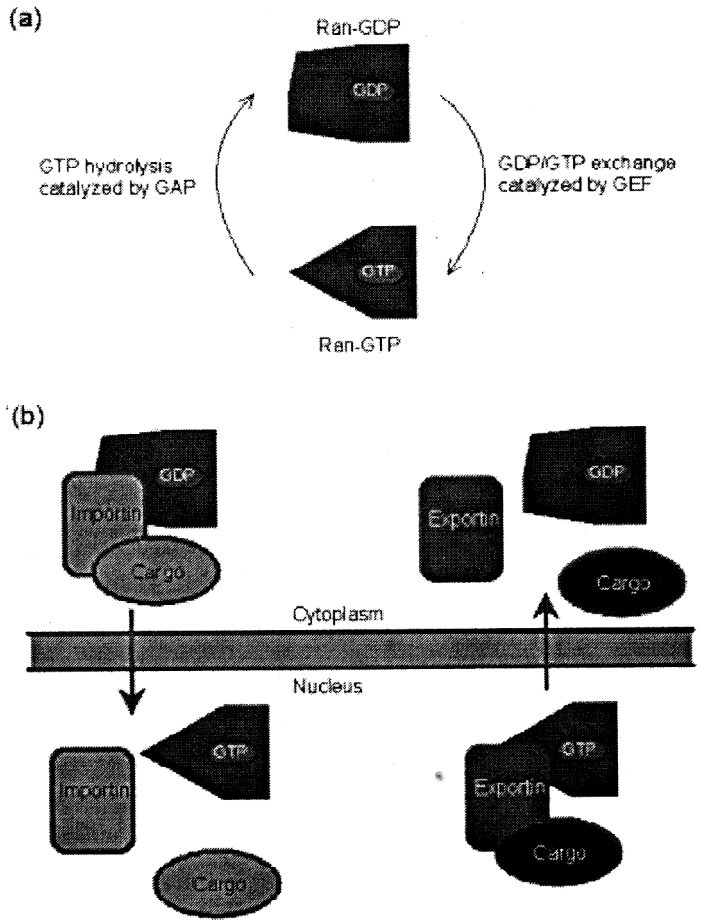


Fig. 4.14: Ran, importin and exportin. **(a)** The two states of Ran: GTP-bound and GDP-bound. **(b)** General function of importins and exportins

the importin and its cargo may move together with RanGDP in the cytoplasm, but the cargo will be released in the nucleus since RanGDP will be converted to RanGTP in the nucleus.

Unit 5 Translation

Structure

5.1 Genetic code

5.1.1 Introduction

A code is a system of symbols that equates information in one language with information in another e.g. Morse code. In living organisms, the hereditary information is written in the language of four nucleic acids, A, G, C, and TAJ. and the language of proteins are written in amino acids. As there are 20 amino acids that are genetically encoded by DNA or RNA sequences, the first question that comes in mind is that how many nucleotides are necessary to code for one amino acid. *Now* ***we know that genetic codes are triplet codons, each of which represents a single amino acid.***

Initially, the number of nucleotides necessary to code for one amino acid was derived by reasoning and later confirmed by experimentation. Scientists reasoned if one nucleotide represented one amino acid, only four nucleotides could be coded. If two nucleotides represented each amino acid, there would be 4^2 =16 possible combination of couplets, still not sufficient to code for 20 amino acids. If the code consisted of groups containing one and two nucleotides, it would have 4 + 16-20, just sufficient for all the 20 amino acids but would fail to recognize the pause signal between two genes. Groups of three nucleotides in a row would provide 4^3 = 64 different triplet combinations, more than enough to code for **all** the amino acids. Theoretically, the above logic appeared to be simple but had to await experimental evidence that proved beyond doubt that groups of three nucleotides indeed are necessary to code for a single amino acid. Each triplet nucleotides in genetics are referred to as **codon.**

5.1.2 Experimental evidences

Crick and his coworkers tested acridine induced mutations in the B cistron of rll locus of T4 phage (acridine mutation is produced by addition or deletion of a

nucle-otide). The mutations acted on the principle of addition or deletion of a single nucleotide pair in DNA molecule. Arbitrarily the mutations were designed as + or -on the basis of their suppression effect on other mutations.

For eg. :

In DNA : TAC TCC CGA ACG ATA CCA GAG

In RNA : AUG AGG GCU UGC UAU GGU CUC

Protein : 

Mutations induced by acridine treatment (arbitrarily designated as 4- mutations)

Point mutation

↓

In DNA : TAC GCC CGA ACG ATA CCA GAG (point mutation; a base is replced)

In RNA : AUG CGG GCU UGC UAU GGU CUC [NO FRAME SHIFT MUTATION]

Protein : 

Second Mutations that suppress the first imitation (designated as - mutations)

Insertion Deletion

↓ ↓ ↓

In DNA : TAC GCC CCGA CGA TAC CAG AG

In RNA : AUG CGG GGU GCU AUG GUC UC

Protein :  FRAME SHIFT MUTATION

Three + Mutations can restore the oriainal mutation

Insertion **[Partial restoration of original]**

In DNA : TAC GCC CCA TCG ATA CCA GAG CCA [three positive mutation restores the

In RNA : AUG CGG CGU ACU UAU GGU CUC GGU original FRAME SHIFT mutation]

Protein : 

Observations:

1. A single + or - mutation is sufficient to alter the wild type trait into mutant
2. The effect of a single + mutation is suppressed either by a - mutation or by 2 further + mutations. Similarly - mutatios can by suppressed by either by a + mutation or by 2 further - mutations.

Explanations & Deductions:

1. Due to deletion or insertion of a nucleotide, codon constitution of reading frame gets altered after the point of such change. For this, the polypeptide product will be different resulting in a mutant trait.
2. If it is insertion mutation, two further insertion mutations can restore the original trait. Similarly, if the mutation is caused by deletion, two further deletion mutations can restore the original trait.

So, three changes of the similar kind or multiples of 3 are necessary to restore to wild type. This can only be possible if the codons are triplet i.e. consists of 3 nucleotide.

5.1.3 Nucleotide sequence is collinear with a polypeptide's amino acid sequence

The nucleotides in RNA or DNA are arranged in a linear order. As DNA codes for proteins, it was assumed that the amino acids in the polypeptide chain must also be arranged in a linear fashion. Although the proteins have a highly complex three dimensional structure under normal circumstances, analysis has revealed that the amino acids in any polypeptide chain are arranged one after another, have definite polarities and show no branching. Thus the linear arrangements of nucleotides and the amino acids led to suggested that there must be one to one corresponding relationship of nucleotides and amino acids during protein synthesis.

5.1.4 Overlapping vs. non-overlapping nature of codons

Logical derivation of 3 nucleotides constituting a codon however could not provide clue as to how the codon are arranged i.e. either overlapping or non-overlapping. Point mutations with mutagens like nitrous acids or provalin were used to decipher the arrangement of codons.

a) If Overlapping, a sequence of 9 nucleotides will code for 7 amino acids as shown below. Alteration of a single nucleotide will produce changes in minimum three codons.

b) If non overlapping, 9 nucleotide will code for only three amino acids and alteration of a single nucleotide will alter only one codon.

Original sequence	***Mutant sequence***
CAGAGCUCA	CAGAACUCA
CAG.....codon 1	CAG.....codon 1
AGA.....codon 2	AGA.....codon 2
GAG.....codon 3	GAG.....codon 3
AGC.....codon 4	AGC.....codon 4
GCU.....codon 5	GCU.....codon 5
CUC.....codon 6	CUC.....codon 6
UCA...codon 7	UCA...codon 7

b) If non overlapping, 9 nucleotide will code for only three amino acids and alteration of a single nucleotide will alter only one codon.

Original sequence			***Mutant sequence***		
CAGAGCUCA			CAGAACUCA		
CAG	AGC	UCA	CAG	AAC	UCA
codon 1	codon 2	codon 3	codon 1	codon 2	codon 3

Experiments revealed that a single-nucleotide substitution mutation caused an amino acid substitution of leucine for proline. The two adjacent amino acids were unchanged by the mutation providing evidence against overlapping codons (Tsugita and Fraenkel, 1960).

	U	C	A	G	
U	UUU, UUC: Phe; UUA, UUG: Leu	UCU, UCC, UCA, UCG: Ser	UAU, UAC: Tyr; UAA, UAG: Stop	UGU, UGC: Cys; UGA: Stop; UGG: Trp	U C A G
C	CUU, CUC, CUA, CUG: Leu	CCU, CCC, CCA, CCG: Pro	CAU, CAC: His; CAA, CAG: Gln	CGU, CGC, CGA, CGG: Arg	U C A G
A	AUU, AUC, AUA: Ile; AUG: Met	ACU, ACC, ACA, ACG: Thr	AAU, AAC: Asn; AAA, AAG: Lys	AGU, AGC: Ser; AGA, AGG: Arg	U C A G
G	GUU, GUC, GUA, GUG: Val	GCU, GCC, GCA, GCG: Ala	GAU, GAC: Asp; GAA, GAG: Glu	GGU, GGC, GGA, GGG: Gly	U C A G

5.1.5 Properties of genetic code

1. Each Genetic Code is a **Triplet Codon** each of which specifies an amino acid.
2. The code is **non-overlapping.** E.g. in mRNA 9 nucleotide sequence 5 GAAGUUGAA3 , will be translated to 3 amino acid sequence corresponding to GAA, GUU and GAA and no more.
3. The **code is degenerate,** which means that in many cases more than one codon specifies the same amino acids.
4. The **code is coma less.** A coma less code means that no punctuations are needed between any two words i.e. after one amino acid is coded; the second amino acid will be automatically coded by the next three nucleotides.
5. The code includes **three stop, or nonsense codons:** UAA, UAG, and UGA. These three codons do not code for any amino acids, rather they terminate translation.
6. The code is **non-ambigious.** Non-ambiguous code means that there is no ambiguity about a particular codon. A particular codon will always code for the same amino acid wherever it is found. *Only exception lies with AUG and GUG at the start point where both code for methionine although GUG actually code for valine.*
7. The **code is universal.** Almost all micro and macro organisms use the same genetic code with a few exceptions. For example, a different code exists in mitochondria of some eukaryotes, so that in cytoplasm and mitochondria same codon may code for different amino acids.
8. During translation, the code is read from 5′ to 3′ direction. Moving from the 5 to the 3 end of an mRNA, each successive codon is sequentially interpreted into an amino acid (N-terminus to C-terminus of a polypeptide).
9. There exists a fixed reading frame for any gene that includes the initiation codon. The initiation codon specifies the first amino acid to be translated, which is usually AUG that codes for methionine. Mutations may modify the message encoded in sequence in three ways:
    a) **Frameshift mutations** where nucleotide insertions or deletions alter the genetic instruction for polypeptide by changing the reading frame.
    b) **Missence mutations** change a codon for one amino acid to a codon for a different amino acid.
    c) **Nonsense mutations** change a codon for an amino acid to stop codon.

Deciphering the Genetic Code

Nirenberg and Matthaei developed the technique in the laboratory of Khorana in 1961 to decipher the genetic code. They artificially synthesized mRNA having only 'Uracil' as the component. Next, they used this mRNA to synthesize polypeptide in cell free extracts of E. coli. When poly U was used, polypeptides consisting entirely of phenylalanine were produced indicating that UUU must code for phenylalanine. Similarly poly 'C RNAs produced polypeptides made entirely of proline, meaning that CCC must code for proline. This method was used to decipher the code of almost all amino acids.

Exceptions:

- In few single celled eukaryotic protozoans known as ciliates, UAA and UAG, which are nonsense codons in most organisms, specify the amino acid glutamine.

- In mitochondria of yeast, CUA specifies threonine instead of leucine.

mRNA codon	virus/pro & Eukaryotes	Mitochondria		
		Yeast	Drosophila	Mammal
AUA	Isoluecine	Methionine	Methionine	Methionine
AGA, AGG	Arginine	Arginine	Serine	Stop
CUA	Leucine	Threonine	Leucine	Leucine
UGA	Stop	Tryptophan	Tryptophan	Tryptophan

5.2 The translation machinery

5.2.1 tRNAs

About 15% of total RNA in a cell is tRNA. tRNAs plays a significant role by serving as adopter molecules that recognize the right enzyme activated amino acid and the anticodon on mRNA. Robert Holly (1965) first elucidated the base sequence of alanine tRNA from yeast. Later, more that 100 tRNA was identified and se~ quenced. All known tRNA share common structural features probably because tRNA molecules must be able to interact in nearly the same way with ribosomes, mRNA and elongation factors. Common features are as follows:

1. All tRNAs are single chains containing 73 to 93 ribonucleotide (~ 25kd)

2. tRNAs possess some unusual bases like inosine, pseudouridine, dihydrouridine, ribothymidine and methylated or derivatives of AUCG. (Methylation prevents the formation of base pairing, rendering them inaccessible for pairing with other base pairs or other type of interaction and also imparts hydrophobic character, important for interactions with synthetase and ribosomal proteins and for folding).

3. The 5′ end of tRNAs is phosphorylated and usually p-Guanine.

4. The base sequence at 3′ end of mature tRNA is always -CCA. Activated amino acid binds to the 3′ -OH group of the terminal adenosine.

5. About half the nucleotides in tRNAs are base paired to form double helix. Five groups are not base paired:

   a) 3′ CCA terminal

   b) T→ C loop (ribothymine-pseudouracil-cytosine)

   c) Extra arm (may have variable no. of residues) (present in only class II tRNA= serene and leucine

   d) Anticondon loop

   e) DHU loop (contains several dihydrouracil)

6. The anticodon loop (Fig. 5.1) consists of seven bases, with the following sequence -5′ Pyrimidine- Pyrimidine-**X-Y-Z**-modified purine- variable base 3′

7. X-ray crystallography study of phenylalanine tRNA by Alexander and Aaron (1974) showed that tRNA is a L-shaped structure. There are two segment of double helix; each having about 10 bases pairs in each turn, in accordance with the cloverleaf mode. Bases in the non-helical region participate in unusual bondings (e.g. GG, AA, CC). Moreover, 2′-OH of the ribose phosphate backbone acts as hydrogen donor and interacts with each other. Most bases are stacked on one another and the hydrophobic interactions between the aromatic rings help to stabilize the architecture of the molecule. Subsequent analyzing of all tRNAs showed that they follow the same basic plan.

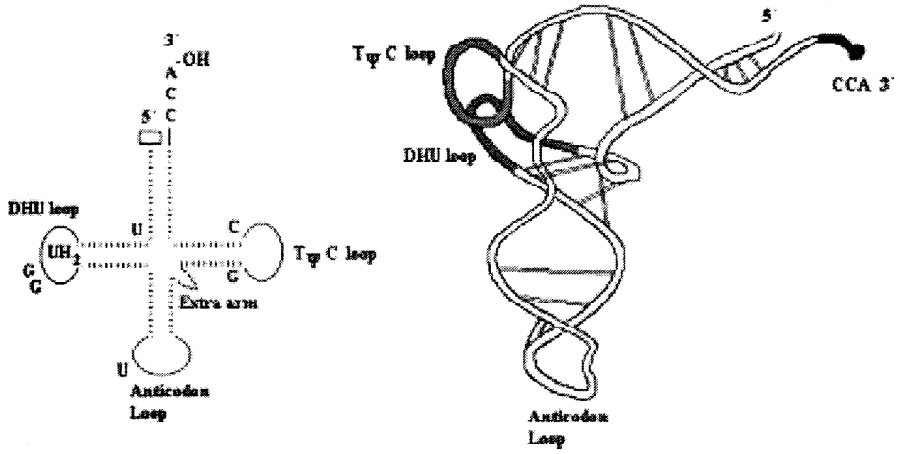


Fig. 5.1: The anticodon loop

In both prokaryotes and eukaryotes tRNA are synthesized as precursor molecules. Some pre t-RNA transcripts have several tRNA sequences, which are cleaved and modified to obtain mature functional tRNA. Some tRNA sequences are present within the pre-RNA transcripts.

5.2.2 Ribozyme

Ribozyme is an enzyme in which RNA component of a protein-RNA complex is responsible for the catalytic activity rather than the protein (Sidney Altman, 1983).

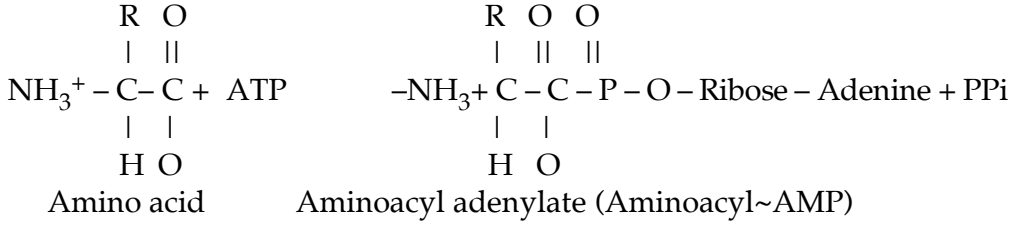
The 5′ end of tRNA is modified by RNase P (ribozyme). Conventional RNase modify the 3′ of end of tRNA which is the by followed by addition of -CCA nucleotides. Some tRNA has the information of-CCA already encoded in the DNA. All tRNAs have -CCA sequence at their 3' end.

In *E. coli,* there are 60 genes arranged in 25 units that synthesize tRNA molecules. One unit transcribes multimeric precursor for 7 tRNAs that are then cleaved by RNase P, RNase D to produce mature tRNA ($tRNA_{leu}$ 2 × $tRNA_{met,}$ 2 x $tRNA_{gtu,}$ 2 x tRNA $_{glu2}$)

5.2.3 Amino acid activation & linkage to specific tRNA by specific synthetase

Thermo dynamically, formation of peptide bond between -NH_2 of one amino acid and -COOH group of another amino acid is not favored. The barrier is overcome by activating the -COOH group of the precursor amino acid. Activation takes place by the following mechanism: -

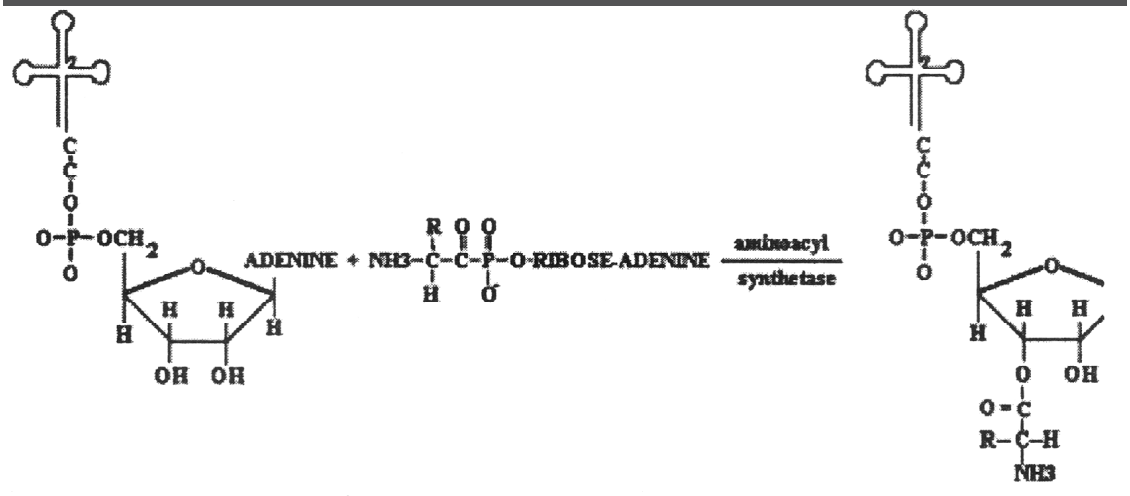
Step 1.



In the first step, an amino acid is linked to the phosphoryl group of AMP and therefore is known as aminoacyl~AMP

Step 2.

In the second step, aminoacyl group of aminoacyl-AMP is transferred to either 2′ or 3′ hydroxyl group of the ribose unit at the 3′ end of tRNA to form aminoacyl tRNA.



The attachment of amino acid to a tRNA is important not only because it activates the carboxyl group but also because amino acids by themselves cannot recognize the codons on mRNA. tRNA serves as an adopter molecule by recognizing the codons on one hand and by bringing specific amino acids, represented by the codons at the site of protein synthesis.

5.2.4 Aminoacyl tRNA synthetase

The amino acid is attached at the 3′ end of the tRNA to either the 2′ hydroxyl or the 3′ hydroxyl.

1. ***Class I*** *amino-acyl tRNA synthetases attach their associated amino acids to the tRNA* ***2' hydroxyl*** (NOTE: typically the hydrophobic amino acids)

2. ***Class II*** *amino-acyl tRNA synthetases attach their associated amino acids to the tRNA* ***3' hydroxyl*** (NOTE: typically hydrophilic amino acids)

For each amino acid, there exists a specific amonoacyl tRNA synthetase. These enzymes have been grouped into two classes on the basis of short signature sequences. In Class I there are 10 aminoacyl synthetases that recognize the larger amino acids and more hydrophobic enzymes. Class II (ancient type) aminoacyl tRNA synthetase recognizes the smaller amino acids. Class I enzymes acetylated the 2 -OH group and have a parallel ? domain (the classical dinucleotide binding fold) while class II enzymes acetylate the 3 -OH group (except phe) on the ribose and have an anti parallel ? domain (aminoacyl activating domain).

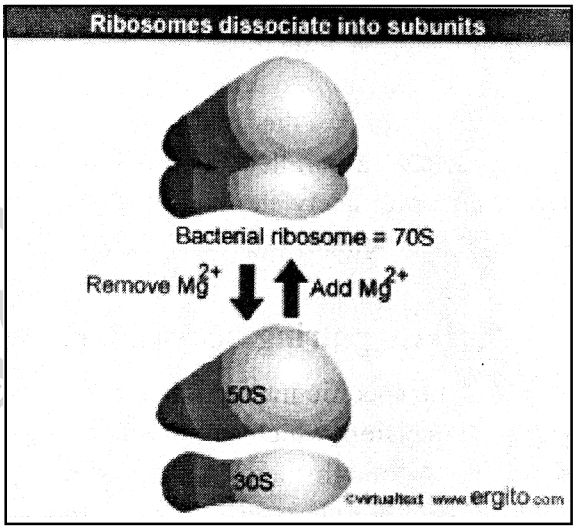
5.2.5 Base pairing between an mRNA codon and tRNA anticodon

It is the specific interaction between tRNA's anticodon and mRNA's codon that makes the decision, which amino acid will be incorporated into the growing polypep-tide chain. Although there is at least one kind of tRNA for each of the 20 amino acids, cells do not necessarily carry tRNAs with anticodons complementary to all of the 61 possible codon triplets in the degenerate genetic code. For e.g. in *E coli,* makes 79 different tRNAs containing 42 different anticodons. Obviously 19 (61-42=19) potential anticodons are not represented. Thus 19 mRNA codons will not find a complementary anticodon in E coli collection of tRNAs although such codons are present and are being coded into polypeptide chains. Therefore, there must be some tRNAs that recognize more than one codon for a particular amino acid. That is, the anticodons of these tRNAs can interact with more than one codon for the same amino acid. Although the exact codon-anticodon interaction of mRNA and tRNA is not very clear, Rrancis Crick, by analyzing the genetic code concluded that $3^{'}$ nucleotide in may codons adds nothing to the specificity of the codon. For example $5^{'}$ GGU$3^{'}$, 5′GGC3′, 5′GGA3′ and 5′GGG3' all encode glycine. It does not matter whether the anticodon on $tRNA^{gly}$ has a complementary base pare for the last codon at 3′ end provided the first two nucleotides are matched properly the tRNA will add glycine to the growing polypeptide chain. The same is true for other amino acids that are encoded by more than one codon. Thus the $5^{'}$ nucleotide

of tRNA's anticodon can often pair with more than one kind of nucleotide in the 3' position of an mRNA's codon. A tRNA charged with a particular amino acid can thus recognize several or even all of the codons for that amino acid. This flexibility in base pairing between the 3′ nucleotide in the codon and 5′ nucleotide in the anticodon is known as **wobble.**

5.2.6 Ribosome

Ribosomes are assembly of rRNA molecules and numerous proteins that synthesizes proteins under the direction from mRNA template. The bacterial system have ribosomes that sediment at 70S and the eukaryotic ribosomes sediment at 80S. Each ribosome can dissociate into two units: 50S and 30S in bacteria, while in eukaryotes it is 60S and 40S. The smaller subunit binds with mRNA to initiate translation.



Fi8- 5-2: Ribosomes dissociate into subunits

The actual number of rRNAs and number of proteins vary among species. For example, mammalian ribosomes have 4 rRNAs and 80 proteins while in *E. coli* the ribosomes have 3 rRNAs and 52 proteins. Although the number of proteins exceeds the number of rRNAs, the rRNAs constitute the major portion of the ribosomes and generally account for over 60% of the mass of the ribosome. The general structures of prokaryotic and eukaryotic ribosomes are more or less similar. Prior to translation, the ribosomes exist as two separate units- small subunit and larger subunit.

		rRNA	No. of proteins
Small subunit	Prokaryotes (30S)	16S	21 proteins
	Eukaryotes (40S)	18S	~30 proteins
Large subunit	Prokaryotes(50S)	23S and 5S	34 proteins
	Eukaryotes (60S)	28S, 5.8S, and 5S	~45 proteins

Each subunit has a specific three-dimensional shape that allows the two subunits to interlock with each other. The complete ribosome in prokaryotes and eukaryotes has a Svedberg value 70S and 80S respectively. There are two sites within the ribosome that can hold tRNAs: A site (aminoacyl or entry site) and the **P site** (peptidyl or donor site). The ribosome assembles on the mRNA with the A site oriented toward the 3′ end of the mRNA and the P site toward the 5′ end. At initiation, the three nucleotides of the initiation codon (AUG) align in the P site, where they pair with the anticodon in the initiator tRNA. This situates the A site over the second codon, which is now ready to receive the appropriate charged tRNA to continue translation.

During translation, several ribosomes may attach one after the other onto an mRNA and proceed through translation as chain of ribosomes. When this happens, the assembled ribosomes and the mRNA together are called polysome.

5.2.7 Ribozyme

Ribozymes are naturally occurring catalytic RNA molecules that have separate catalytic and substrate binding domains. The substrate binding domain binds to specific sequences of substrate RNA molecules by nucleotide complementarity and the catalytic domain cleaves the target RNA at a specific site.

The substrate binding domain can be engineered to bind to any target RNA and thus can be utilized a therapeutic agent. However, susceptibility of RNA molecules to enzymatic degradation in target cells and difficulty associated with the production of large scale synthetic RNA molecules has been solved by synthesizing an oligodeoxynucleotide with a ribozyme catalytic domain (~20 nucleotide) flanked by sequences that hybridize to the target mRNA after transcription. Such synthetic oligonucleotides are amplified and is cloned into eukaryotic expression vector. Transfection of target cells by engineered vectors will produce ribozymes that can cleaves the target mRNA, thereby suppressing the translation of the protein that is responsible for the disorder. Various cancers and viral diseases could be treated with genetically engineered riboszymes.

5.3 Prokaryotic and eukaryotic translation

Eukaryotic mRNA generally encodes a single polypeptide chain but prokaryotic are sometimes polysistronic. A mature mRNA of both prokaryotes and eukaryotes has coding and non-coding sequences.

The 5′ non-coding sequences are referred to as **'5′ untranslated region' (5'UTR).**

In prokaryotes, 5′ UTR have specific sequences called **'Shine Delgarno'** sequence, which just precedes the coding sequence. This sequence aligns the mRNA on 3′ end of 16S rRNA present in small unit of the ribosome.

.....**AGGAGGU**UUGACCU**AUG**..... pro-mRNA

.....**UCCUCCA**..... 16S rRNA

In Eukaryotes, the ribosomes recognize the 7-methylguanosine at the 5′ end. The ribosomes then scan downstream of the 5′ cap until they encounter an initiation codon i.e. AUG.

In both prokaryotic and eukaryotic cells, translation always initiates with the amino acid methionine, usually encoded by AUG. Alternative initiation codons, such as GUG that normally code for valine or CUG arginine are used by bacteria to code for N-fromyl-methionine at the initiation point.

The translation process is a complex mechanism that operates in the cytosol and requires the presence of mRNA, aminoacyl tRNA, ribosomes and many different protein factors. Translation takes place on ribosomes; which can be conceived as a moving protein-synthesizing machine. The entire translation process in both eukaryotes and prokaryotes can be distinguished into **initiation, elongation** and **termination.** Because more is known about translation in bacteria, the process described here will primarily focus bacterial translation. In most respects, eukaryotic translation is similar, although there are some significant differences that will be noted as we proceed through the stages of translation.

5.3.1 Initiation of translation in prokayotes

Initiation steps involves the association of specific **methionyl tRNA** (*f*-met-tRNA), **mRNA** and **ribosome subunits, initiation factors, guanosine triphosphate (GTP)** and recognition of the first codon, which in most cases is **AUG.**

- **IF1** facilitates the separation of the two ribosomal subunits
- At first the small **30S** subunit of the ribosome binds to the protein initiation factor **IF3** (Fig. 5.3a)

- **IF3** and **30S** subunit complex then binds to the Shine-Dalgarno sequence (AGGAGG) present on the mRNA at the 5′ end, approximately 7 nucleotide upstream the start codon **AUG** (AGGAGG: is complementary to the six nucleotides 3′ UCCUCC5′ on the **16SrRNA** at the 3end) (more than one Shine-Dalgarno sequence may be present in a single mRNA and therefore in most cases they are polycistronic).

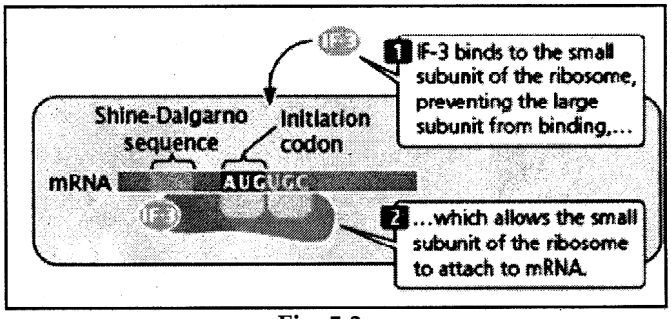


Fig. 5.3a

- Another initiation factor **IF2** that has a GTP bound to it, specifically recognizes and binds to the **initiator *f*-met-tRNA** and facilitates the binding of the *f*-met-tRNA to mRNA 3OS complex (Fig. 5.3b).

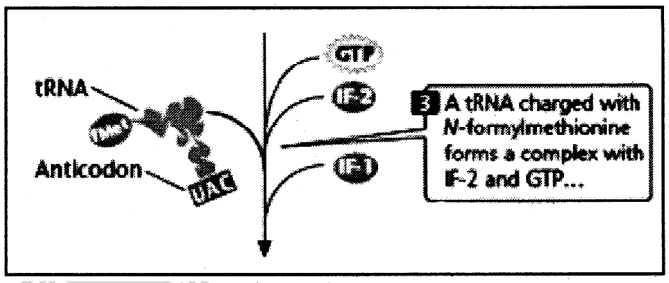


Fig. 5.3b

- The anticodon pairs with the initiation codon AUG on mRNA but lying within the 30S complex
- Binding of the *f*-met-tRNA releases the **IF3** from the initiation complex
- Release of **IF3** allows the 50S ribosomal subunit to bind to the complex
- 50S ribosome then triggers the hydrolysis of the GTP molecule bound to **IF2**
- Hydrolysis of the GTP results in the formation of **70S** complex
- Formation of **70S** ribosome and

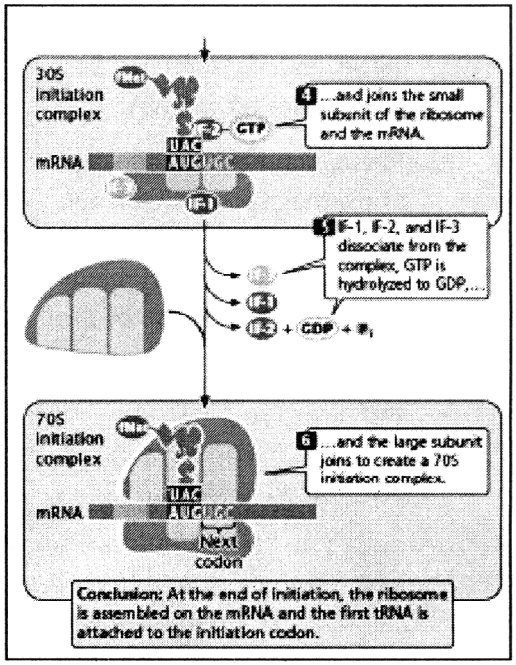


Fig. 5.3c

the binding of *f*-met-tRNA to mRNA at the initiation codon complete the initiation process that is now ready to begin peptide bond formation during elongation. (Fig. 5.3c)

- At initiation, the three nucleotides of the initiation codon **AUG** on mRNA align in the P site of ribosome, where they pair with the anticodon in the initiator tRNA. This arrangement of AUG in ribosome places the second codon in the A site, which is now ready to receive the appropriate charged tRNA to continue with the process of elongation of the polypeptide chain.

5.3.2 Initiation of translation in eukaryotes

In eukaryotes, Shine-Dalgarno sequences are absent and therefore the initiation occurs in a different way.

First, a eukaryotic initiation factor **elF4A,** a multimerie protein has a cap binding protein (CAP), that binds to the cap at the 5′ end of the mRNA. There exists several other factors with helicase activity that helps to unwind the secondary structures that may exists on mRNA.

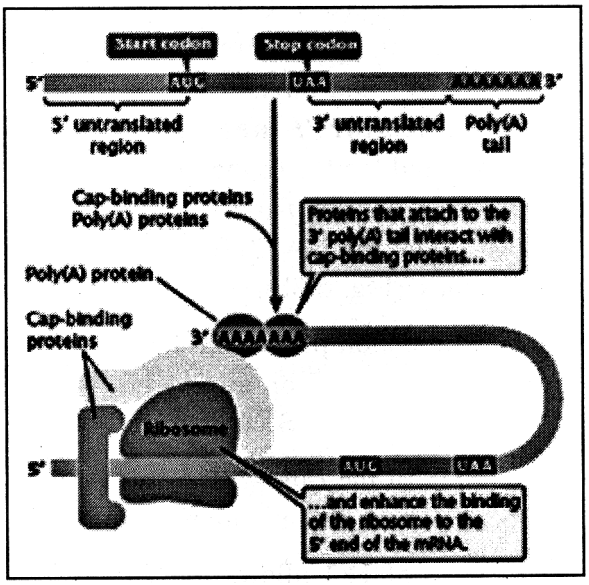


Fig. 5.4: Binding *of* the ribosome to the 5′ end of the mRNA

Then, a complex of the 40S ribosomal subunit with the initiator Met-tRNA.Met, along with several elF binds at the methylated cap of the 5′ end of the mRNA

The ribosomal subunit the scans down the mRNA to locate the initator codon AUG in the concensus in the Kozak sequence **ACC(AUG)G.**

The poly(A) tail at the 3′ end of eukaryotic mRNA also plays a role in the initiation of translation. Proteins that attach to the poly(A) tail interact with proteins that bind to the 5′ cap, enhancing the binding of the small subunit of the ribosome to the 5′ end of the mRNA. This interaction between the 5′ cap and the 3′ tail suggests that the mRNA bends backward during the

initiation of translation, forming a circular structure **(FIG** 2) A few eukaryotic mRNAs contain internal ribosome entry sites, where ribosomes can bind directly without first attaching to the 5' cap.

Once AUG is located, the 40S ribosome firmly binds with it and then 60S ribosomal subunit binds by displacing the elFs, producing the 80S initiation complex.

At initiation, the three nucleotides of the initiation codon (AUG) on mRNA align in the P site of ribosome, where they pair with the anticodon in the initiator tRNA. This arrangement of AUG in ribosome places the second codon in the A site, which is now ready to receive the appropriate charged tRNA to continue with the process of elongation of the polypeptide chain.

5.3.3 Elongation of the polypeptide chain

After the initiation process is over, the elongation of the polypeptide chain begins by joining charged amino acids. The process requires:

(1) the 70S complex;
(2) tRNAs charged with their amino acids;
(3) several elongation factors (EF-Ts, EF-Tu, and EF-G); and
(4) GTP.

A ribosome has three sites that can be occupied by tRNAs; the **aminoacyl,** or A, **site,** the **peptidyl,** or **P, site,** and the **exit,** or **E, site** (Fig. **5.5).**

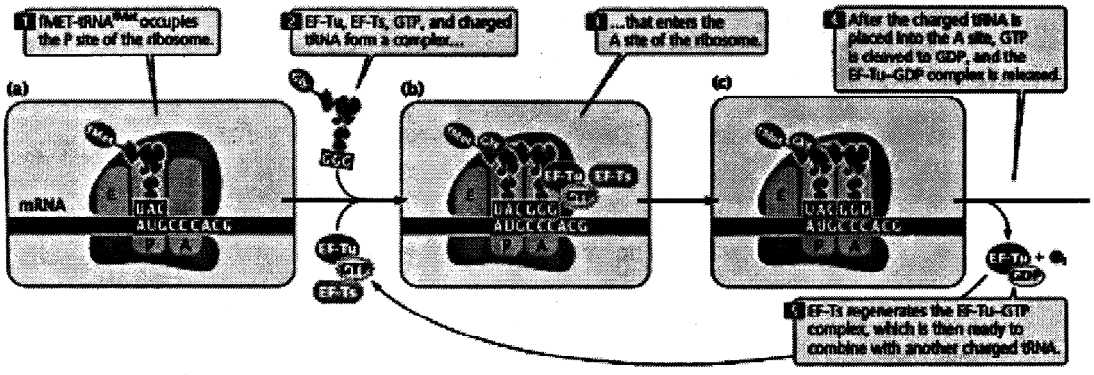


Fig. 5.5 : Sites of a ribosome

At first the ribosome occupies a position on mRNA in such a way that the P site is positioned over AUG and the adjacent A site is unoccupied (Fig 5.5a). The initiator tRNA immediately occupies the P site (the only site to which the fMet-tRNAfMet is capable of binding), but all other tRNAs first enter the A site.

Elongation occurs in three steps.

1. The first step is the delivery of a charged tRNA with its amino acid attached to the A site. This requires the presence of elongation factors **EF-Tu, EF-Ts,** and **GTP** (Fig. 5.5b).

   EF-Tu first joins with GTP and then binds to a charged tRNA to form a three-part complex.

   This three-part complex enters the A site of the ribosome, where the anticodon on the tRNA pairs with the codon on the mRNA.

   After the charged tRNA is in the A site, GTP is cleaved to GDP, and the EF-Tu-GDP complex is released (Fig. 5.5c)

   Factor EF-Ts regenerates EF-Tu-GDP to EF-Tu-GTP.

   In eukaryotic cells, a similar set of reactions delivers the charged tRNA to the A site.

2. The second step of elongation is the creation of a peptide bond between the amino acids that are attached to tRNAs in the P and A sites **(Fig. 5.6**). The

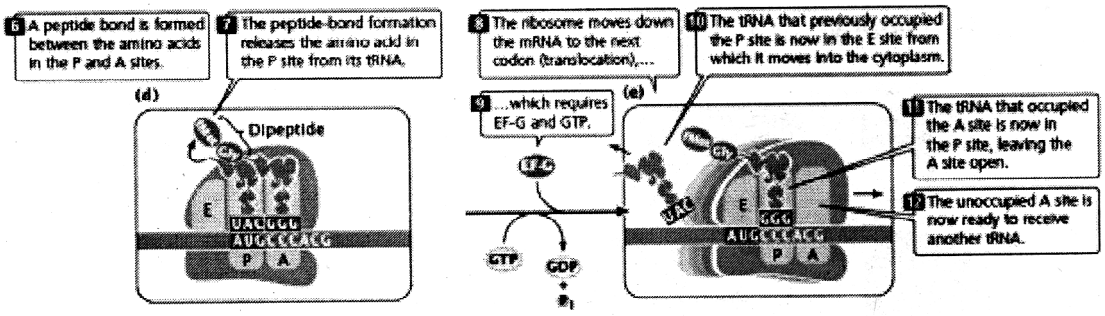


Fig. 5.6: The second step of elongation

formation of this peptide bond releases the amino acid in the P site from its tRNA. The activity responsible for peptidebond formation in the ribosome is referred to as **peptidyl transferase.** The peptidyl bond formation is catalyzed by the rRNA of the large subunit of the ribosome.

The third step in elongation is **translocation, (Fig. 5.7)** where the ribosome moves down the mRNA in the 5′—3′ direction. The A site of the ribosome moves

forward to occupy the next codon. The process requires **EF-G** (elongation factor) and the hydrolysis of **GTP** to **GDP.** Because the tRNAs in the P and A site are still attached to the mRNA through codon- anticodon pairing, they do not move with the ribosome as it translocates. Consequently, the tRNA that previously was occupying the P site now occupies the E site, from which it moves into the cytoplasm. tRNA that occupied the A site now occupies the P site, leaving the A site open for the next incoming tRNA specified by the mRNA's codon sequence. Thus, the progress of each tRNA through the ribosome during elongation can be summarized as follows:

Cytoplasm → A site → P site → E site → cytoplasm

Throughout the cycle, the polypeptide chain remains attached to the tRNA in the P site. The ribosome moves down the mRNA in the 5′ —3′ direction, adding amino acids one at a time.

Elongation in eukaryotic cells takes place in a similar manner.

5.3.4 Termination

Addition of new amino acids stops when the A site of ribosome translocates to a termination codon. Because there are no tRNAs with anticodons complementary to the termination codons, no tRNA enters the A site of the ribosome when a termination codon is encountered (**Fig. 5.7a**). Instead, proteins called **release factors** bind to the ribosome (**Fig. 5.7b**). *E. coli* has three release factors—RF1, RF2, and RF3. Release factor 1 recognizes the termination codons UAA and UAG, and RF2 recognizes UGA and UAA. Release factor 3 forms a complex with GTP and binds to the ribosome. The release factors then promote the cleavage of the tRNA in the P site from the polypeptide chain; in the process, the GTP that is

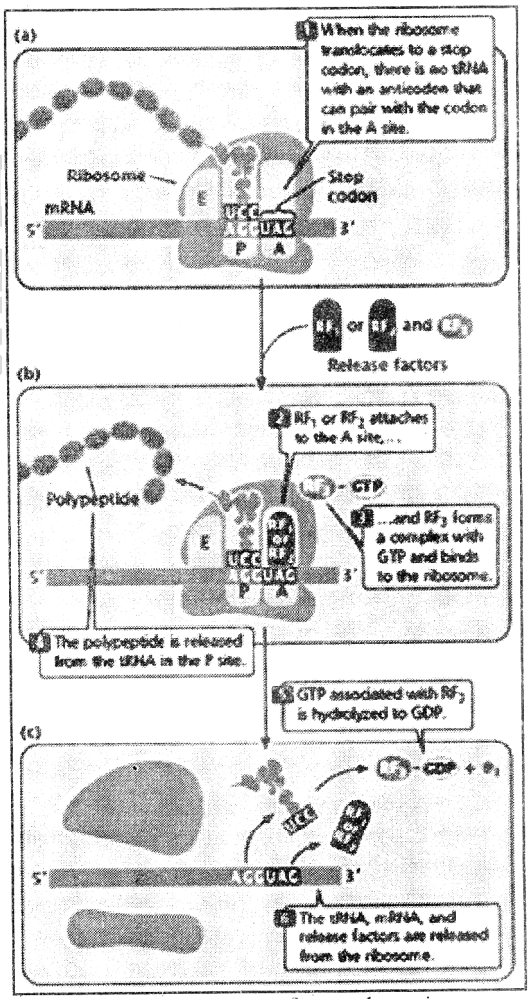


Fig. 5.7: Steps of translocation

complexed to RF3 is hydrolyzed to GDP. Additional factors help bring about the release of the tRNA from the P site, the release of the mRNA from the ribosome, and the dissociation of the ribosome **(Fig. 5.7c).** Findings from recent studies suggest that the release factors bring about the termination of translation by completing a final elongation cycle of protein synthesis. In this model, RF1 and RF2 are similar in size and shape to tRNAs and occupy the A site of the ribosome, just as the amino acid-tRNA-EF-Tu-GTP complex does during an elongation cycle. Release factor 3 is structurally similar to EF-G; it then translocates RF1 and RF2 to the P site, as well as the last tRNA to the E site, in a way similar to that in which EF-G brings about trans location. When both the A site and the P site of the ribosome are cleared of tRNAs, the ribosome can dissociate. Research findings also indicate that some of the sequences in the rRNA play a role in the recognition of termination codons.

Translation in eukaryotic cells terminates in a similar way, except that there are two release factors: eRFl, which recognizes all three termination codons, and eRF2, which binds GTP and stimulates the release of the polypeptide from the ribosome.

5.4 Regulation of translation

Regulation of gene expression refers to the cellular control of the amount and timing of changes to the appearance of the functional product of a gene. Gene regulation gives the cell control over its structure and function, and is the basis for cellular differentiation, morphogenesis and the versatility and adaptability of any organism. Prokaryotic and eukaryotic gene expressions are regulated at several different levels of flow of information from gene to their final product. Any step of gene expression may be modulated, from the DNA-RNA transcription step to post-translational modification of a protein. Stages where gene expression is regulated are:

Regulate gene expression by chemical and structural modification of DNA or chromatin (e.g. by methylation, phosphorylation, acetylation or structural changes)

By controlling when and how often a given gene is transcribed (during transcription e.g. by repressor or activator proteins, attenuation etc.)

By modifying the primary RNA transcript (post transcriptional modifications)

Selection of transcribed and processed RNA (transport from nucleus to cytoplasm)

During translation of mRNA (e.g. by an antisense RNA)

Post-translational control (e.g. by proteolysis or modification of the gene product)

5.4.1 DNA modification

1. Chemical modification of DNA

Methyiation of DNA refers to addition of a methyl group to the number 5 carbon of the cytosine pyrimidine ring by the enzyme methyl transferase. Generally methyiation occurs on the cytosine in the CpG dinucleotide sequence (CpG islands). DNA methyiation pattern can be inherited without changing the DNA sequence. As such, it is part of the epigenetic code and is the most characterized epigenetic (changes in phenotype without any alteration in the genomic material) mechanism.

DNA methyiation has been found in all vertebrate. In humans, approximately 1% of DNA bases undergo DNA methyiation (~60-70% of all CpGs are methylated mainly 5′ regulatory regions). In adult somatic tissues, DNA methyiation typically occurs in a CpG dinucleotide context; non-CpG methyiation is prevalent in embryonic stem cells. In plants, cytosines are methylated both symmetrically (CpG or CpNpG) and asymmetrically (CpNpNp), where N can be any nucleotide.

DNA methyiation may impact the transcription of genes in two ways.

1. First, the methyiation of DNA may itself physically impede the binding of transcriptional proteins to the gene, thus blocking transcription.

2. Second, and likely more important, methylated DNA may be bound by proteins known as Methyl-CpG-binding domain proteins' (MBDs). MBD proteins then recruit additional proteins to the locus, such as histone deacetylases and other chromatin remodelling proteins that can modify histones, thereby forming compact, inactive chromatin termed silent chromatin.

This link between DNA methyiation and chromatin structure is very important. . In many disease processes such as cancer, gene promoter CpG islands acquire abnormal hypermelhylation, which results in heritable transcriptional silencing In particular, loss of Methyl-CpG-binding Protein 2 (MeCP2) has been

implicated in Rett syndrome and Methyl-CpG binding domain protein 2 (MBD2) mediates the tran-scriptional silencing of hypermethylated genes in cancer

Analysis of the pattern of methylation in a given region of DNA (generally a promoter) can be achieved through a method called bisulfite mapping. Methylated cytosine residues are unchanged by the treatment, whereas unmethylated ones are changed to uracil. The differences are analyzed **in** sequencing gels, Abnormal methylation patterns are thought to be involved in carcinogenesis.

2. Structural modification of DNA

Transcription of DNA is highly dependent on the secondary structure of DNA molecule. Histone proteins, responsible for supercoiling of DNA can modify the structure temporarily or more permanently depending on the phosphorylation or methylation of the histone proteins respectively. Such modifications influences the level of gene expression. In general, the density of its packing is indicative of the frequency of transcription.

Histone acetylation is also an important process in transcription. Histone acetyltransferase enzymes (HATs) such as CREB-binding protein also dissociate the DNA from the histone complex, allowing transcription to proceed. Often, DNA methylation and histone acetylation work together in gene silencing. The combination of the two seems to be a signal for DNA to be packed more densely, lowering gene expression.

5.4.2 Regulation of transcription

Transcription is an important level of control in eukaryotic cells. Transcription regulation of a gene by RNA polymerase can be regulated by at least five mechanisms:

Specificity factors alter the binding specificity of RNA polymerase for a given promoter or set of promoters, making it more or less likely to bind to them (i.e. sigma factors used in prokaryotic transcription).

Repressors bind to non-coding sequences on the DNA strand that are close to or overlapping the promoter region, impeding RNA polymerase's progress along the strand, thus impeding the expression of the gene.

Basal factors These transcription factors position RNA polymerase at the start of a protein-coding sequence and then release the polymerase to transcribe the mRNA. Recruitment of these proteins at the promoter region affects the RNA polymerase activity.

Activators enhance the interaction between RNA polymerase and a particular promoter, encouraging the expression of the gene. Activators do this by increasing the attraction of RNA polymerase for the promoter, through interactions with subunits of the RNA polymerase or indirectly by changing the structure of the DNA.

Enhancers are sites on the DNA helix that are bound to by activators in order to loop the DNA bringing a specific promoter to the initiation complex.

Examples :

The $ó^{32}$ subunit of RNA polymerase changes itself in such a way that the enzyme binds to a specialized set of promoters when E. coli bacteria are subjected to heat stress producing heat-shock response proteins.

When there is excess tryptophan in the cell, the amino acid binds to a specialized repressor protein, changing the structural conformity of the repressor such that it binds to the operator region for the operon that synthesizes tryptophan, preventing their expression and thus suspending production which also represents a form of negative feedback mechanism.

In bacteria, the lac repressor protein blocks the synthesis of enzymes that digest lactose when there is no lactose to feed on. When lactose is present, it binds to the repressor, causing it to detach from the DNA strand.

5.4.3 Gene regulation can be summarized as how they respond

Inducible systems - An inducible system is off unless there is the presence of some molecule (called an inducer) that allows for gene expression. The molecule is said to "induce expression".

Repressible systems - A repressible system is on except in the presence of some molecule (called a corepressor) that suppresses gene expression. The molecule is said to "repress expression". In both the cases, the control mechanism varies in prokaryotic and eukaryotic cells.

5.4.4 Post-transcriptional regulation

The cells regulate the posttranscriptional activity by several way to check how much the mRNA should be translated into proteins. Cells do this by Capping, Splicing, and the addition of a Poly(A) Tail. These processes occur only in

eukaryotes because in prokaryotes, the transcription and translation is coupled.

Capping changes the five prime end of the mRNA to a three prime end by 5'-5' linkage, which protects the mRNA from 5' exonuclease, which degrades foreign RNA. The cap also helps in ribosomal binding.

Splicing removes the introns, noncoding regions that are transcribed into RNA, in order to make the mRNA able to create proteins. Cells do this by spliceosome's binding on either side of an intron, looping the intron into a circle and then cleaving it off. The two ends of the extrons are then joined together.

Addition of poly(A) tail a poly(A) tail is just junk RNA added to the 3' end in order to slowly be degraded by a 3' exonuclease in order to increase the half life of mRNA.

Unit 6 Antisense and Ribozyme Technology

Structure

6.1 Antisense molecules and their mechanism of action

6.1.1 Introduction

Expression of some genes may be regulated or suppressed with the aid of small RNA molecules - a process termed as RNA silencing, also known as RNA interference or posttranscriptional gene silencing. Although many of the details of this mechanism are still poorly understood, it appears to be widespread, existing in fungi, plants, and animals. It may also prove to be a powerful tool for artificially regulating gene expression in genetically engineered organisms.

The discovery of antisense RNA was preceded first by observation of transcriptional inhibition by antisense RNA expressed in transgenic plants. In an attempt to alter flower colors in petunias, researchers introduced additional copies of a gene encoding chalcone synthase, a key enzyme for flower pigmentation into petunia plants of normally pink or violet flower color. The scientists' goal was to produce petunia plants with improved flower colors but instead produced less pigmented, fully or partially white flowers, indicating that the activity of chalcone synthase had been substantially decreased. Similar suppression of gene activity was also observed in fungus *Neurospora crassa.* This phenomenon was called *co-suppression of gene expression,* but the molecular mechanism remained unknown. Sometime later, plant virologists noted that plants carrying only short transgenic non-coding regions of viral RNA sequences would showed enhanced levels of protection. The reverse experiment, in which short sequences of plant genes were introduced into viruses, showed that the targeted gene was suppressed in an infected plant. This phenomenon was labeled "virus-induced gene silencing" (VIGS), and the set of such phenomena were collectively called post transcriptional gene silencing. Craig C. Mello and Andrew Fire's 1998 published in *Nature* the potent gene silencing effect after injecting double stranded RNA into C. *elegans.* In investigating the regulation of muscle protein production, they observed that neither mRNA nor antisense RNA injections had an effect on protein production,

but double-stranded RNA successfully silenced the targeted gene. As a result of this work they were awarded Nobel Prize in the year 2006.

6.1.2 Antisense RNA molecules

Antisense molecules are nucleotide sequences that interact with complementary strands of nucleic acids and modify expression of genes. For example, Antisense RNA is single-stranded RNA that is complementary to an mRNA strand transcribed within a cell and capable of blocking the translation machinery. Antisense molecules occur naturally. For example, in both mice and humans, the gene for the insulin-like growth factor 2 receptor (Igf2r) that is inherited from the father synthesizes an antisense RNA that appears to block synthesis of the mRNA for Igf2r.

Historically, the effects of antisense RNA have often been confused with the effects of RNA interference, a related process in which double-stranded RNA fragments called small interfering RNAs trigger catalytically mediated gene silencing, most typically by targeting the RNA-induced silencing complex (RISC) to bind to and degrade the mRNA.

6.1.3 RNA interference (RNAi)

In the course of working with artificially synthesized single stranded antisense RNA molecules, it was discovered that double stranded RNA (dsRNA) molecule can also act as a powerful suppressant of genes expression. dsRNA may arise in several ways: by the transcription of inverted repeats in DNA into a single RNA molecule that base pairs with itself; by the simultaneous transcription of two different RNA molecules that are complementary to one another and pair; or by the replication of double-stranded RNA viruses (Fig 6.1).

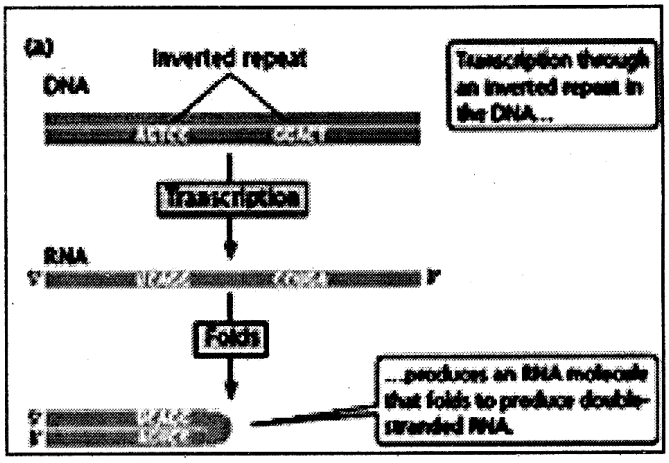


Fig. 6.1: Replication of double-stranded RNA viruses

Drosophila, an enzyme called Dicer cleaves and processes the double-stranded RNA to produce In fact, the suppressive effect of antisense RNA probably also depends on its ability to form dsRNA. The ability of dsRNA to suppress the

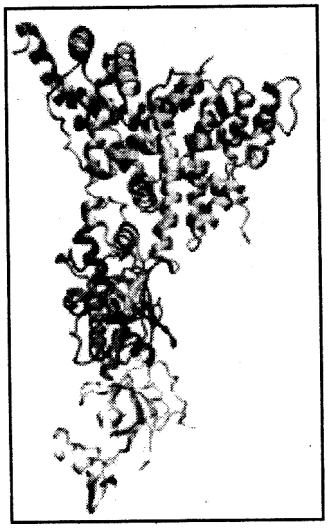


Fig 6.1 a.

expression of a gene corresponding to its own sequence is called RNA interference (RNAi). It is also called post-transcriptional gene silencing or PTGS.

6.1.4 Mechanism of RNAi

Normally single-stranded RNA molecules are found in the cytoplasm of a cell. If double-stranded RNA (dsRNA) formation occurs in the cell, the enzyme called Dicer to cleave the dsRNA into fragments containing 19 base pairs (~2 turns of a double helix) with two additional nucleotides at the opposite end of each strand. The two strands of each fragment then separate — releasing the antisense strand.

Dicer *is a ribonuclease in the RNase III family that cleaves double-stranded RNA (dsRNA) and pre-microRNA (miRNA) into short double-stranded RNA fragments called small interfering RNA (siRNA) about 20-25 nucleotides long, usually with a two-base overhang on the 3' end. Dicer contains two RNase III domains and one PAZ domain; the distance between these two domains of the molecule is determined by the length and angle of the connector helix and determines the length of the siRNAs it produces. Dicer catalyzes the first step in the RNA interference pathway and initiates formation of the RNAAnduced silencing complex (RISC), whose catalytic component argonaute is an endonuclease capable of degrading messenger RNA (mRNA) whose sequence is complementary to that of the siRNA guide strand.*

With the aid of a protein, it binds to a complementary sense sequence on a molecule of mRNA. If the base-pairing is exact, the mRNA is destroyed. Because of their action, these fragments of RNA have been named "short (or small)

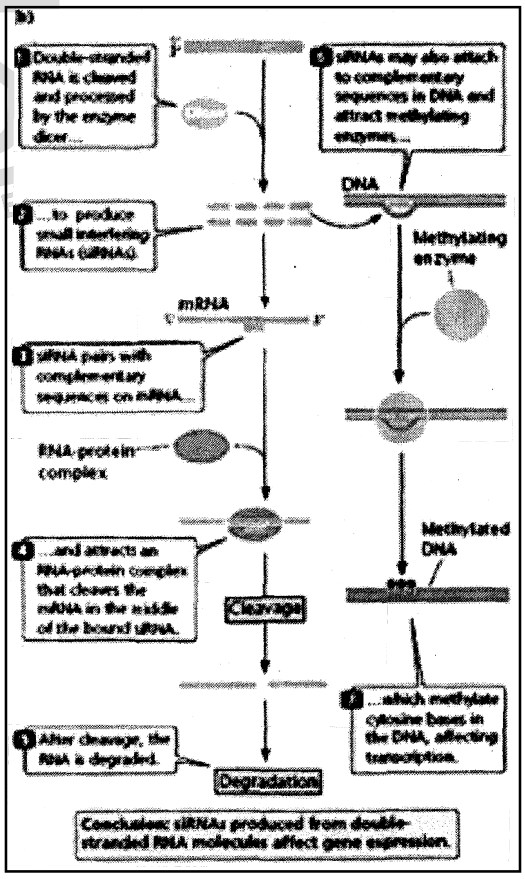


Fig. 6.2b:

interfering RNA" **(siRNA).** The complex of siRNA and protein is called the "**R**NA-induced silencing complex" **(RISC)** (Fig. 6.2b).

In fission yeast, evidences indicate that siRNAs can also inhibit the transcription of genes perhaps by binding to complementary sequences on DNA or by binding to the nascent RNA transcript as it is being formed. Synthetic siRNA molecules that bind to gene promoters can repress transcription by methylation of the DNA in the promoter and, perhaps, methylation of histones in the vicinity. The siRNA forms a complex called the RITS complex ("RNA-induced initiation of transcriptional gene silencing") with at least three different proteins. How these siRNAs, synthesized in the cytosol gain access to the DNA in the nucleus is unknown.

Example :

The rice plant of strain LGC-1 produces abnormally low levels of proteins called glutelins although there are several glutelin genes. Interestingly, it was observed that two closely-similar glutelin genes are located back to back on the same chromosome and a deletion has occurred in the 3' region of the first glutelin gene that has removed the stop signal. As a consequence, RNA polymerase II transcribes right past the first gene and on into the second. The result is a messenger RNA with almost-identical sequences running in opposite directions. Such a composition of the mRNA molecule allows it to fold up into a molecule of double-stranded RNA (dsRNA). A Dicer-like enzyme cuts up the dsRNA into small interfering RNAs (siRNAs) that suppress further transcription of those genes as well as other glutelin genes

Some related. RNA molecules produced through the cleavage of double-stranded RNA bind to complementary sequences in the 3′UTR of mRNA and inhibit their translation. RNA silencing is thought to have evolved as a defense against RNA viruses and transposable elements that move through an RNA intermediate (see Chapter 20). The extent to which it contributes to normal gene regulation is uncertain, but dramatic phenotypic effects result from some mutations that occur in the enzymes that carry out RNA silencing.

Amplification of RNAi : In C. elegans, plants, and Neurospora, the introduction of a few molecules of dsRNA has a potent and long-lasting effect. In plants, the gene silencing spreads to adjacent cells (through plasmodesmata) and even to other parts of the plant (through the phloem). RNAi within a cell can continue after mitosis in the progeny of that cell. Triggering of RNAi in C. elegans can even pass through the germline into its descendants. Such amplification of an initial trigger signal suggests a catalytic effect. It turns out that these organisms have RNA-dependent RNA poly-merases (RdRPs) that uses the mRNA targeted

by the initial antisense siRNA as a template for the synthesis of more siRNAs. Synthesis of these "secondary siRNAs even occurs in adjacent regions of the mRNA. So not only can these secondary siRNAs target additional areas of the original mRNA, but they are potentially able to silence mRNAs of other genes that may carry the same sequence of nucleotides. This phenomenon, called **"transitive RNAi".**

In mammalian cells, introducing dsRNA fragments only reduces gene expression temporarily. However, Brummelkamp *et. al.* report in the 19 April 2002 issue of **Science** that they have succeeded in introducing into (mammalian) cells a **DNA vector** that can continuously synthesize a siRNA corresponding to the gene that they want to suppress. Two months later the cells still failed to manufacture the protein whose gene had been turned off by RNAi.

6.1.5 MicroRNAs (miRNAs)

MicroRNAs were first described by Lee, et al. in 1993. **MicroRNAs (miRNA)** are single-stranded RNA molecules of about 21-23 nucleotides in length thought to regulate the expression of other genes. miRNAs are encoded by genes that are transcribed from DNA but not translated into protein (non-coding RNA); instead they are processed from primary transcripts known as *pri-miRNA* to short stem-loop structures called *pre-miRNA* and finally to functional miRNA, Mature miRNA molecules are partially complementary to one or more messenger RNA (mRNA) molecules, and they function to downregulate gene expression.

The genes encoding miRNAs are much longer than the processed miRNA molecule. miRNAs are first transcribed as primary transcripts or pri-miRNA and processed to short, 70-nucleotide stem-loop structures known as pre-miRNA in the cell nucleus by protein complex known as the Microprocessor complex, consisting of the nuclease Drosha and the double-stranded RNA binding protein Pasha. These pre-miRNAs are then processed to mature miRNAs in the cytoplasm by interaction with the endonuclease Dicer, which also initiates the formation of the RNA-induced silencing complex (RISC). This complex is responsible for the gene silencing observed due to miRNA expression and RNA interference. The pathway in plants varies slightly due to their lack of Drosha homologs; instead, Dicer homologs alone effect several processing steps.

In *C. elegans,* successful development through its larval stages and on to the adult requires the presence of at least two **"microRNAs"** ("miRNAs") — single-stranded RNA molecules containing about 22 nucleotides and thus about the same size as siRNAs. These small single-stranded transcripts are generated by the cleavage of larger precursors using the C. elegans version of Dicer. The miRNA acts by either destroying or inhibiting translation of several messenger RNAs in

the worm (by binding to a region of complementary sequence in the 3' untranslated region [3^1-UTR] of the mRNA). miRNA genes have also been discovered in humans Dro-sophila, mice, frogs, fish, and plants (Arabidopsis) as well as in C. elegans.

6.1.6 Biological functions of antisense molecules

Immunity :

RNA interference provides immunity to plants especially against viruses and other foreign genetic materials. It is also suggested the RNA interference may also prevent self-propagation by transposons. Even before the RNAi pathway was fully understood, it was known that induced gene silencing in plants could spread throughout the plant in a systemic effect, and could be transferred from stock to scion plants via grafting. This phenomenon is a feature of the plants innate immune system, and allows the entire plant to respond to a virus after an initial localized encounter.

Although animals generally express fewer variants of the dicer enzyme than plants, RNAi in some animals has also been shown to produce an antiviral response. In both juvenile and adult *Drosophila*, RNA interference is important in antiviral innate immunity and is active against pathogens such as Drosophila X virus. A similar role in immunity may operate in *C. elegans*, as argonaute proteins are upregulated in response to viruses and worms that overexpress components of the RNAi pathway are resistant to viral infection.

The role of RNA interference in mammalian innate immunity is poorly understood and the hypothesis of RNAi-mediated immunity in mammals has been challenged as relatively little data is available. However, alternative functions for RNAi in mammalian viruses exist, such as miRNAs expressed by the herpes virus that may act as heterochromatin organization that triggers to mediate viral latency.

Genome maintenance

Components of the RNA interference pathway are used in many eukaryotes in the maintenance of the organisation and structure of their genomes. Modification of histones and associated induction of heterochromatin formation serves to downregulate genes pre-transcriptionally and this process is referred to as RNA-induced transcrip-tional silencing (RITS), and is carried out by a complex of proteins called the RITS complex. In fission yeast this complex contains argonaute, a chromodomain protein Chpl, and a protein called Tas3 of unknown function. As a consequence, the induction and spread of heterochromatic regions requires the argonaute and RdRP proteins. Indeed, deletion of these genes in the fission yeast *S. pombe* disrupts histone meth-ylation and centromere formation, causing slow or

stalled anaphase during cell division. Thus repression of gene expression by miRNAs appears to be a mechanism to ensure proper, coordinated gene expression as cells differentiate along particular paths. For example, when zygote. genes begin to be turned on in the zebrafish blas-tula, one of them encodes a miRNA that triggers the destruction of the maternal mRNAs that have been running things up to then. So miRNAs may play as important role as transcription factors in coordinating the expression of multiple genes in a particular type of cell at particular times.

6.2 Splicing repressors and activators control splicing at alternative sites

6.2.1 Introduction

Gene splicing mechanism exists in eukaryotes that facilitate the removal of intron from pre-mRNA. There are also evidences to suggest that alternative splicing enable the synthesis of two different proteins from the same peptide. For the alternative splicing mechanism to operate, it is believed that there exists splicing repressor proteins in the cells and function to produce alternative peptides from the same genes. For example, in *Drosophila* Sxl inhibit splicing at specifice sites, causing exons to be skipped, whereas Tra promotes splicing. As a consequence,

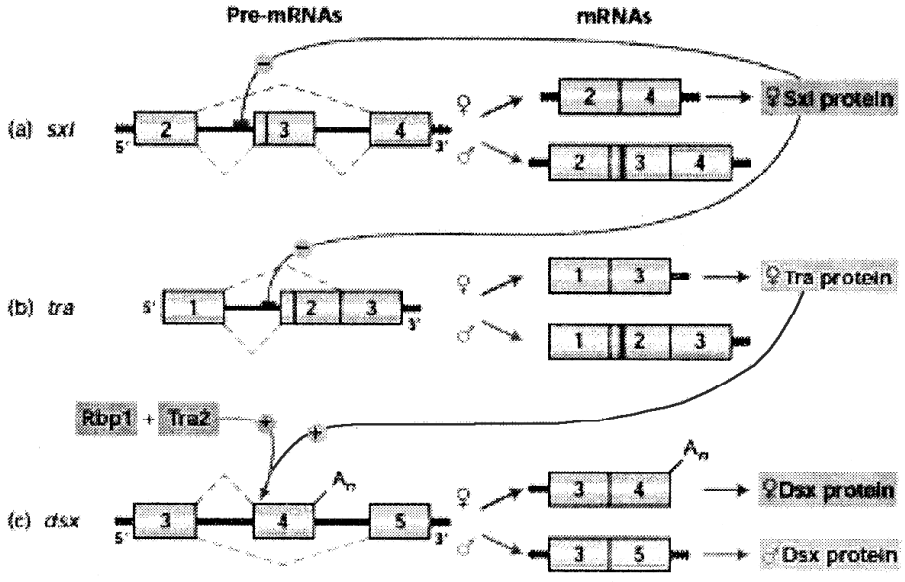


Fig. 6.3 : Alternative spilicing due to repression at splicing sites lead to the formation of alternative Dsx protein in Drosophila

distinct Dsx proteins are produced in female and male embryos leading to sexual differentiation (Fig. 6.3). Sxl-like splicing repressor expressed in hepatocytes might bind to splice sites for the EIIIA and EIIIB exons in the fibronectin pre-mRNA, causing them to be skipped during RNA splicing. Experimental examination in some systems has revealed that inclusion of an exon in some cell types versus skipping of the same exon in other cell types results from the combined influence of several splicing repressors and enhancers.

The action of similar proteins may explain the cell-type specific expression of fibronectin isoforms in humans. Splicing repressors expressed in hepatocytes might **bind** to splice sites for the EIIIA and EIIIB exons in the fibronectin pre-mRNA, causing them to be skipped during RNA splicing (Figure: 6.4). Alternatively, a Tra-like splicing activator expressed in fibroblasts might activate the splice sites associated with the fibronectin EIIIA and EIIIB exons, leading to inclusion of these exons in the mature mRNA.

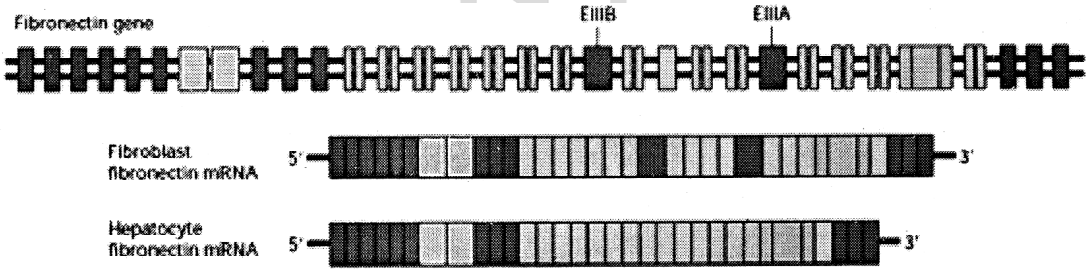


Fig.6.4: Alternative splicing fibronectin pre mRNA in human hepatocytes can produce different isoforms.

6.2.2 Repression of translation of mRNAs

Micro RNAs (miRNAs) are efficient tools for the sells to regulate mRNA translation. In C. elegans, the genes lin-4 and let-7 produce small RNA sequences 21 and 22 nucleotides long, respectively, that hybridize to the 3' untranslated regions of specific target mRNAs. For example, the *lin-4* miRNA, which is expressed early in embryogenesis, hybridizes

to the 3' untranslated regions of both the *lin-14* and *lin-28* mRNAs, thereby repressing translation of these mRNAs by an as yet unknown mechanism. Expression of *lin-4* miRNA ceases later in development, allowing translation of newly synthesized *lin-14* and *lin-28* mRNAs at that time.

In *C. elegans;* about 100 different miRNAs have been found in *C. elegans*, and

at least as many in humans. All miRNAs appear to be formed by processing of ~70-nucleotide precursor RNAs that form hairpin structures with a few base-pair mismatches in the stem of the hairpin. A ribonuclease called *Dicer*, cleaves the double-stranded RNA to produce miRNAs precursors. Interestingly, the base pairing between a miRNA and the 3' untranslated region of its target mRNAs is not precisely complementary and some base-pair mismatches occur in the hybridized region. This mismatching distinguishes miRNA-mediated translational repression from the related phenomenon of RNA interference, which we describe next.

6.2.3 Degradation of mRNAs in the cytoplasm

Concentration of any mRNAs in the cell depends not only on the rate of its transcription but also on the rate of its degradation. With more stable mRNAs, protein synthesis persists long after transcription of the gene is repressed. Usually, bacterial mRNAs are unstable and decay exponentially, probably because they need to switch genes rapidly in response to the change in environment. In multicellular organism, the cells reside in a more stable environment and do not require frequent adjustment to the changing environment. However, some proteins in eukaryotic cells are required only for short periods of time and must be degraded immediately. For example, during cell cycle activity, synthesis of cyclins occurs in burst at intervals. The mRNA of cyclins must be degraded very quickly. mRNA of other proteins like c-Fos and c-Jun, synthesized during S phase need to be degraded immediately after the function is over. mRNAs of such proteins have half life less than 30 minutes.

There are three main pathway that lead to the degradation of cytoplasmic mRNAs as shown in (Fig. 6.5).

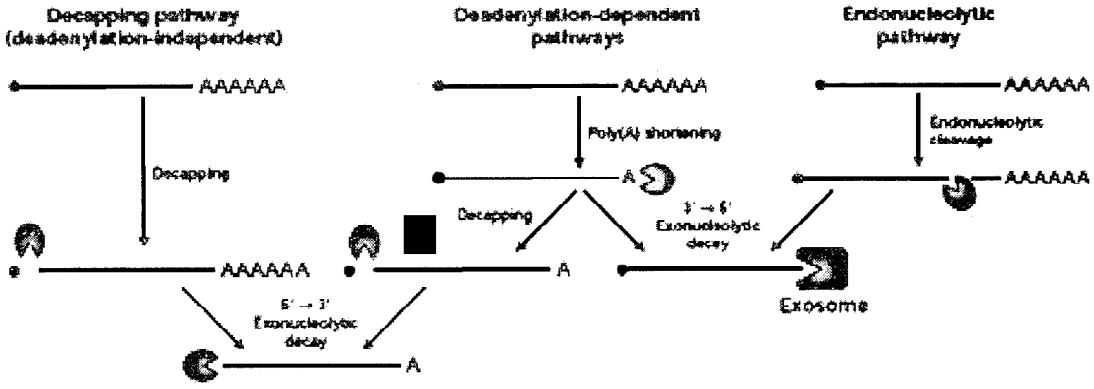


Fig.6.5: Main pathway that lead to degradation of cytoplasmic mRNAs

For most mRNAs, the length of the poly(A) tail gradually decreases with time through the action of a deadenylating nuclease. When it is shortened sufficiently, the PABPI molecules that bind during polyadenylation of mRNA can no longer bind and stabilize interaction of the 5′ cap and initiation factors (Fig. 6.6). The exposed cap then is removed by a decapping enzyme, and the unprotected mRNA is degraded by a 5′→3′ exonuclease. Removal of the poly(A) tail also makes mRNAs susceptible to degradation by cytoplasmic exosomes containing 3′→5′ exonucleases. The 5′→3′ exonucleases predominate in yeast, and the 3′→5′ exosome apparently predominates in mammalian cells. The rate of deadenylation determines the rate of degradation of mRNA. Recent experiments suggest that the bound proteins interact with a deadenylating enzyme and with the exosome, thereby promoting the rapid deadenylation and subsequent 3′→5′ degradation of these mRNAs. In this mechanism, the rate of mRNA degradation is uncoupled from the frequency of translation. Thus mRNAs containing the AUUUA sequence can be translated at high frequency, yet also degraded rapidly, allowing the encoded proteins to be expressed in short bursts.

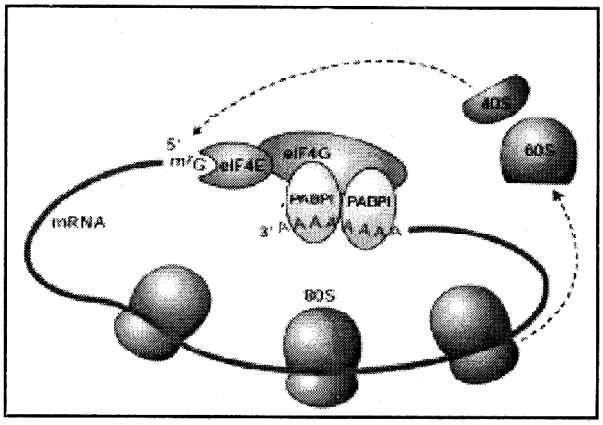


Fig. 6.6 : Initiation factors

Some mRNAs are degraded by decapping the mRNA before the deadenylation process is initiated. It appears that certain mRNA sequences make the cap sensitive to the decapping enzyme, but the precise mechanism is unclear.

In the other alternative pathway, mRNAs first are cleaved internally by endonucleases. The RNA-induced silencing complex (RISC) discussed earlier is an example of such an endonuclease. The fragments generated by internal cleavage then are degraded by exonucleases.

6.2.4 Regulation of mRNA translation and degradation

Ironresponse element-binding protein *(IRE-BP)* can be a classis example of protein that regulates the translation. Intracellular iron concentrations the protein in way that can regulate the translation of one mRNA and the degrade another. When the intracellular iron falls below the threshold level, IRE-BP proteins releases free irons in the system for the enzymes that require Fe to function. Again, when

the concentration of free iron increases within the system, the proteins bind to free Fe to prevent accumulation and toxicity.

Production of *ferritin*- an intracellular iron-binding protein is regulated by IRE-BP. The 5′ UTR of ferritin mRNA has a stem and loop structure containing a iron response element ***(IRE).*** The IRE-BP recognizes five specific bases in the IRE loop and the duplex nature of the stem. At low iron concentrations, IREBP is in active conformation and binds to the IREs. The bound IRE-BP blocks the 40S ribosomal subunit from scanning for the AUG start codon, thereby inhibiting translation initiation. The resulting decrease in ferritin means less iron is complexed with the ferritin and is therefore available to iron-requiring enzymes. At high iron concentrations, IRE-BP is in an inactive conformation that does not bind to the 5' IREs, so translation initiation can proceed. The newly synthesized ferritin then binds free iron ions, preventing their accumulation to harmful levels (Fig. 6.7a).

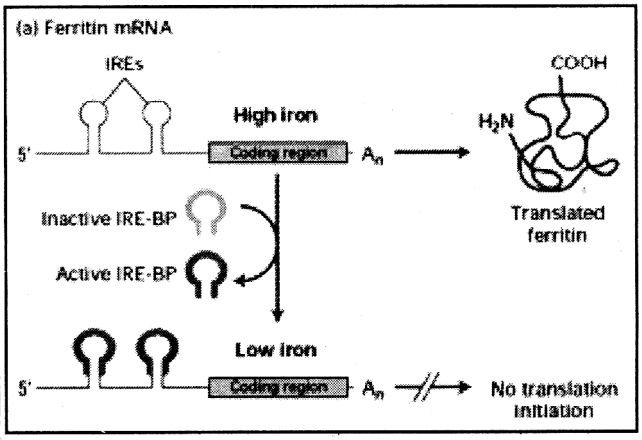


Fig. 6.7 a :

In vertebrates import of iron is also regulated. Ingested iron is carried through the circulation bound to a protein called *transferrin.* After binding to the transferrin receptor (TfR) in the plasma membrane, the transferrin-iron complex is brought into cells by receptor-mediated en-docytosis. The 3′-UTR of TfR mRNA contains IREs whose stems have AU′rich destabilizing sequences (Fig. 6.7b), At high iron concentrations, when the IRE-BP is in the inactive, nonbinding conformation, these AU-rich sequences are thought to promote degradation of TfR mRNA by the same

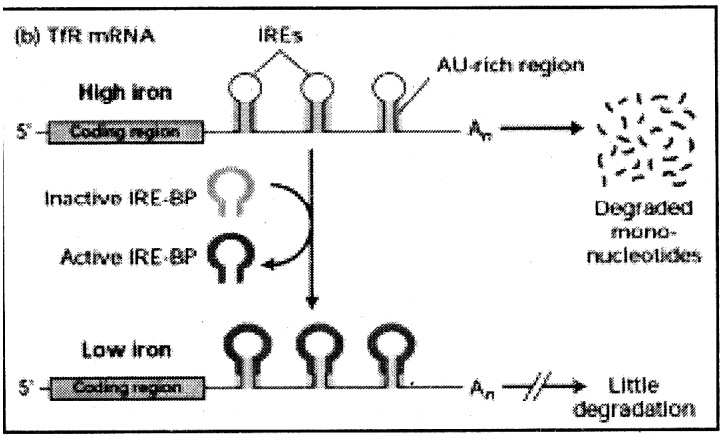


Fig. b.7b.

mechanism that leads to rapid degradation of other shortlived mRNAs, as described previously. The resulting decrease in production of the transferrin receptor quickly reduces iron import, thus protecting the cell. At low iron concentrations, however, IRE-BP can bind to the 3' IREs in TfR mRNA. The bound IRE-BP is thought to block recognition of the destabilizing AU-rich sequences by the proteins that would otherwise rapidly degrade the mRNAs. As a result, production of the transferrin receptor increases and more iron is brought into the cell.

6.2.5 Nonsense-mediated decay and other mRNA

Improperly processed mRNA cannot be translated and should be eliminated out of the system which otherwise can lead to production of an abnormal protein that interferes with functioning

of the normal protein. This effect is equivalent to dominant-negative mutations. Several mechanisms collectively termed **mRNA surveillance** help cells avoid the translation of improperly processed mRNA molecules.

Nonsense-mediated decay mechanism is another way how the cells get cleared of the wrongly processed mRNAs in which one or more exons have been skipped during splicing. During splicing, improper exon skipping sometime introduce stop codons. Nonsense-mediated decay results in the rapid degradation of mRNAs with stop codons that occur before the last splice junction in the mRNA. Analysis of yeast mutants suggests that some of the proteins in exon-junction complexes function in nonsensemediated decay.

6.3 Ribozymes

6.3.1 Concept of ribozymes

Until about 20 years ago, all known enzymes were proteins. But then it was discovered that some RNA molecules can act as enzymes; that is, catalyze covalent changes in the structure of substrates, RNA molecules that can catalyze a chemical reaction are called RIBOZYMES. The first ribozymes were discovered in the 1980s by Thomas R. Cech, who was studying RNA splicing in the ciliated protozoan *Tetrahymena therrnophila.* Subsequently, Sidney Altman, discovered bacterial RNase P complex. The ribozymes were found in the intron of an RNA **transcript,** which removed itself from the transcript and in the RNA component of the RNase P complex, which is involved in the maturation of pre-tRNAs. Many natural ribozymes catalyze either their own cleavage or the cleavage of other RNAs, but they have also been found to. catalyze the aminotransferase activity of the ribosome. Although most ribozymes are quite rare in the cell, their roles are sometimes essential to life.

Five classes of ribozymes have been described based on their unique characters in the sequences as well as three-dimensional structures (Bunnell, 1997). They are denoted as (1) the Tetrahymena group I intron, (2) RNase P, (3) the hammerhead ribozyme, (4) the hairpin ribozyme, and (5) the hepatitis delta virus ribozyme. They may catalyze self-cleavage as well as the cleavage of external substrates.

1. Group One Intron: The splicing reaction is self-contained; that is, the intron - with the help of associated proteins - splices itself out of the precursor RNA. the action is catalyzed by the RNA, only a single molecule of substrate is involved (unlike protein enzymes that repeatedly catalyze a reaction). **However, synthetic versions of Group I introns made in the laboratory can - *in vitro* - act repeatedly; that is, like true enzymes. The DNA of some Group I introns includes an open reading frame (ORF) that encodes a transposase-like protein that can make a copy of the intron and insert it elsewhere in the genome**

2. Ribonuclease P: All living things synthesize an enzyme - called Ribonuclease P - that cleaves the head (5′) end of the precursors of transfer RNA (tRNA) molecules. Ribonuclease P is a heterodimer containing a molecule of RNA and one protein. When the RNA is separated from the protein, the RNA retains its ability to catalyze the cleavage step (although less efficiently than the intact dimer), but the protein alone cannot do the job.

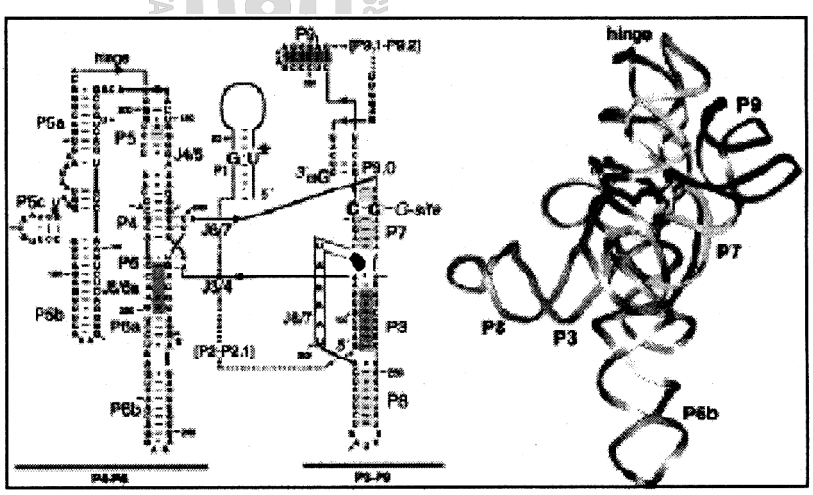


Fig. 6.8: Structure of Ribonuclease P

3. **Hammerhead ribozyme:** The name of hammerhead ribozyme is given by the similarity between its secondary structure and the shape of a hammerhead. They are the best understood subcategory of all ribozymes. As well as other

ribozymes, the hammerhead ribozyme is an antisense RNA. Some of the ribonucleotides within the sequence selectively form Watson-Crick base pairs with others to form a stem, while the rest stay in single stranded state called loop. These loops and stems can be predicted at the secondary structure level using conformational energy analysis, such as RNAdraw and mfold; and three dimensional structures were obtained mainly by X-ray crystallography

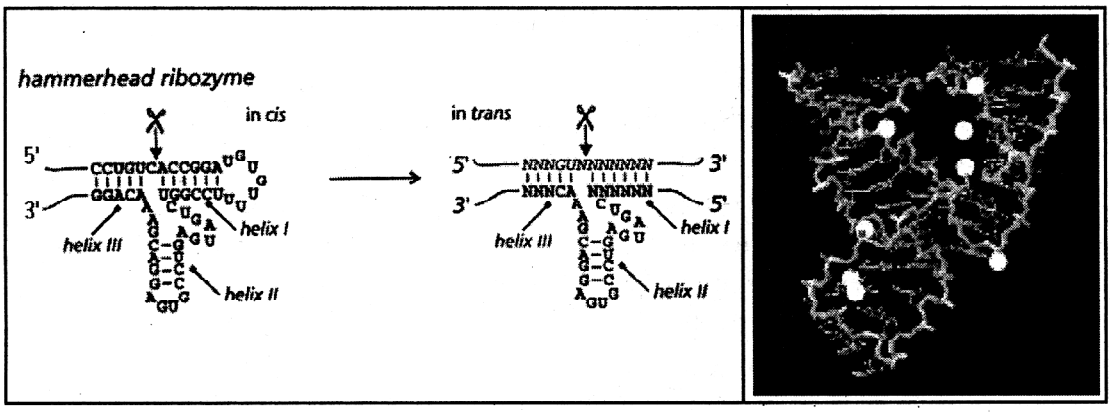


Fig. 6.9: Hammerheaded ribozyme

4. The hepatitis delta virus ribozyme : Some RNA viruses, such as the hepatitis delta virus, also include a ribozyme as part of their inherited RNA molecule. During replication of the viral RNA, long strands containing repeats of the RNA genome (viral genetic information) are synthesized. The ribozyme then cleaves the long multimeric molecules into pieces that contain one genome copy, and fits that RNA piece into a virus particle.

Since the discovery of ribozymes that exist in living organisms, there has been interest in the study of new synthetic ribozymes made in the laboratory. Ribozymes can be produced in the laboratory which, are capable of catalyzing their own synthesis under very specific conditions, such as an RNA polymerase ribozyme; although the polymerase activity is very limited. Such RNA polymerase ribozymes are able to add up to 14 nucleotides to a primer template in 24 hours until it is decomposed by hydrolysis of the phosphodiester bonds. Another artificially produced self cleaving RNAs ribozyme was produced by Tang and Breaker by in vitro selection of RNAs originating from random-sequence RNAs. Some of the synthetic ribozymes that were produced had novel structures, while some were similar to the naturally occurring hammerhead ribozyme.

The techniques used to discover synthetic ribozymes involve Darwinian evolution. This approach takes advantage of RNA's dual nature as both a catalyst and an informational polymer, making it easy for an investigator to produce vast

populations of RNA catalysts using polymerase enzymes. The ribozymes are mutated by reverse transcribing them with reverse transcriptase into various cDNA and amplified with mutagenic PCR. The selection parameters in these experiments often differ. One approach for selecting a ligase ribozyme involves using biotin tags, which are covalently linked to the substrate. If a molecule possesses the desired ligase activity, a streptavidin matrix can be used to recover the active molecules.

6.3.2 Ribozymes for human therapy

The application of ribozymes for gene therapy of autosomal dominant diseases has become popularized in recent years. Further this technology has widespread utility in the treatment of any disease, acquired or inherited, by inhibition of gene expression. The design of ribozymes is usually accomplished using computer assisted design programs, however they are not very useful in predicting the behavior of the ribozyme in the in vivo setting. To overcome this technical challenge, methods and strategy are being evolved to accurately assess the efficiency of ribozyme cleavage in vivo situations that significantly enhances the computer based design programs.

Already, a synthetic ribozyme that destroys the **mRNA** encoding a receptor of Vascular Endothelial Growth Factor (**VEGF**) is being readied for clinical trials. VEGF is a major stimulant of angiogenesis, and blocking its action may help starve cancers of their blood supply.

6.4 Antisense technology

6.4.1 Introduction

Over the last several decades, the knowledge of DNA/RNA physiology has been applied in a variety of ways. One of the more productive applications is the development of antisense technology. The basic idea is that if an oligonucleotide (a short RNA or DNA molecule complementary to a mRNA produced by a gene) can be introduced into a cell, it will specifically bind to its target mRNA through the exquisite specificity of complementary-based pairing—the same mechanism which guarantees the fidelity of DNA replication and of RNA transcription from the gene. This binding forms an RNA dimer in the cytoplasm and halts protein synthesis. This occurs because the mRNA no longer has access to the ribosome and because dimeric RNA is rapidly degraded in the cytoplasm by ribonuclease H. Therefore, the introduction of short chains of DNA complementary to mRNA will lead to a specific diminution, or blockage, of protein synthesis by a particular gene. In effect, the gene will be turned off.

The technical problems associated with the use of this technology are many. First, sufficient amounts of antisense oligonucleotide must be administered to the vicinity of target cells and, more importantly, must be taken up by those cells. Second, the antisense oligonucleotide should, ideally, have a long enough half-life within the cell to successfully impair mRNA translation into protein over a significant period of time. Finally, the oligonucleotide must also be nontoxic and sufficiently specific so as not to interfere with other cellular functions. In many applications, these hurdles have been overcome and antisense technology has developed into a productive branch of biology.

These technical challenges can be overcome in various ways depending on the specific application at hand. Oligonucleotides can be mixed with a variety of lipids to form complexes that are more easily incorporated by cell membranes, facilitating the entry of associated oligonucleotides into the cells. A number of other techniques have also been developed to facilitate the uptake of oligonucleotides by cells. Chemical modification of the antisense oligonucleotides can render them more stacble in cells and blood by increasing their resistance to ribonuclease digestion. Also, complementary DNAs or fragments of complementary DNAs can be incorporated in reverse sense in order to generate antisense RNA products in the host cell itself. This results in a long-term inhibition of the synthesis of the target protein.

6.4.2 Application of antisense technology *in vitro*

Antisense technology has been applied successfully in two general areas. The first is in fundamental research where the introduction of antisense oligonucleotides can help determine the role of a specific gene in a specific physiological process. For example, introduction of antisense oligonucleotides to inhibit the synthesis of angiotensinogen, the substrate from which cells make angiotensin II actually was found to stop the synthesis of angiotensinogen that resulted in a decline in cell growth. The introduction of angiotensin II to the cells restored this growth.

6.4.3 Therapeutic application of antisense technology

A second application of this technology, and one that is potentially of more immediate relevance to the practicing physician, is the use of this technology in therapy. In principle, antisense oligonucleotides complementary to viral RNAs can suppress a wide variety of viral infections; a tremendous amount of research is ongoing in this area. Similarly, antisense oligonucleotides directed towards the products of oncogenes can play a role in reducing the growth of cancer cells, and this lead is being hotly pursued.

Perhaps the most widely discussed application of antisense technology lies in its applications to gene therapy. In this case, a variety of vectors is used to introduce antisense-encoding genes into a large number of cells in a patient or animal to produce long-term inhibition of a protein. For example, in animal models the introduction of vectors encoding antisense angiotensin II receptor sequences results in long-term normotension in spontaneously hypertensive animals.

These are but a few of the possible applications of antisense technology. As familiarity with the relevant chemistry increases, it is likely that more effective oligonucleotides and gene vectors will be developed, thereby providing the ability to interfere at will with the translation of specific mRNAs.

6.4.4 Triplex antisense technology

In the face of all this progress, still newer technologies are being developed based on concepts related to antisense biology For example, it is known that oligonucleotides can, in certain instances, bind to duplex DNA molecules through an unusual kind of base pairing. In this triplex binding mode, oligonucleotides insert themselves into the major groove of the DNA double helix on a reasonably specific basis determined by the nucleotide sequence of the target DNA. This triplex technology provides the opportunity to reduce gene transcription jtself rather than to destroy mRNA once it is produced. Because the triplex oligonucleotides can be made to permanently alter the DNA after localizing to specific target sites, the technology actually has the potential to permanently silence genes.

6.4.5 RNA Inhibition

It has recently been shown that double-stranded RNA in the cytoplasm triggers an as yet poorly understood cascade of events leading to the suppression of the transcription of the gene producing the specific mRNA involved in the cytoplasmic RNA duplex. This could potentially lead to the development of new pharmacological agents.

Antisense technology is a formidable tool for investigating physiologic and pathologic processes. In addition, it is soon likely to become a mainstay of therapy, particularly in infectious diseases, with wider applications in the future as gene therapy techniques are developed further. Antisense Pharmaceuticals will soon be available for the routine care of patients and are expected to .prove to be effective, specific agents with favorable therapeutic profiles.

Unit 7 Recombination and Repair

Structure

7.1 Holiday junction in recombination

Recombination is a process or set of processes by which DNA molecules interact with one another to bring about a rearrangement of the genetic information or content in an organism. Although recombination as a process has been known for a hundred years, the real reason has not been appreciated until relatively recently. It now seems clear that recombination reactions exist to repair DNA. In bacteria, and probably also in eukaryotes, recombination mechanisms exist to repair stalled DNA replication forks. In the simplest sense, recombination is an exchange of both strands between two DNA molecules

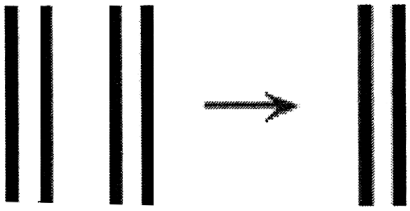


Fig. 7.1: Exchange of DNA strands between two homologous chromosomes.

In eukaryotic systems, you will be familiar with recombination as the process that is responsible for **crossing-over** during meiosis. Crossing-over has been well-documented genetically and is used to map the relative locations of genes on a chromosome.

7.2 The holiday model of genetic recombination

The model for recombination of two individual DNA strand was first proposed by Robin Holliday in 1964 and re-established by David Dressier and Huntington Potter in 1976 who demonstrated that the proposed physical intermediates existed.

In the most simplified explanation, two homologous DNA molecules align themselves which is followed by a nick at the same place on the two molecules as

shown in the figure below. This must happen in strands with the same polarity. The nicked strands then exchange themselves. The intermediate structure that is formed during such exchange is called a Holliday intermediate or Holliday structure. The shape of this intermediate in vivo is similar *to* that of the greek letter chi, hence this is also called a chi form.

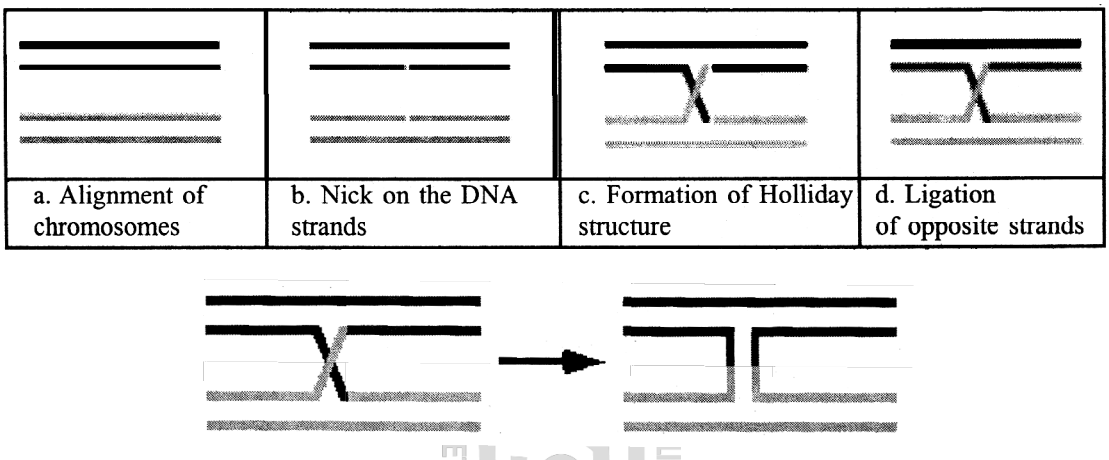


Fig. 7.2 : Initiation of recombination process and formation of Holiday structure

There are two ways in which the holiday structure can resolve itself to return back to its original conformation after the recombination process. If the same strands are cleaved a second time then the original two DNA molecules are generated (Fig. 7.3a). But, if the other strands are cleaved, then **recombinant** molecules are generated in a manner as shown in the figure below (Fig. 7.3b).

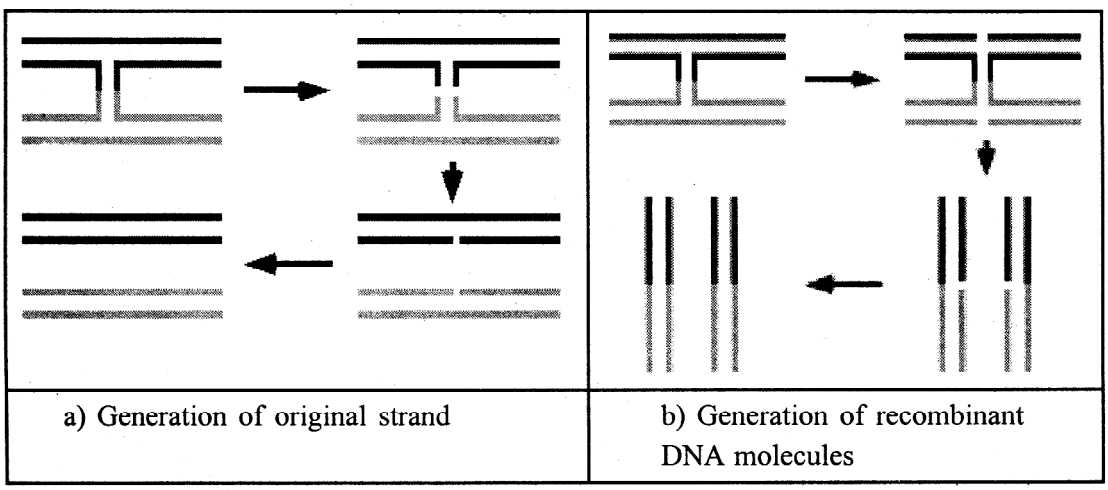


Fig. 7.3 : Two alternatives of recombination after the formation of holiday structure

In reality, a more complex mechanism operate to obtain different types of recom-bined strands as was observed in the Meselson-Weigle experiment where they located two different recombinant bacteriophages in a single plaque. These can be explained by modifying the above model slightly. As before, two homologous DNA molecules must be aligned and nicked at the same place. Following strand exchange the intermediate Holliday structure is formed. After the formation of the Holliday structure and ligation of the strands, the branch migrates, which can take place in either direction. The result is a physical transfer of part of one of the strands of one molecule with that of the other:

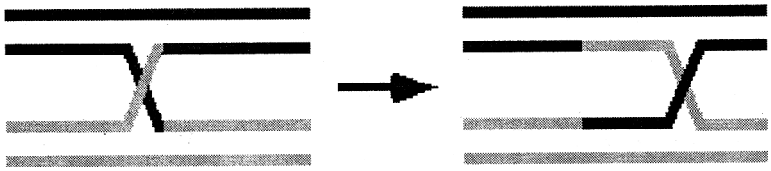


Fig. 7.4 : Two alternatives of recombination after the formation of holliday structure

For better understanding of the subsequent steps, one molecule is now rotated through 180° with respect to the other (Fig. 7.4).

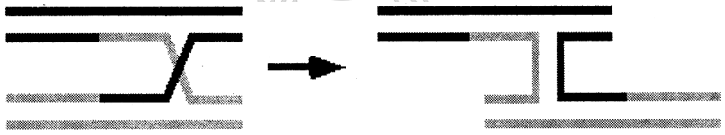
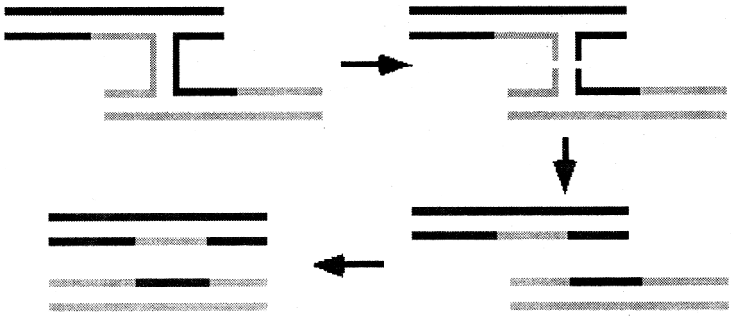


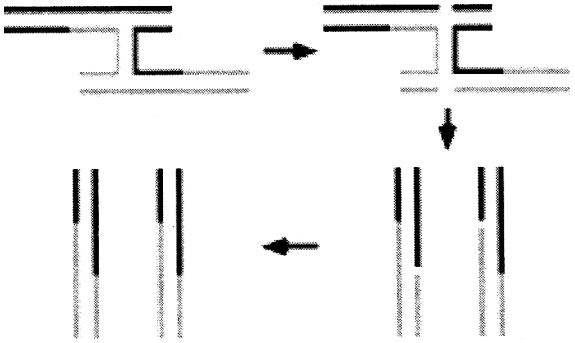
Fig. 7.5 : Rotation of one strand through 180° for better understanding

As described above, there are two possibilities of recombination which may result in two different consequences.

1. If the same strands are cleaved a second time then n on recombinant DNA molecules are generated but they each contain a region of heteroduplex DNA that spans the region of branch migration:



2. If the other strands are cleaved, then **recombinant** molecules are generated as before, however, each will also contain a region of **heteroduplex** DNA that spans the region of branch migration:



Double stranded nicks :

While the single-strand nick model provides a simple explanation for recombination, recent work in the yeast Sacchromyces cerevisiae shows that recombination is actually a response to Double-stranded breaks in the DNA molecule. The double-strand break model begins with the introduction of a double stranded break in one of the paired homologs (Fig. 7.6).

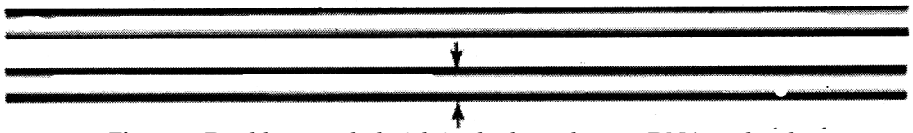


Fig. 7.6 : Double-stranded nick in the homologous DNA molecule

Introduction of the double-stranded break results in exonucleolytic degradation of the adjacent strand in a 5' to 3′ direction resulting in two single-stranded.whiskers (Fig. 7.7) which can now invade the paired homolog (via the action of a RecA like protein).

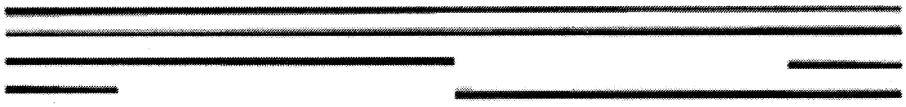


Fig. 7.7.: Exonucleolytic degradation of the adjacent strand in a 5' to 3′ direction

The DNA strand which was undisturbed serve as template and the invading 3' ends serve as primers for DNA synthesis and extends it self as shown below (Fig. 7.8).

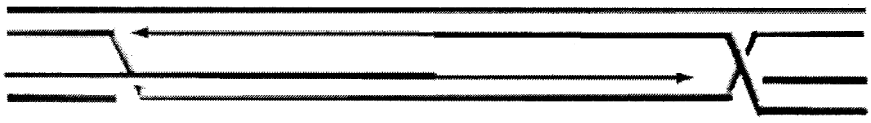


Fig. 7.8 : Extension of the overhands of the nicked strand

The 3' ends of the newly synthesized strand are then ligated with 5' ends of the degraded red homolog to form the **Double Holiday Junction** shown below (Fig. 7.9)

Fig. 7.9 : Formation of Double Holiday Junction

These Holiday Junctions are free to migrate as before to generate the heteroduplex regions as shown below (Fig. 7.10)

Fig. 7.10 : Free migration of the Holiday Junction

As before, each Holiday Junction can be resolved in two ways...

Resolution I: involves breaking and rejoining the two strands that cross between the two homologs.

Resolution II involves breaking and rejoining the two strands that do not cross between the two homologs

Since there are TWO Holiday Junctions, the exchange of flanking markers depends on how each Holiday Junction is resolved.

If both Holiday Junctions resolve via Resolution I, no exchange of flanking markers is observed as shown below. Note that the region between the two junctions involves heteroduplex which can be corrected to produce Gene Conversion (Fig. 7.11).

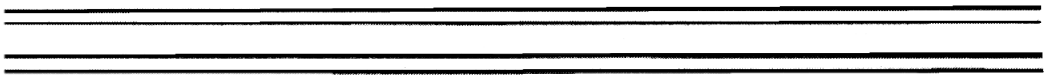


Fig. 7.11 : Formation of small region of heteroduplex if resolved through resolution 1

If Holiday Junction 1 undergoes Resolutation I and Holiday Junction 2 undergoes Resolutation II, or Holiday Junction 1 undergoes Resolution II and

Holiday Junction 2 undergoes Resolution I exchange of flanking markers is observed as shown below (Fig. 7.12).

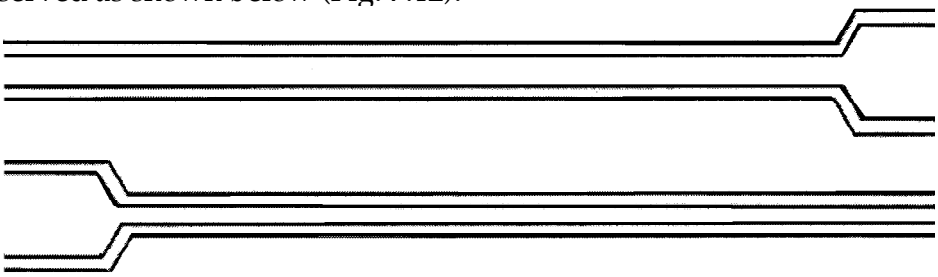


Fig. 7.12 : When both the Holliday junctions involve both type 1 and type 2 resolution, flanking markers are present

Finally, if both Holiday Junctions undergo Resolution II, no exchange of flankinj markers is observed.

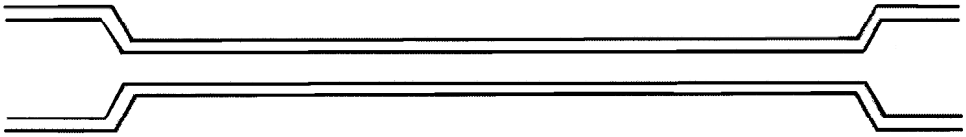


Fig. 7.13 : When both the Holliday junctions, undergo resolution 2 resolution no flanking markers is observed and the result is similar to fig. 11

7.3 Recombination proteins in *E. coli*

A number of the key proteins required for recombination. The structures and biochemical function of these proteins has been characterized which, is now helping us to understand details of the mechanism of recombination. The most important proteins are **RecA, RecBCD, RuvA, RuvB** and **RuvC.**

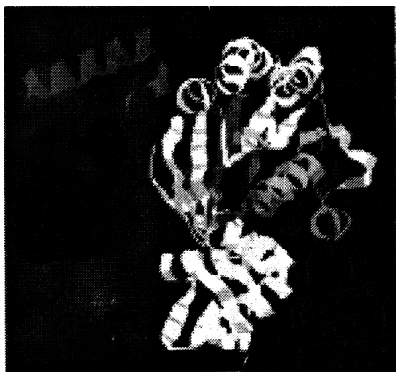


Fig. 7.14a :

RecA

The RecA protein is a multifunctional powerhouse! It has strand-exchange, ATPase and co-protease activities all packed into a compact 352 amino-acid, 38 kDa structure. It is required for all recombination pathways in *E. coli.* The RecA protein is a critical enzyme in this process, as it

catalyzes the pairing of ssDNA with complementary regions of dsDNA. The RecA monomers (Fig. 7.14a) first polymerize to form a helical filament around ssDNA, binding to a span of 4-6 nucleotides. Assembly of the nucleoprotein complex proceeds in a 5′→3′ direction. The complex is both fairly stable (half-life is 30 min) and is the active species that will promote strand exchange. RecA filament that forms is helical with a pitch of 82.7 Å and it consists of 6 monomer units per turn (Fig. 7.14b)

Fig. 7.14B :

During this process, RecA extends the ssDNA by 1.6 angstroms per axial base pair. Duplex DNA is then bound to the polymer. Bound dsDNA is partially unwound to facilitate base pairing between ssDNA and duplexed DNA. Once ssDNA has hybridized to a region of dsDNA, the duplexed DNA is further unwound to allow for branch migration. RecA has a binding site for ATP, the hydrolysis of which is required for release of the DNA strands from RecA filaments. ATP binding is also required for RecA-driven branch migration, but non-hydrolyzable analogs of ATP can be substituted for ATP in this process, suggesting that nucleotide binding alone can provide conformational changes in RecA filaments that promote branch migration. (Fig. 7.15)

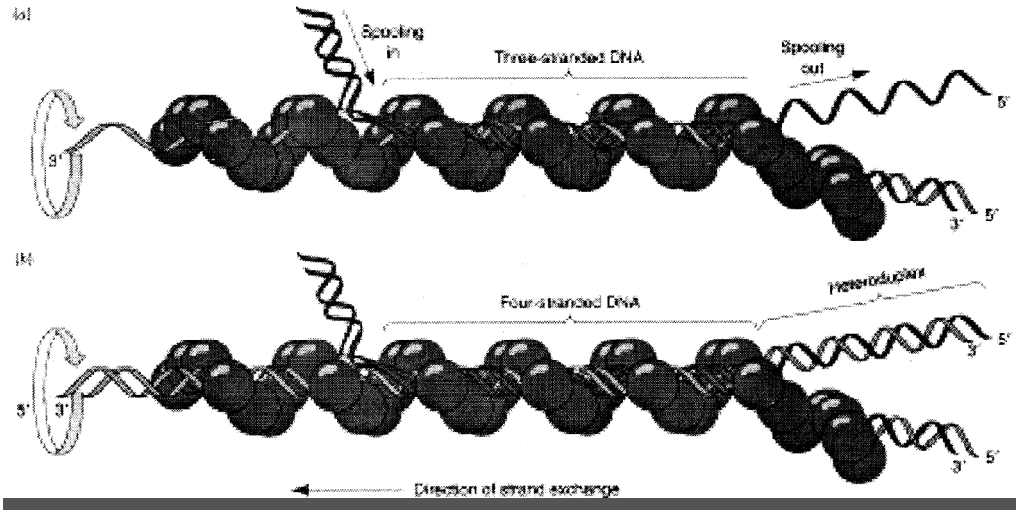


Fig. 7.15 : The strand exchange reaction probably involves the following steps: (a) RecA binds to the ssDNA partner, (b) The two molecules are aligned possible through the formation of a triple-stranded intermediate, (c) Displacement of one of the old strands. This requires concurrent migration of the RecA nucleoprotein filament along the molecule - which proceeds in one direction only (5′→3′) - and consequent winding/unwinding. ATP hydrolysis takes place during this step

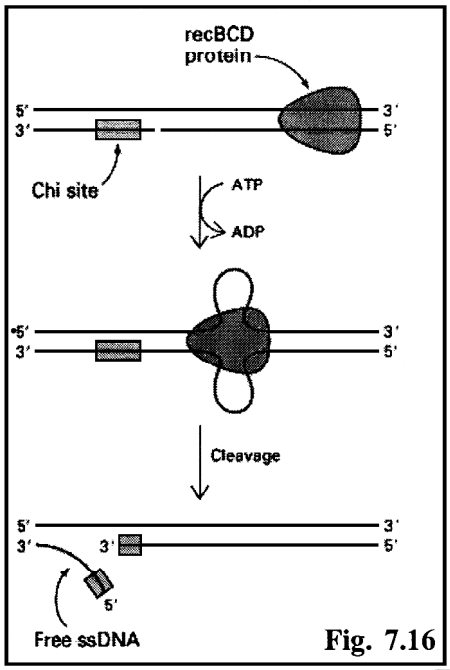


Fig. 7.16

RecBCD

The *recB, recC & recD* genes code for the three subunits of the RecBCD enzyme which has five activities: exonuclease V; a helicase activity; an endonuclease activity; an ATPase activity; and, an ssDNA exonuclease activity. The RecBCD helicase activity can unwind DNA faster than it rewinds. Thus as it travels along a DNA molecule, it can generate ssDNA loops (Fig. 7.16).

The RecBCD complex functions as a DNA exonuclease. It will bind to double-stranded breaks in DNA and degrade both strands simultaneously (Fig. 7.17). However, when **RecBCD** encounters a **Chi** sequence, its activity changes, the RecBC proteins act as a helicase to unwind the The RecD subunit is released and DNA in an ATP dependent reaction. This generates a ssDNA region that can serve (along with RecA) to initiate strand exchange and a recombination reaction

7.3.1 Proteins required for resolving holliday junctions in *E. coli*

RuvA

RuvA is a small protein whose function is to recognize a Holliday junction thereby assisting the **RuvB** helicase to promote branch migration.

The RuvA protein is 203 amino acids in length, but only 190 of them could be ssigned in the crystal structure. Most of the missing assignments represent amino

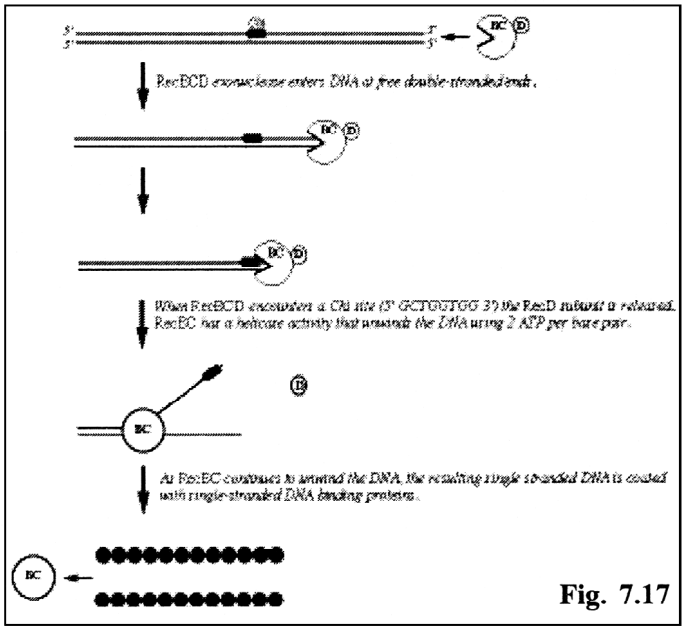
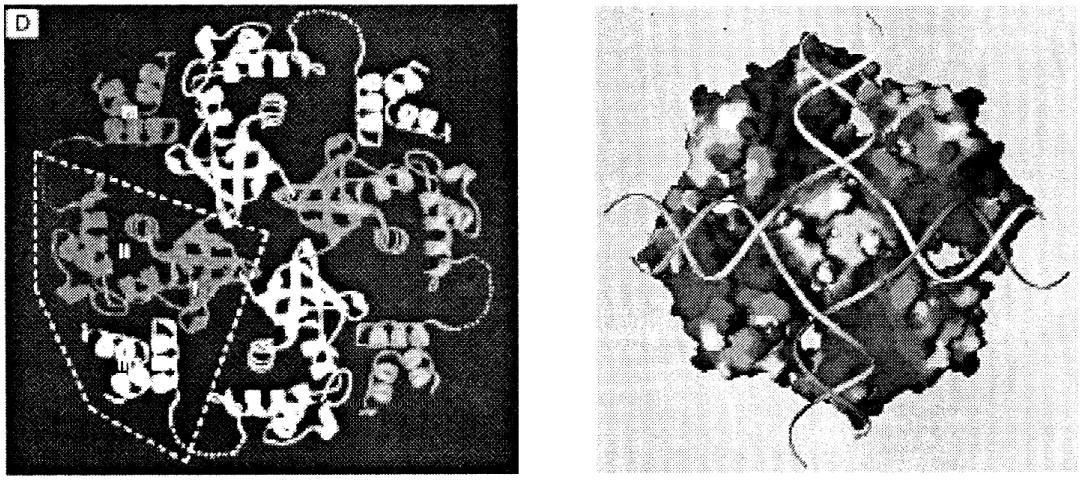


Fig. 7.17

acids in a flexible part of the protein. The crystal structure of the *E. coli* **RuvA** protein was solved at a resolution of 1.9 Å. The protein forms a tetramer in an unusual manner - though one that is ideally suited to its function.



the DNA strands at Holliday Junction

RuvB

The **RuvB** protein is a **helicase** that catalyzes branch migration of Holliday junctions. By itself it cannot bind to DNA efficiently. It functions in combination with **RuvA.** Like other helicases, RuvB functions as a hexamer; but, unlike other helicases, RuvB encloses double-stranded DNA not ssDNA.

Electron microscopy has shown that RuvB is a heptamer in solution and that it converts to a hexamer ring when it binds to DNA. Electron microscopy has also shown that the two hexamer rings of RuvB lie contacting RuvA on the two opposite sides of a RuvAB-Holliday junction complex (Fig. 7.19).

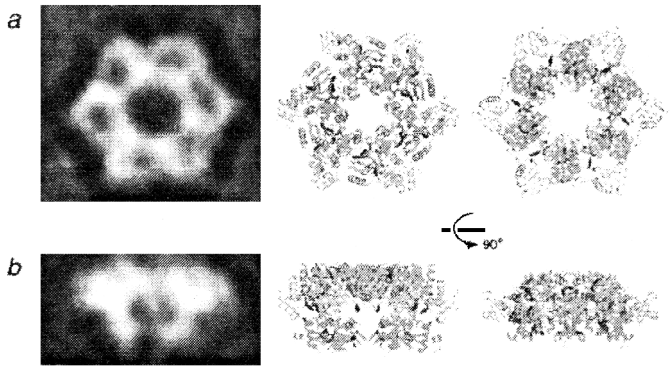


Fig. 7.19 : A hypothetical hexamer model of RuvB derived from the electron microscopy images

RuvC

The **RuvC** protein resolves the Holliday intermediate. It functions as a dimer to cleave two of the four strands that make up the central part of the intermediate. Since binding is symmetrical, **RuvC** can bind to the Holliday intermediate in two equally likely ways. Hence, Holliday intermediates can be resolved in two different, but equally likely, ways. The interaction of **RuvC** with Holliday junction is shownin Fig. 7.20.

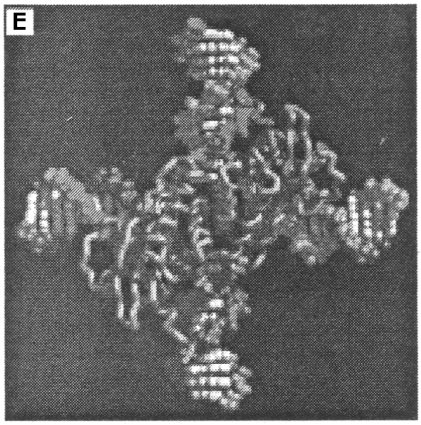


Fig. 7.20

RuvC does have some sequence specificity. It cleaves DNA at the 3′-side of thymidine, preferentially at the consensus 5′-$^{A}/_{T}$TT|$^{C}/_{G}$-3′ where '|' indicates the site of cleavage

7.4 DNA repair mechanisms

7.4.1 Introduction

Cells are always subjected to different types of stresses some of which cause alteration to the DNA molecules. For the genetic information encoded in the DNA is to remain uncorrupted, any chemical changes must be corrected. Surprisingly, both eukaryotic and prokaryotic cells have evolved an efficient DNA repair mechanism that function to maintain the integrity of the DNA molecules but allows subtle changes required to bring about variations and evolutionary changes. The recent publication of the human genome has already revealed 130 genes whose products participate in DNA repair. More will probably be identified soon.

7.4.2 Agents that damage DNA

- Certain wavelengths of **radiation**
  - ionizing radiation such as gamma rays and x-rays
  - **ultraviolet rays,** especially the UV-C rays (-260 nm) that are absorbed strongly by DNA but also the longer-wavelength UV-B that penetrates the ozone shield.
- Highly-reactive **oxygen radicals** produced during normal cellular respiration as well as by other biochemical pathways.
- Chemicals in the **environment**
  - many hydrocarbons, including some found in cigarette smoke
  - some plant and microbial products, e.g. the aflatoxins produced in moldy peanuts
- Chemicals used in **chemotherapy,** especially chemotherapy of cancers

7.4.3 Types of DNA damage

1. All four of the bases in DNA (A, T, C, G) can be covalently modified at various positions.

   o One of the most frequent is the loss of an amino group ("deamination") — resulting, for example, in a **C** being converted to a **U**.

2. During DNA replication DNA polymerase *(E. coli)* - inserts one incorrect nucleotide for every 10^5 nucleotides due to tautomeric "flickering" of the bases

3. **Mutations;**

   - Base Substitution by
     - **o point mutations**
       - *transitions*
         - pyrimidine-to-pyrimidine substitutions (T6C)
         - purine-to-purine substitutions (A6G) ?
       - *transversions*
         - pyrimidine-to-purine substitutions (T6G or A)
         - purine-to-pyrimidine substitutions (A6C or T)
   - Insertions or deletions of a small number of nucleotides
     - o **frame shift mutations**
       - recombination errors
       - transposons

Spontaneous rate of mutation at a given site on a chromosome is approximately 10^{-6} to 10^{-11} per round of replication. It is

- species and site specific
- "hot spots"
  - DNA microsatellites
    - repetitive sequences errors dues to "slippage" of DNA polymerase during replication
    - low frequency sites

4. **Mismatches** of the normal bases because of a failure of proofreading during DNA replication.
   - o example: incorporation of the pyrimidine **U** (normally found only in RNA) instead of **T.**

5. **Breaks** in the backbone : caused frequently by Ionizing radiation or by chemicals
   - o Can be limited to one of the two strands (a single-stranded break, **SSB)** or
   - o on **both strands** (a double-stranded break **(DSB).**

6. **Crosslinks** Covalent linkages can be formed between bases
   - o on the same DNA strand ("intrastrand") or
   - o on the opposite strand ("interstrand").

   Several chemotherapeutic drugs used against cancers crosslink DNA.

7.4.4 Repairing damaged bases

Damaged or inappropriate bases can be repaired by several mechanisms:

- **Direct chemical reversal** of the damage
- **Excision Repair,** in which the damaged base or bases are removed and then replaced with the correct ones in a localized burst of DNA synthesis. There are three modes of excision repair, each of which employs specialized sets of enzymes.
  1. **Base Excision Repair (BER)**
  2. **Nucleotide Excision Repair (NER)**
  3. **Mismatch Repair (MMR)**

7.4.5 Direct reversal of base damage

Perhaps the most frequent cause of point mutations in humans is the spontaneous addition of a methyl group (CH_3-) (an example of alkylation) to Cytosine followed by deamination to a Thyamine. Fortunately, most of these changes are repaired by enzymes, called glycosylases, that remove the mismatched T restoring the correct C. This is done without the need to break the DNA backbone (in contrast to the mechanisms of excision repair described below).

Some of the drugs used in cancer chemotherapy ("chemo") also damage DNA by alkylation. Some of the methyl groups can be removed by a protein encoded by our ***MGMT*** gene. However, the protein can only do it once, so the removal of each methyl group requires another molecule of protein.

This illustrates a problem with direct reversal mechanisms of DNA repair: they are quite wasteful. Each of the myriad types of chemical alterations to bases requires its own mechanism to correct. What the cell needs are more general mechanisms capable of correcting all sorts of chemical damage with a limited toolbox. This requirement is met by the mechanisms of **excision repair.**

7.4.6 Base excision repair (BER)

The steps and some key players:

1. Removal of the damaged base (estimated to occur some 20,000 times a day in each cell in our body!) by a DNA glycosylase. There exist at least 8 genes encoding different DNA glycosylases each enzyme responsible for identifying and removing a specific kind of base damage.
2. Removal of its deoxyribose phosphate in the backbone, producing a gap. We have two genes encoding enzymes with this function.
3. Replacement with the correct nucleotide. This relies on **DNA polymerase beta,** one of at least 11 DNA polymerases encoded by our genes.
4. Ligation of the break in the strand. Two enzymes are known that can do this; both require ATP to provide the needed energy.

7.4.7 Nucleotide excision repair (NER)

NER differs from BER in several ways.

- It uses different enzymes.
- Even though there may be only a single "bad" base to correct, its nucleotide is removed along with many other adjacent nucleotides; that is, NER removes a large "patch" around the damage.

The steps and some key players:

1. The damage is recognized by one or more protein factors that assemble at the location.
2. The DNA is unwound producing a "bubble". The enzyme system that does this is **Transcription Factor IIH, TFIIH,** (which also functions in normal transcription).

3. Cuts are made on both the 3' side and the 5' side of the damaged area so the tract containing the damage can be removed.
4. A fresh burst of DNA synthesis — using the intact (opposite) strand as a template — fills in the correct nucleotides. The DNA polymerases responsible are designated polymerase **delta** and **epsilon.**
5. **A DNA ligase** covalent binds the fresh piece into the backbone.

7.4.8 Transcription-coupled NER

Nucleotide-excision repair proceeds most rapidly

- in cells whose genes are being actively transcribed
- on the DNA strand that is serving as the template for transcription.

This enhancement of NER involves XPB, XPD, and several other gene products. The genes for two of them are designated ***CSA*** and ***CSB*** (mutations in them cause an inherited disorder called **Cockayne's syndrome).** The CSB product associates in the nucleus with **RNA polymerase II,** the enzyme responsible for synthesizing **messenger RNA** (mRNA), providing a molecular link between transcription and repair. One plausible scenario: If RNA polymerase II, tracking along the template (antisense) strand), encounters a damaged base, it can recruit other proteins, e.g., the CSA and CSB proteins, to make a quick fix before it moves on to complete transcription of the gene.

7.4.9 Mismatch repair (MMR)

Mismatch repair deals with correcting mismatches of the **normal bases;** that escapes proofreading mechanism of correction, and fails to maintain normal Watson-Crick base pairing (A•T, C•G)

It can enlist the aid of enzymes involved in both base-excision repair (BER) and nucleotide-excision repair (NER) as well as using enzymes specialized for this function.

- **Recognition** of a mismatch requires several different proteins including one encoded by ***MSH2.***

Cutting the mismatch out also requires several proteins, including one encoded by ***MLH1.***

Eg. In *E. coll*

- **MutS** dimer scans DNA for mismatches which distort DNA backbone
- Binds ATP at conformation changed sites where MutS bends the DNA
- ATP-MutS complex recruits **MutL & MutH.** ATP hydrolysis required for loading
- MutL activates **MutH endonuclease** activity nicks one strand near the mismatch
- DNA is unwound by **helicase UvrD** from the incision to the site of the mismatch
- **Exonuclease** digests displaced strand gap filled in by **Pol III** nick sealed by **ligase**

Mutations in either of these genes predispose the person to an inherited form of colon cancer. So these genes qualify as tumor suppressor genes. Synthesis of the repair patch is done by the same enzymes used in NER: **DNA polymerase delta** and **epsilon.** Cells also use the MMR system to enhance the fidelity of recombination; i.e., assure that only homologous regions of two DNA molecules pair up to crossover and recombine segments (e.g., in meiosis).

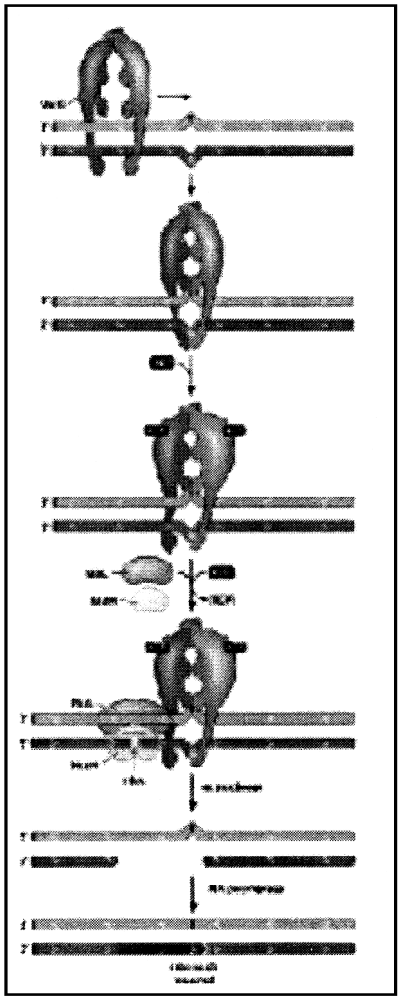


Fig. 7.21

Repairing Strand Breaks

Ionizing radiation and certain chemicals can produce both single-strand breaks **(SSBs)** and double-strand breaks **(DSBs)** in the DNA backbone.

Single-Strand Breaks (SSBs)

Breaks in a single strand of the DNA molecule are repaired using the same enzyme systems that are used in Base-Excision Repair (BER).

Double-Strand Breaks (DSBs)

There are two mechanisms by which the cell attempts to repair a complete break in a DNA molecule:

- **Direct joining** of the broken ends. This requires proteins that recognize and bind to the exposed ends and bring them together for ligating. They

would prefer to see some complementary nucleotides but can proceed without them so this type of joining is also called **Nonhomologous End-Joining (NHEJ).**

- Errors in direct joining may be a cause of the various **translocations** that are associated with cancers.
- Examples:
  - Burkitt's lymphoma
  - the Philadelphia chromosome in chronic myelogenous leukemia (CML)
  - B-cell leukemia

Meiosis I with the alignment of homologous sequences provides a mechanism for repairing damaged DNA; that is, mutations, in fact, many biologists feel that the main function of sex is to provide this mechanism for maintaining the integrity of the genome. However, most of the genes on the human Y chromosome have no counterpart on the X chromosome, and thus cannot benefit from this repair mechanism. They seem to solve this problem by having multiple copies of the same gene ––oriented in opposite directions. Looping the intervening DNA brings the duplicates together and allowing repair by homologous recombination.

7.4.10 Gene conversion

If the sequence used as a template for repairing a gene by homologous recombination differs slightly from the gene needing repair; that is, is an allele, the repaired gene will acquire the donor sequence. This **nonreciprocal transfer** of genetic information is called gene conversion. Gene conversion during meiosis alters the normal mendelian ratios. Normally, meiosis in a heterozygous **(A,a)** parent will produce gametes or spores in a 1:1 ratio; e.g., 50% **A;** 50% **a.** However, if gene conversion has occurred, other ratios will appear. If, for example, an **A** allele donates its sequence as it repairs a damaged **a** allele, the repaired gene will become **A**, and the ratio will be 75% **A;** 25% **a.**

Human diseases caused by loss of DNA repair systems: DNA repair systems play a major role in normal human health. Two examples of human pathology caused by loss of repair systems are described below.

Xeroderma pigmentosum: Xeroderma pigmentosum is a human genetic disease (or more correctly, a family of closely related genetic diseases), in which there is abnormal sensitivity to ultraviolet radiation. A number of different genes appear to be involved. Some patients exhibit defects in photoreactivation, but loss of

excision repair is more common. Mutations in at least seven different genes coding for proteins involved in excision repair can cause afflicted individuals to exhibit the symptoms of xeroderma pigmentosum.

Cockayne syndrome: This human genetic disease, whose symptoms include mental retardation, dwarfism, and premature aging, appears to be primarily due to failure of transcription-repair coupling.

Unit 8 Molecular Mapping of Genome

Structure

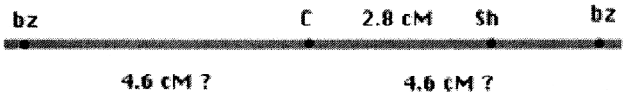
8.1 Genetic and physical maps

With the rebirth of genetics in the 20th century, it quickly became apparent that Mendel's second rule does not apply to many matings of dihybrids. In many cases, two alleles inherited from one parent show a strong tendency to stay together as do those from the other parent. This phenomenon is called **linkage.** The linkage phenomenon was utilized to construct gene maps.

8.1.1 Genetic map

A genetic map is a representation of the genes on a chromosome arrayed in linear order with distances between loci expressed as percent recombination **(map units, centimorgans).** It is also called a **linkage map.** The relative position of genes on a chromosome in a Genetic map is determined by counting the phenotypes from a cross. One **map unit =** one **centimorgan (cM)** = 1% recombination between loci. The farther apart two loci are, the more likely that a crossover will occur between them. Conversely, if two loci are close together, a crossover is less likely to occur between them. A recombination rate of 50% corresponds to independent assortment. Therefore, only distances less than 50 map units can be measured directly. Greater distances can be constructed by adding up distances between closer loci. They have been prepared for many eukaryotes, including corn, Drosophila, the mouse, and the tomato. A genetic map of chromosome 9 (the one that carries the **C, Sh,** and **bz** loci) of the corn plant (**Zea mays)** is shown on the right.

For example, test crossing a com plant that is dihybrid for the C,c (yellow/ colourless kernels) alleles and the alleles for bronze color (**Bz, bz**) produces 4.6% recombinants. So these two loci are 4.6 cM apart. However, is the **bz** locus on the same side of **c** as **sh** (smooth/shrunken) or is it on the other side? consider a cross between two different strains of corn (maize)



The answer can be found by test crossing the dihybrid **Shsh, Bzbz.** If the percentage of recombinants is less than 4.6%, then bz must be on the same side of locus c as locus **sh.** If greater than 4.6%, it must be on the other side. In fact, the recombination frequency is less than 1.8%, telling us that the actual order of loci is c — **sh** — **bz**

But there are certain difficulties with such genetic maps. Mapping by linkage analysis is best done with loci that are relatively close together; that is, within a few centimorgans of each other. Why? Because as the distance between two loci increases, the probability of a second crossover occurring between them also increases and therefore interpretation becomes difficult. There are other problems with preparing genetic maps of chromosomes.

- The probability of a crossover is not uniform along the entire length of the chromosome.
  - o Crossing over is inhibited in some regions (e.g., near the centromere).
  - o Some regions are "hot spots" for recombination (for reasons that are not clear). Approximately 80% of genetic recombination in humans is confined to just one-quarter of our genome.
- In humans, the frequency of recombination of loci on most chromosomes is higher in females than in males. Therefore, genetic maps of female chromosomes are longer than those for males.

8.1.2 Chromosome maps

The chromosome map (or cytogenetic map) is based on the karyotype of an organism. For example: All mouse chromosomes are defined at the cytogenetic

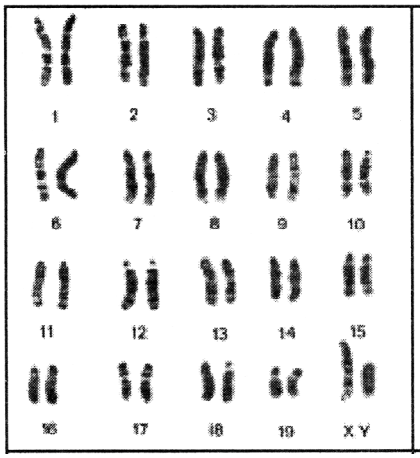


Fig. 8.1 : Karyotype of a diploid set of normal raetaphase chromosomes of *Mus musculus.*

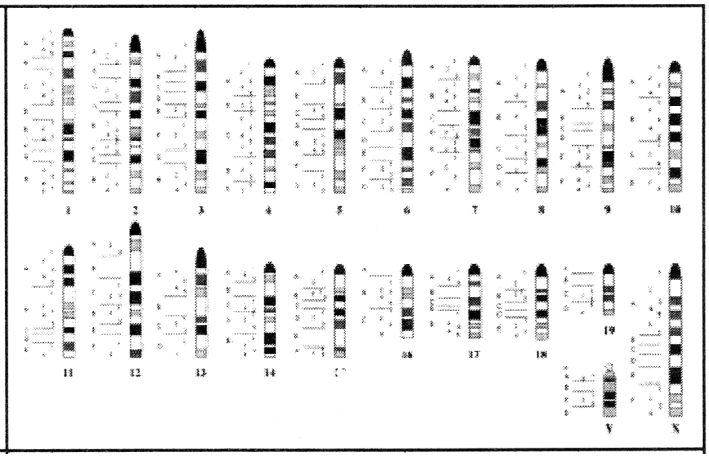


Fig 8.2 : Mouse chromosome idiograms. (Giemsa banding patterns associated with each chromosome in a normal karyotype

level according to their size and banding pattern (Fig. 8.1) and ultimately, all chromosomal assignments are made by direct cytogenetic analysis or by linkage to a locus that has previously been mapped in this way. Chromosomal map positions are indicated with the use of band names (Fig. 8.2).

Today, several different approaches, with different levels of resolution, can be used to generate chromosome maps.

1. Human/mouse cell hybrids tend to lose human chromosomes at random, leading eventually to hybrid cell lines that have one or a few human chromosomes. If a gene is always present or absent when one particular chromosome is present or absent, it can be concluded that the gene is on that chromosome.
2. Fluorescent or radioactive probes that bind to a particular gene can be observed microscopically and can be used to localize the gene on a metaphase spread.
3. Chromosomes from cells in metaphase can be sorted with high-speed electronic sorters. One can make preparations of a particular chromosome. If a particular gene can be shown to be in the preparation, it must be located on that chromosome.

8.1.3 Physical map

All physical maps are based on the direct analysis of DNA. Physical distances between and within loci are measured in basepairs (bp). Physical maps are arbitrarily divided into short range and long range. 1. Short range mapping is commonly pursued over distances ranging up to 30 kb. In very approximate terms, this is the average size of a gene and it is also the average size of cloned inserts obtained from cosmid-based genomic libraries. Cloned regions of this size can be easily mapped to high resolution with restriction enzymes or by sequencing.

2. Direct long-range physical mapping can be accomplished over megabase-sized regions with the use of rare-cutting restriction enzymes together with various methods of gel electrophoresis referred to generically as *pulsed field gel electrophoresis* or PFGE, which allow the separation and sizing of DNA fragments of 6 mb or more in length.

3. Long-range mapping can also be performed with clones obtained from large insert genomic libraries such as those based on the yeast artificial chromosome (YAC) cloning vectors, since regions within these clones can be readily isolated for further analysis.

8.1.4 Connections between maps

In theory, linkage, chromosomal, and physical maps should all provide the same information on chromosomal assignment and the order of loci (Fig. 8.3). However, the relative *distances* that are measured within each map can be quite different. Only the physical map can provide an accurate description of the actual length of DNA that separates loci from each other. This is not to say that the other two types of maps are inaccurate. Rather, each represents a version of the physical map that has been modulated according to a different parameter. Cytogenetic distances are modulated by the relative packing of the DNA molecule into different chromosomal regions. Linkage distances are modulated by the variable propensity of different DNA regions to take part in recombination events.

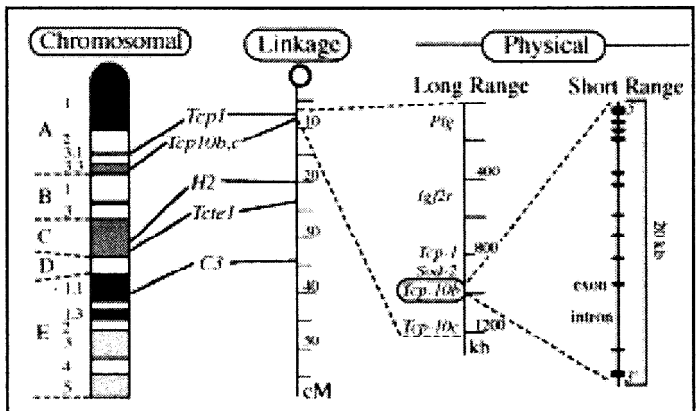


Fig. 8.3: Comparison of chromosomal, linkage, and physical maps. The three types of maps are shown for loci on Chr 17 of mouse.

In practice, genetic maps of the mouse are often an amalgamation of chromosomal, linkage, and physical maps, but at the time of this writing, it is still the case that classical recombination studies provide the great bulk of data incorporated into such integrated maps. Thus, the primary metric used to chart interlocus distances has been the centimorgan. However, it seems reasonable to predict that, within the next five years, the megabase will overtake the centimorgan as the unit for measurement along the chromosome

8.1.5 Gene mapping has important applications

A. It is useful for locating the position of genes on chromosomes, e.g. if two genes are closely linked and the position of one is known, then the other must also be nearby.

B. It is useful in estimating genetic risk, e.g. if a gene cannot be tested directly, then variation at a closely linked locus may indicate the presence or absence of a detrimental allele.

C. A major goal of the Human Genome Project is the mapping of all human

genes (as well as those of mice, Drosophila, *Caenorabditis elegans* (a nematode), *Arabidopsis thaliana* (a small plant), yeast, and the bacterium *Escherichia coli.* As of 1999, yeast, *E. coli, C. elegans,and* about a dozen other bacteria have been completely sequenced and all their genes identified, although the functions of most are unknown. Major progress has been made in mapping human genes, and a "rough draft" of the human genome is anticipated by 2000. Understanding of function of the many newly discovered human genes is being greatly aided by the studies of yeast, which has many genes similar to those of humans.

8.2 Gene cloning

8.2.1 Procedure of cloning

Gene cloning involves separating a specific gene or DNA segment from a larger chromosome, attaching it to a small molecule of carrier DNA, and then replicating this modified DNA thousands or millions of times through both an increase in cell number and the creation of multiple copies of the cloned DNA in each cell. The result is selective amplification of a particular gene or DNA segment. Cloning of DNA before the 1970 was a difficult tusk. Unlike a protein, a gene does not exist as a discrete entity in cells, but rather as a small region of a much larger DNA molecule. But with the discovery of restriction nucleases and other enzymes like ligase, poly-merases, the tusk of cloning targeted gene became rather easy. The key development that made recombinant DNA technology possible was the discovery in the late 1960s of restriction enzymes or called restriction endonucleases. The specialty of the endo-nucleases is that it recognizes and makes double-stranded cuts in the sugar-phosphate backbone of DNA molecules at specific nucleotide sequences. These enzymes are produced naturally by bacteria, where they are used in defense against viruses. In bacteria, restriction enzymes recognize particular sequences in viral DNA and then cut it up. A bacterium protects its own DNA from a restriction enzyme by modifying the recognition sequence, usually by adding methyl groups to its DNA. The endonucleases can be used to cut purified DNA at targeted sites for partially purifying the gene from a mixture.

Cloning of any DNA fragment essentially involves four steps: fragmentation, ligation, transfection, and screening/selection. Although these steps are invariable among cloning procedures a number of alternative routes can be selected, these are summarised as a 'cloning strategy'. The protocol for isolation of genes can be broken down into several steps.

Step 1.

At first, DNA need to be isolated from the desired cell, purified to ensure there is no protein contamination in the DNA sample. Presence of protein or any undesired chemicals can effect the subsequent steps of cloning.

Step 2.

DNA sample having the gene of interest requires to be segmented into suitable size. Preparation of DNA fragments for cloning is frequently achieved by means of PCR, but it may also be accomplished by restriction enzyme digestion, DNA soni-cation and fractionation by agarose gel electrophoresis. Restriction enzyme digestion, for example with EcoRl will produce small fragments of DNA with sticky ends. If there is no EcoRl site in the gene of interest, then a the DNA fragment carrying the gene will have flanking region with sticky ends.

Step 3.

The next step involves the insertion of the DNA fragment into a vector. A vector is usually a plasmid- circular DNA, which is linearised by means of restriction enzymes. The DNA fragments and the linearised vector is incubated together under appropriate condition in presence of the enzyme DNA ligase. Sticky ends or the single stranded DNA overhangs allow annealing of the DNA fragment with the vector sequence. Sticky ends may also be produced by chemical modification and attachment of adapter molecules. 'Sticky ends' allow for both higher efficiency transformations and directional insertion of the insert into the vector, thus minimising the need for subsequent screening.

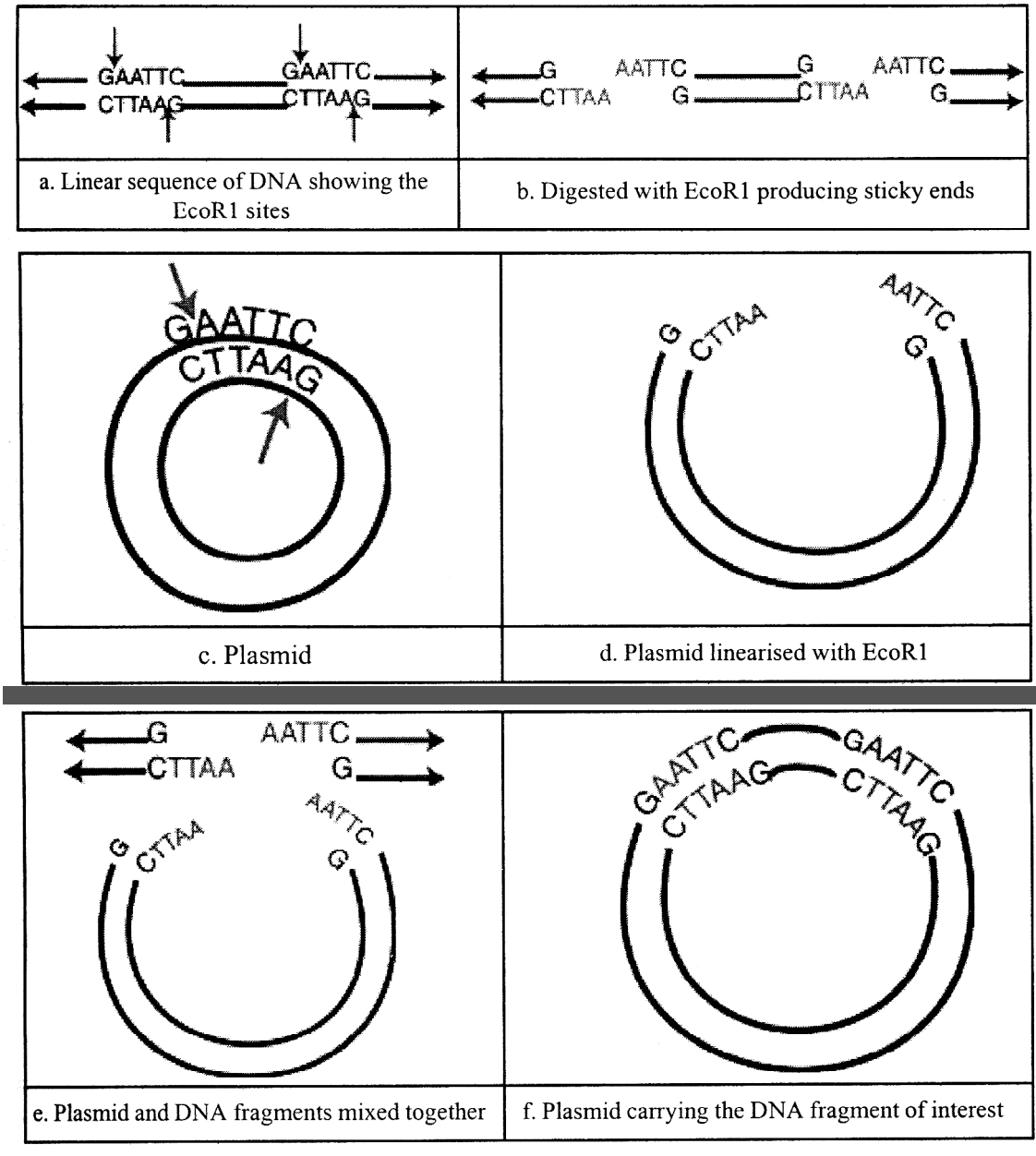
Step 4.

After the ligation precedure, the vectors with successful inserts are first identified and then is transfected into host cells. A number of alternative techniques are available, such as chemical sensitization of cells, electroporation and biolistics. Chemical sensitization of cells is frequently employed since this does not require specialised equipment and provides relatively high transformation efficiencies. Electroporation is employed when extremely high transformation efficiencies are required, as in very inefficient cloning strategies. Biolistics are mainly used in plant cell transformations, where the cell wall is a major obstacle in DNA uptake by cells.

Step 5.

Finally, the transfected cells are cultured and screened to identify the clone carrying the gene of interest. If the starting material is a PCR product, the screening step is not required. Successfully transformed carrying the gene of interest is identified primarily by hybridization technique. The required cells will be those that have been successfully transfected with the vector construct containing the

desired insertion sequence in the required orientation. Modern cloning vectors include selectable antibiotic resistance selection marker, which allow only cells in which the vector has been transfected, too grow. Additionally, the cloning vectors may contain colour selection markers which provide blue/white screening



a. Linear sequence of DNA showing the EcoR1 sites

b. Digested with EcoR1 producing sticky ends

c. Plasmid

d. Plasmid linearised with EcoR1

e. Plasmid and DNA fragments mixed together

f. Plasmid carrying the DNA fragment of interest

Fig. 8.4 : Schematic representation of the procedure of cloning

(á-factor complementation) on X-gal medium. Nevertheless, these selection steps do not absolutely guarantee that the DNA insert is present in the cells obtained. Further investigation of the resulting colonies is required to confirm that cloning was successful

8.2.2 cDNA Libraries

A collection of clones containing all the DNA fragments from one source is called a DNA library. For example, we might isolate genomic DNA from human cells, break it into fragments, and clone all of them in bacterial cells or phages. The set of bacterial colonies or phages containing these fragments is a human genomic library, containing all the DNA sequences found in the human genome. A genomic library must contain a large number of clones to ensure that all DNA sequences in the genome are represented in the library. A library of the human genome formed by using cosmids, each carrying a random DNA fragment from 35,000 to 44,000 bp long, would require about 350,000 cosmid clones to provide a 99% chance that every sequence is included in the library.

An alternative to creating a genomic library is to create a library consisting only of those DNA sequences that are transcribed into mRNA (called a cDNA library because all the DNA in this library is *complementary* to mRNA). Much of eukaryotic DNA consists of repetitive (and other DNA) sequences that are not transcribed into mRNA and such sequences are not represented in a cDNA library.

One of the most challenging tasks in molecular biology is the synthesis and cloning of cDNA. A complex series of enzymatic steps is involved in copying mRNA into double-stranded cDNA and subsequently preparing the termini for vector ligation. Many approaches have been used to generate cDNA libraries. Most cDNA molecules produced will lack a few nucleotides corresponding to the 5' end of the mRNA because second-strand replacement only proceeds from 3'-OH RNA primers. However, since all eukaryotic mRNA molecules appear to have 5' noncoding leader sequences, which commonly range from 40-80 nucleotides, it is likely that the vast majority of double-stranded cDNA will contain all of the coding sequences present in the initial cellular mRNA molecules.

However, cDNA library has two advantages. First, it is enriched with fragments from actively transcribed genes. Second, introns do not interrupt the cloned sequences and therefore easy to clone in bacterial system for expression.The disadvantage of a cDNA library is that it contains only sequences that are present in mature mRNA. Introns and any other sequences that influence transcription procedure are not present; sequences, such as promoters and enhancers, etc. It is

also important to note that the cDNA library represents only those gene sequences expressed in the tissue from which the RNA was isolated which is again dependent on the frequency of mRNA trans cribbed in the tissue. In contrast, almost all genes are present at the same frequency in a genomic DNA library.

8.2.3 Choice of primers

The classical method of cDNA synthesis uses the Oligo(dT) Primer to prime first-strand synthesis (Fig. 8.5). This method is suitable in most cases where poly(A)+ RNA of high quality can be prepared from the cell line or tissue of interest. Random Hexameric Primers (hexadeoxyribonucleotides) provide an alternative procedure (Fig. 8.6) by which first-strand cDNA synthesis is initiated from internal sites within the mRNA molecule. Random Primers can be used to prime cDNA synthesis from mRNA molecules that do not possess a poly(A)+ tail or for RNA isolated from prokaryotic sources. Random Primers also provide a scheme by which cDNA can be synthesized representing mRNA with strong 5′ secondary structure.

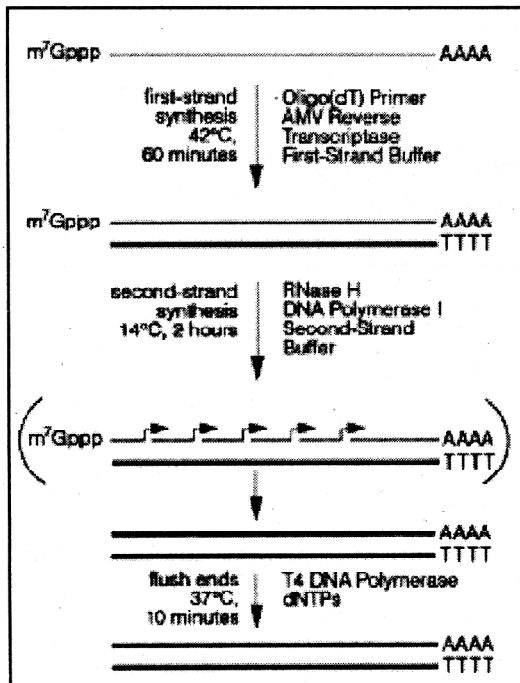


Fig.8.5: cDNA synthesis using the Oligo(dT) Primer

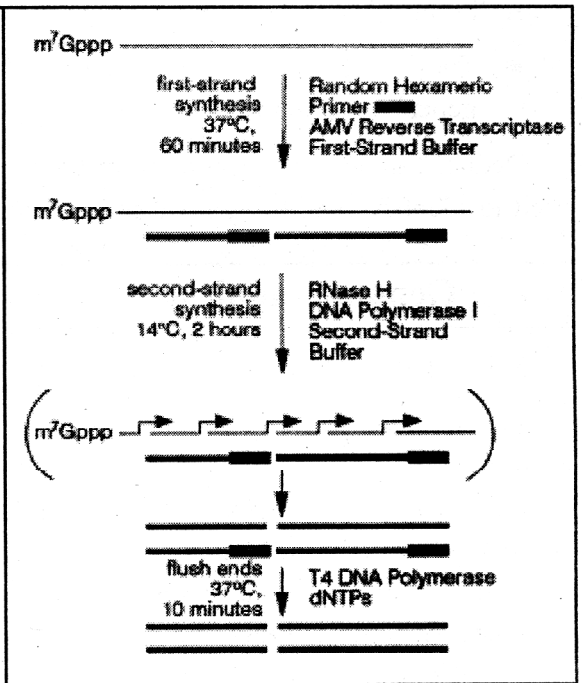


Fig. 8.6: cDNA synthesis using Random Hexameric Primers.

The mRNA molecules are then copied into cDNA by reverse transcription. Reverse transcriptase, an enzyme isolated from retroviruses, synthesizes single-stranded complementary DNA from the RNA template by adding DNA nucleotides to the 3'-OH group of the primer (Fig. 1 & 2)

The resulting RNA-DNA hybrid molecule is then converted into a double-stranded cDNA molecule by one of several methods. One common method is to treat the RNA-DNA hybrid with RNase to partly digest the RNA strand. Partial digestion leaves gaps in the RNA-DNA hybrid, allowing DNA polymerase to synthesize a second DNA strand by using the short undigested RNA pieces as primers and the first DNA strand as a template. DNA polymerase eventually displaces all the RNA fragments, replacing them with DNA nucleotides, and nicks n the sugar-phosphate backbone are sealed by DNA ligase. The cDNA thus synthesized are then subjected to end modifications and cloned into vectors for further analysis.

8.2.5 Libraries

A "library" is a convenient storage mechanism of genetic information.

- They are typically either "genomic" or "cDNA" (i.e. mRNA in DNA form) genetic information.

Deduced genetic sequences from corresponding polypeptide information can be used to identify specific genetic information within a library.

8.3 Genomic analysis

8.3.1 Introduction

The field of genomics comprises focuses on the content and organization of genomic information, and attempts to understand the function of information in genomes. Genomics is trying to look at all the genes as a dynamic system, over time, and determine how they interact and influence biological pathways and physiology, in a much more global sense. Genetics looks at single genes, one at a time, as a snapshot. Genetics is much more linear than genomics, complicated but not as complex as genomics.

The genetic information possessed by each individual is termed its **genotype** and can refer to the entirety of its genetic information or a part of it. The set of characteristics expressed by an individual is termed as **phenotype**. When two

individuals possessing the same genotype also have the same phenotype, regardless of the environmental conditions in which they exist, the character expressed is termed as **genetic trait.** Determination of the mode of inheritance of the genetic trait is called inheritance analysis. Inheritance analysis of a trait or phenotypic character is a **genetic marker.** If each phenotype can be unambiguously assigned to exactly one genotype, then the genetic marker defines a **gene marker.**

The advent of recombinant DNA technology in population genetics in the mid-1980's, gradually led to the development of **DNA markers.** However, the repertoire of genetic markers available for population genetic studies continues to increase enormously and is still relevant of genetic analysis. DNA marker analysis, though costly have some added advantage over the classical genetic markers and is rapidly replacing the old system of genetic analysis.

Technological advancement in the DNA sequencing technique contributed to development of the Genomics as a subject. Agenomic sequence is, by itself, of limited use. Functional genomics is, in essence, probing genome sequences for meaning— identifying genes, identify the unique sequences which can serve as DNA markers, recognizing their organization, and understanding their function etc. The goals of functional genomics include identifying all the RNA molecules transcribed from a genome (the **transcriptome**) and all the proteins encoded by the genome (the **proteome**).' Functional genomics exploits both bioinformatics and laboratory-based experimental approaches in its search to define the function of DNA sequences.

8.3.2 Predicting function from sequence

Several methods for identifying genes and assessing their functions have been discussed earlier. The methods include in situ hybridization, DNA footprinting, experimental mutagenesis, and the use of transgenic animals and knockouts. These methods can provide important information about the locations and functions of genetic information and can be applied to study to large numbers of genes simultaneously.

However, this biochemical approach to understanding gene function is both time consuming and expensive. A major goal of functional genomics has been to develop computational methods that allow gene function to be identified from DNA sequence alone, bypassing the laborious process of isolating and characterizing individual proteins.

8.3.3 Search for homology

One computational method for determining gene function is to conduct a homology search, which relies on comparing DNA and protein sequences from the same and different organisms. Databases containing sequences of genes and proteins for a wide array of organisms are available in gene banks which are in public domain and can be accessed for homology searches. Powerful computer programs have been developed for scanning these databases to look for particular sequences. A commonly used homology search program is BLAST used to align sequences from different or same species. If a function is known for one of these sequences, that function may provide information about the function of the newly discovered protein. Similar programs are also available that can analyze two sequences and predict the evolutionary relationship from which phylogenetic trees can be established.

8.3.4 Drug designing

Computer programs are also available that can detect single nucleotide polymorphism. One. can use the information from the analysis to predict the causes of various diseases and also can use the information to design drugs for treatment of diseases.

8.3.5 Gene expression and microarrays

The advent and developmet of the **Microarray** technique made it possible to study hundred and thousands of gene at the same time. Many important clues about gene function come from knowing when and where the genes are expressed. The microarray technique enables us to get such clues. **Microarrays** rely on nucleic acid hybridization in which a known DNA fragment is used as a probe to find complementary sequences (Fig. 8.7). In a microarray, numerous known DNA fragments are fixed to a solid support in an orderly pattern or array, usually as a series of dots. These DNA fragments (the probes) usually correspond to known genes. When the microarray has been constructed, mRNA, DNA, or cDNA isolated from experimental cells is labeled with fluorescent nucleotides and applied to the array. Any of the DNA or RNA molecules that are complementary to probes on the array will hybridize with them and emit luorescence, which can be detected by an automated scanner. An array containing tens of thousands of probes can be applied to a glass slide or silicon wafer just a few square centimeters in size.

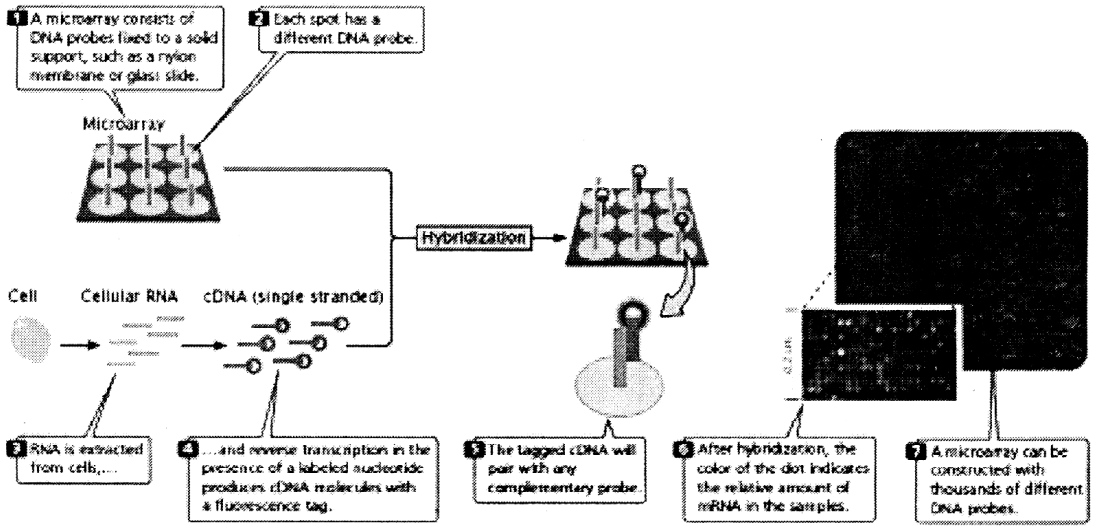


Fig 8.7 : Microarrays, used to detect the expression of many genes

For example, the experimental cells are stimulated from which mRNA is isolated and converted into cDNA labeled with red fluorescent nucleotides. mRNA from control cells is converted into cDNA and labeled with green fluorescent nucleotides. The labeled cDNAs are mixed and hybridized to the DNA chip, which contains DNA probes from different genes from the same organism. Hybridization of the red (experimental) and green (control) cDNAs is proportional to the relative amounts of mRNA in the samples. The fluorescence of each spot is assessed with microscopic scanning and appears as a single color. Red indicates the overexpression of a gene in the experimental cells relative to that in the control cells (more red-labeled cDNA hybridizes), whereas green indicates the underexpression of a gene in the experimental cells relative to that in the control cells (more green-labeled cDNA hybridizes). Yellow indicates equal expression in experimental and control cells (equal hybridization of red- and greenlabeled cDNAs), and no color indicates no expression in either experimental or control cells (Fig. 8.8).

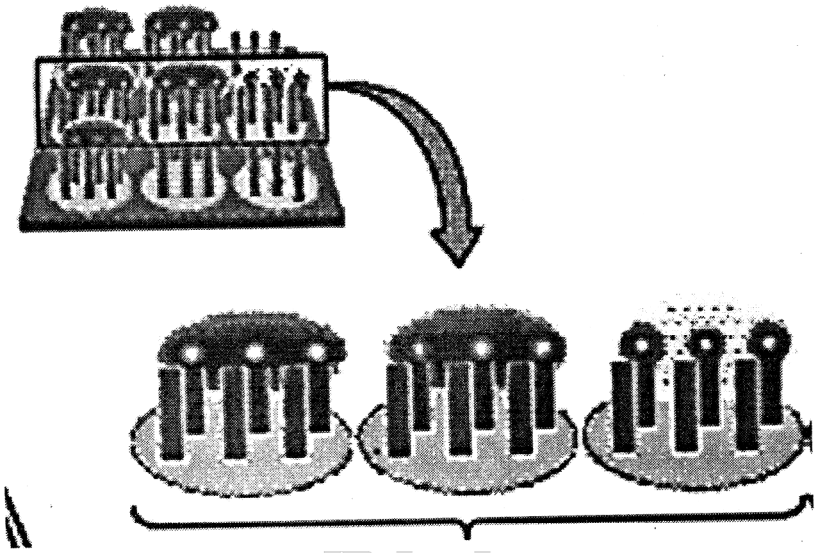


Fig: 8.8: Emission of fluorescence depends on the nature of hybridization

Microarrays allow the expression of thousands of genes to be monitored simultaneously, enabling scientists to study which genes are active in particular tissues. They can also be used to investigate how gene expression changes in the course of biological processes such as development or disease progression.

8.4 Restriction fragment length polymorphisms (RFLP)

A restriction fragment length polymorph of alternative alleles associated with restriction fragments that differ in size from each other. RFLPs are visualized by digesting DNA from different individuals with a restriction enzyme, followed by gel elec-trophoresis to separate fragments according to size, then blotting and hybridization to a labeled probe that identifies the locus under investigation. An RFLP is demonstrated whenever the Southern blot pattern obtained with one individual is different from the one obtained with another individual (Fig. 8.9). In this example, DNA samples from five individual mice were

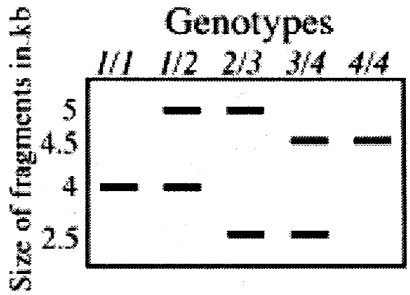


Fig. 8.9: Demonstration of RFLPs by Southern blot analysis.5 different Southern blot genotypes detected with a single probe is shown.

digested with the same enzyme, and after electrophoresis were probed with the same clone of a single-copy DNA sequence. The five patterns detected are all different from each other and are representative of five different genotypes. The first lane and the last lane are homozygous for the genes while the remaining individuals are heterozygous.

RPLPs were the predominant form of DNA variation used for linkage analysis until the advent of PCR. Even now, in the PCR age, RFLPs provide a convenient means for turning an uncharacterized DNA clone into a reagent for the detection of a genetic marker. The main advantage of RFLP analysis over PCR-based protocols is that no prior sequence information, nor oligonucleotide synthesis, is required. Furthermore, in some cases, it may not be feasible to develop a PCR protocol to detect a particular form of allelic variation. Nevertheless, if and when a PCR assay for typing a particular locus is developed, it will almost certainly be preferable over RFLP analysis. The detection of a RFLP does not provide information as to the mechanism by which it was created. Moreover, from RFLP data, it is also not possible to predict how the individuals differ from each other at the molecular level.

Attempts to identify RFLPs between different inbred strains of mice often meet with limited success even after, testing with large numbers of enzymes. In one study, RFLPs were identified at only 30% of the single copy loci tested with 22 different restriction enzymes. Furthermore, when RFLPs are identified, they are almost always 'di-allelic binary systems — the insertion, deletion, or restriction site change is either present or absent. Unfortunately, di-allelic loci can only be mapped in crosses where the two parental chromosomes carry the two alternative alleles. Thus, even if a RFLP is identified between two inbred strains of mice, there is no guarantee that another pair of strains will also happen to- carry alternative alleles. As a consequence, only a subset of the RFLP markers developed for analysis of one cross between traditional mouse strains will be of use for mapping in a cross between any other pair of inbred strains.

8.4.2 Choice of restriction enzymes to use for RFLP detection

With so many restriction enzymes available, how does one decide which ones are the best to use in the search for RFLPs? Obviously, cost is an important consideration. Another consideration is whether the enzyme is optimally active with genomic DNA obtained from animal tissues. However, a critical consideration is the rate at which RFLPs can be detected based on the enzyme that is chosen.

A systematic study of RFLP detection between B6 and *M. spretus* DNA subsequent to digestion with one often different enzymes has been reported. One

hundred and ten anonymous DNA sequences of less than 4 kb in length were used as probes. The highest rate of RFLP detection — 63% — was observed with DNA digested with *Taql.* The second highest rate — 56% — was observed with *Mspl.* In decreasing order of effectiveness were the enzymes *BainRl* (50%), *Xbal* (47%), *Pstl* (44%), *Bg*II (41%), *Hind* III (39%), *Pvull* (38%) *Rsa* I (38%), and *Eco* RI (33%). It is ironic that of the ten enzymes tested, the one most commonly used in molecular biological research — *EcoRI* — was the worst one, by a long shot, at detecting polymorphisms.

8.4.3 Minisatellites: variable number tandem repeat loci

In contrast to traditional RFLPs caused by basepair changes in restriction sites, a special class of RFLP loci present in all mammalian genomes is highly polymorphic with very large numbers of alleles. These "hypervariable" loci were first exploited in a general way by Jeffreys (1985) and his colleagues for genetic mapping in humans.

Hypervariable RFLP loci of this special class are known by a number of different names including variable number tandem repeat (VNTR) loci and *minisatellites,* which is the more commonly used term today. Minisatellites are composed of unit sequences that range from 10 to 40 bp in length and are tandemly repeated from tens to thousands of times. Although various functions have been suggested for mini satellite loci as a class, none of these has withstood the test of further analysis. Rather, it appears most likely that minisatellite loci evolve in a neutral manner through expansion and contraction caused by unequal crossing over between out-of-register repeat units. Recombination events of this type will yield reciprocal products which both represent new alleles with a change in the *number* of repeat units.

The frequency with which new alleles are created at minisatellite loci — on the order of 10^{-3} per locus per gamete — is much greater than the classical mutation rate of 10^{-5} to 10^{-6}. This leads to a much higher level of polymorphism between unrelated individuals within a population. At the same time, one change in a thousand gametes is low enough so as to not interfere with the ability to follow minisatellite alleles in classical breeding studies.

Length polymorphisms at minisatellite loci are most simply detected by digestion of genomic DNA samples with a restriction enzyme that does not cut within the minisatellite itself but does cut within closely flanking sequences. As with all other RFLP analyses, the restriction digests are fractionated by gel electrophoresis, blotted and hybridized to probes derived from the polymorphic locus. However, unlike traditional point mutation RFLPs, minisatellites are caused

by, and reflect, changes in the actual size of the locus itself.

The best restriction enzymes to use for minisatellite analysis are those with 4 bp recognition sites such as *Hae*III, *Hin*fI or *Sau*3A; it is likely that one of these enzymes will not cut within the relatively short minisatellite unit sequence, but will cut within several hundred basepairs of flanking sequence on both sides. Standard 1% agarose gels with maximal separation in the 1-4 kb range are usually best for the resolution of minisatellite bands; however, conditions can be optimized for each minisatellite system under analysis.

8.4.4 RAPD

RAPD stands for random amplification of polymorphic DNA. It is a type of PCR reaction, where random segments of genomic DNA are amplified with single primer of arbitrary nucleotide sequence and which are able to differentiate between genetically distinct individuals, although not necessarily in a reproducible way. By resolving the resulting patterns, a semi-unique profile can be gleaned from a RAPD reaction (Fig. 8.10).

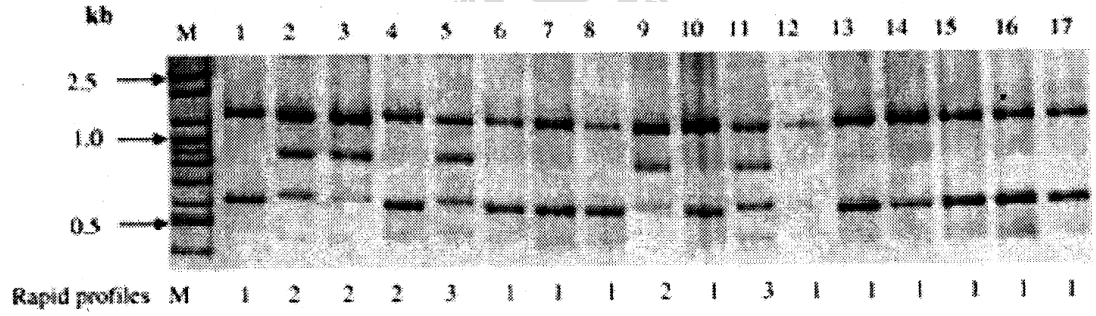


Fig. 8.10: Randomly amplified genomic DNA with short primers

No knowledge of the DNA sequence for the targeted gene is required, as the primers will bind **somewhere** in the sequence, but it is not certain exactly where. This makes the method popular for comparing the DNA of biological systems that have not had the attention of the scientific community, or in a system in which relatively few DNA sequences are compared (it is not suitable for forming a DNA databank). Due to the fact that it relies on a large, intact DNA template sequence, it has some limitations in the use of degraded DNA samples. Its resolving power is much lower than targeted, species specific DNA comparison methods, such as

short tandem repeats. In recent years, RAPD is used to characterize, and trace, the phylogeny of diverse plant and animal species.

Limitations of RAPD

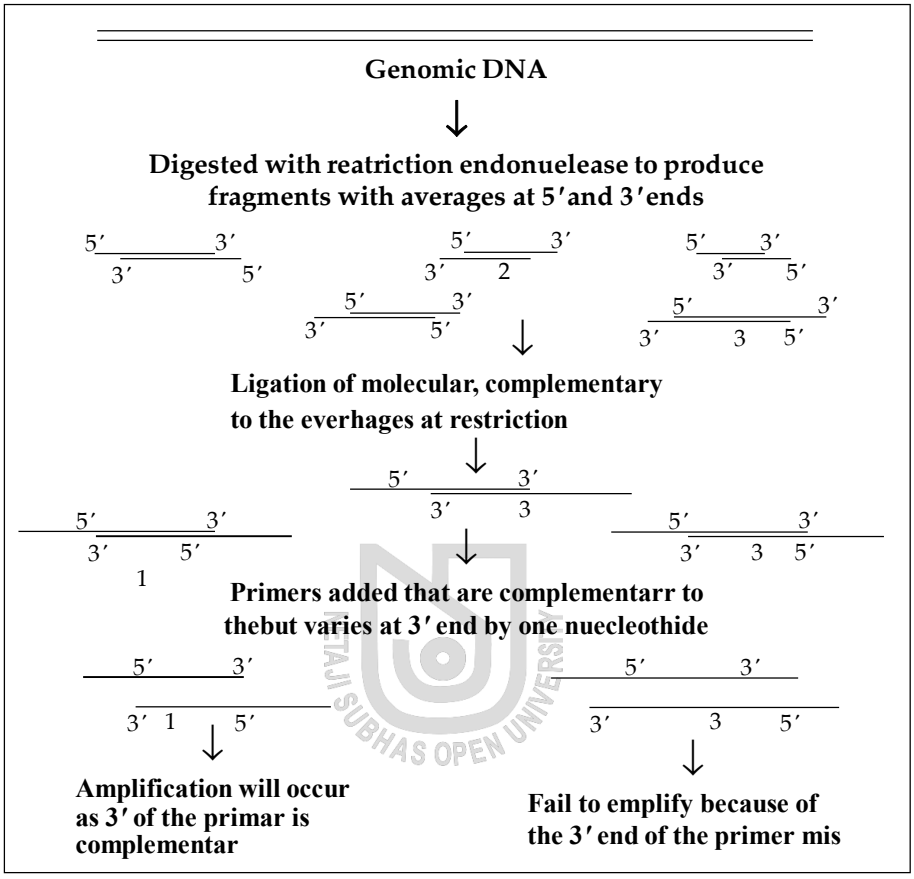
- Nearly all RAPD markers are dominant, i.e. it is not possible to distinguish whether a DNA segment is amplified from a locus that is heterozygous (1 copy) or homozygous (2 copies). Co-dominant RAPD markers, observed as different-sized DNA segments amplified from the same locus, are detected only rarely.

- PCR is an enzymatic reaction, therefore the quality and concentration of template DNA, concentrations of PCR components, and the PCR cycling conditions may greatly influence the outcome. Thus, the RAPD technique is notoriously laboratory dependent and needs carefully developed laboratory protocols to be reproducible.

- Mismatches, between the primer and' the template may result in the total absence of PCR product as well as in a merely decreased amount of the product. Thus, the RAPD results can be difficult to interpret.

8.4.5 AFLP

AFLP stands for Amplified Fragment Length Polymorphism which is a hybrid of RFLP and RAPD techniques. Genomic DNA is cut with restriction enzymes, as in RFLP. Typically, two different restriction enzymes are used. The idea is to produce a large number of fragments. Some of the fragments are selectively amplified with PCR using "random" primers, as in RAPD. The primers are not really random, however. Specific oligonucleotide "adapters" (these are complementary to the restriction sites) of 25-30 bp are ligated to the restricted DNA fragments. The primers are complementary to these adapters. However, the primers vary at their 3'-end, such that they will amplify only a subset of the restricted DNA fragments. Typically 50-100 restriction fragments are amplified and detected on denaturing polyacrylamide gels. AFLPs have typically been used to study variation among individuals of a species, most commonly for producing genetic maps (and in trying to find genes responsible for certain traits). They have received



limited attention as tools in systematics, perhaps because the method is relatively labor intensive when compared with other methods. The power of AFLP is based upon the molecular genetic .variations that exist between closely related species, varieties, orcultivars. These variations in DNA sequence are exploited by the AFLP technology such that "fingerprints" of particular genotypes can be routinely generated. These "fingerprints" are simply RFLPs visualized by selective PCR amplification of DNA restriction fragments.