M.Sc. Previous Year

Chemistry, MC-05

BIOLOGY FOR CHEMISTS

मध्यप्रदेश भोज (मुक्त) विश्वविद्यालय – भोपाल

MADHYA PRADESH BHOJ (OPEN) UNIVERSITY – BHOPAL

COURSE WRITERS

Dr. Pradeep Kumar, Former Faculty Department of Paramedical sciences at Management and Computer Science Institute (MCSI), New Delhi
Units: (1.0-1.1, 1.2-1.2.3, 1.2.4, 1.4.4, 1.5, 1.7-1.11, 2.0-2.7, 2.9-2.13, 3.0-3.1, 3.4-3.5.2, 3.6-3.10, 4, 5.0-5.2.5, 5.3, 5.4.6-5.4.9, 5.5-5.9)
Dr. Pradeep Kumar, Former Faculty Department of Paramedical sciences at Management and Computer Science Institute (MCSI), New Delhi
Rahul Dev Ambedkar, Assistant professor Depatrment of Botany, Sri Aurobindo College, Delhi
Units: (3.2, 3.3, 1.3, 1.6, 3.5.3)
Preetika Chandna, Sr. Lecturer, Subharti Dental College, Swami Vivekanand Subharti University, Meerut
Vivek Kumar Adlakha, Lecturer, Subharti Dental College, Swami Vivekanand Subharti University, Meerut
Units: (1.4-1.4.3)
Vijay Kumar Jha, Lecturer, Department of Biochemistry, Netaji Subash Medical College and Hospital, Bihar
Units: (2.8)
Dr. Radhika Sharma, Assistant Professor (Zoology), Department of Biology and Environmental Sciences, College of Basic Sciences, CSK Himachal Pradesh Krishi Vishvidayala, Palampur, H.P.
Units: (5.4-5.4.5)

Published by Registrar, MP Bhoj (Open) University, Bhopal in 2020

Vikas® is the registered trademark of Vikas® Publishing House Pvt. Ltd.

VIKAS® PUBLISHING HOUSE PVT. LTD.
E-28, Sector-8, Noida - 201301 (UP)
Phone: 0120-4078900 • Fax: 0120-4078999
Regd. Office: A-27, 2nd Floor, Mohan Co-operative Industrial Estate, New Delhi 1100 44
• Website: www.vikaspublishing.com • Email: helpline@vikaspublishing.com

SYLLABI-BOOK MAPPING TABLE

Biology for Chemists

Syllabi	Mapping in Book
UNIT - 1: CELL STRUCTURE AND FUNCTIONS Structure of prokaryotic and eukaryotic cells, intracellular organelles and their functions, comparison of plant and animal cells. Overview of metabolic processes - catabolism and anabolism. ATP - the biological energy currency. Origin of life - unique properties of carbon, chemical evolution and rise of living systems. Introduction to biomolecules, building blocks of bio-macromolecules.	**Unit-1:** Cell Structure and Functions **(Pages 3-66)**
UNIT - 2: CARBOHYDRATES Conformation of monosaccharides, structure and functions of important derivatives of monosaccharides like glycosides, deoxy sugars, myo-inositol, amino sugars. N-acetylmuramic acid, sialic acid, disaccharides and polysaccharides. Structural polysaccharides - cellulose and chitin. Storage polysaccharides - starch and glycogen. Structure and biological functions of glucosaminoglycans or mucopolysaccharides. Carbohydrates of glycoproteins and glycolipids. Role of sugars in biological recognition. Blood group substances. Ascorbic acid. Carbohydrate metabolism-Kreb's cycle, glycolysis, glycogenesis and glycogenolysis, gluconeogenesis, pentose phosphate pathway.	**Unit-2:** Carbohydrates **(Pages 67-120)**
UNIT - 3: LIPIDS Fatty acids, essential fatty acids, structure and function of triacylglycerols, glycerophospholipids, sphingolipids, cholesterol, bile acid, prostaglandins, Lipoproteins - composition and function, role in atherosclerosis. Properties of lipid aggregates-micelles, bilayers, liposomes and their possible biological functions. Biological membranes. Fluid mosaic model of membrane structure. Lipid metabolism - β -oxidation of fatty acids.	**Unit-3:** Lipids **(Pages 121-189)**
UNIT - 4: AMINO-ACIDS, PEPTIDES AND PROTEINS Chemical and enzymatic hydrolysis of proteins to peptides, amino acid sequencing. Secondary structure of proteins, forces responsible for holding of secondary structure. α-helix, β-sheets, super secondary structure, triple helix structure of collagen. Tertiary structure of protein-folding and domain structure. Quaternary structure. Amino acid metabolism - degradation and biosynthesis of amino acids. Sequence determination: chemical/enzymatic/ Mass spectrometry, racemization/detection, Chemistry of Oxytocin and Tryptophan Releasing Hormone (TRH).	**Unit-4:** Amino Acids, Peptides and Proteins **(Pages 191-260)**

UNIT - 5: NUCLEIC ACIDS

Purine and pyrimidine bases of nucleic acids, base pairing via H-Bonding. Structure of Ribonucleic Acids (RNA) and Deoxyribonucleic Acid (DNA), double helix model of DNA and forces responsible for holding it. Chemical and enzymatic hydrolysis of nucleic acids. The chemical basis for heredity, and overview of replication of DNA, transcription, translation and genetic code. Chemical synthesis of mono and trinucleoside.

Unit-5: Nucleic Acids
(Pages 261-318)

CONTENTS

INTRODUCTION

NOTES

The study of chemical processes within and pertaining to living beings is known as biochemistry or biological chemistry. Biochemistry is divided into three fields: structural biology, enzymology, and metabolism. It is a sub-discipline of both chemistry and biology. Biochemistry has proven effective at understanding life processes through these three disciplines in the later decades of the twentieth century. Biochemical methods and study are being used to explore and develop almost every aspect of the biological sciences. Biochemistry focuses on understanding the chemical base that permits biological molecules to give birth to the activities that occur within living cells and between cells, which has a lot to do with the form and function of tissues and organs. Biochemistry is intimately linked to molecular biology, which is the study of molecular mechanisms of biological phenomena.

The structures, bonds, activities, and interactions of biological macromolecules such as proteins, nucleic acids, carbohydrates, and lipids are the focus of most of biochemistry. They give cells their structure and conduct many of the processes that make life possible. The chemistry of the cell is also influenced by small molecule and ion processes. These might be inorganic (for example, water and metal ions) or organic (for example, the amino acids, which are used to synthesize proteins). Metabolism is the process through which cells employ chemical processes to extract energy from their surroundings. Biochemistry's results are mostly used in medicine, nutrition, and agriculture. Biochemists study illness causes and treatments in medicine. Nutrition is the study of how to keep one's health and well-being, as well as the consequences of dietary deficiencies. Biochemists study soil and fertilizers in agriculture. Crop production, storage, and pest control are all objectives.

This book, *Biology for Chemists*, has been designed keeping in mind the Self-Instruction Mode (SIM) format and follows a simple pattern, wherein each unit of the book begins with the Introduction followed by the Objectives for the topic. The content is then presented in a simple and easy-to-understand manner, and is interspersed with Check Your Progress questions to reinforce the student's understanding of the topic. A list of Self-Assessment Questions and Exercises is also provided at the end of each unit. The Summary and Key Terms further act as useful tools for students and are meant for effective recapitulation of the text.

UNIT 1 CELL STRUCTURE AND FUNCTIONS

NOTES

Structure

1.0 INTRODUCTION

In biology, cell theory is the historic scientific theory and is universally accepted that all living organisms are made up of 'cells'. The scientists have defined that the cells are the basic structural and organizational unit of all living organisms, and that all cells originate from pre-existing cells. Principally, the cells are the basic unit of structure in all organisms and also the basic unit of reproduction. With the discovery and advancement of microscopes and magnification technology over time, Robert Hooke was the first who studied the cork cells under the microscope and gave the very significant theory on scientific study of cells, also known as 'cell biology' or 'cell theory'. Cell theory was eventually formulated in 1839. The three principles to the cell theory include that, all living organisms are composed of one or more cells, the cell is the basic unit of structure and organization in organisms and the cells arise from pre-existing cells. In this unit, we will discuss the structure of cell including the structure and functions of prokaryotic and eukaryotic cells and intracellular organelles, along with the comparison of animal and plant cells. It will also focus on the metabolic processes and origin of life, along with the concept of biomolecules.

NOTES

1.1 OBJECTIVES

After going through this unit, you will be able to:

- Describe the structure of cell
- Explain the structure and functions of prokaryotic and eukaryotic cells and intracellular organelles
- Explain the metabolic processes and origin of life
- Discuss the concept of biomolecules

1.2 FUNDAMENTALS OF CELL THEORY

By definition, a 'cell' is the fundamental and structural unit of all living organisms. It is the smallest biological, structural and functional unit of all plants and animals. Therefore, cells are called the 'building blocks of life' or the 'basic units of life'. Organisms made up of a single cell are termed as 'unicellular' whereas organisms made up of many cells are termed as 'multicellular'. Cells perform many different functions within a living organism, such as digestion, respiration, reproduction, etc., and keep it alive. For example, within the human body there are various types of cells which perform specific functions. Basically, the cells form tissues and multiple tissues make up an organ, different organs create an organ system, such as digestive system, respiratory system, circulatory system, nervous system, etc., to perform specific functions in the human body and any other living organism. Therefore,

$$\text{Cells} \rightarrow \text{Tissue} \rightarrow \text{Organ} \rightarrow \text{System} \rightarrow \text{Organism/Human Body}$$

Fundamentally, all organisms are composed of structural and functional units of life called 'cells'. The body of some organisms like bacteria, protozoans and some algae is made up of a single cell while the body of fungi, plants and animals are composed of many cells. Human body is built of about one trillion cells. Cells vary in size and structure as they are specialized to perform different functions. But the basic components and functions of the cell are common to all cells.

Landmarks in Cell Study

Soon after Antonie van Leeuwenhoek invented the microscope, Robert Hooke in 1665 observed a piece of cork under the microscope and found it to be made of small compartments which he called 'cells', in Latin 'cell' means 'small room'. In 1672, Leeuwenhoek observed bacteria, sperm and red blood corpuscles, all of which were cells. In 1831, Robert Brown, an Englishman observed that all cells had a centrally positioned body which he termed as nucleus.

Antonie van Leeuwenhoek is another scientist who saw these cells soon after Hooke. He made use of a microscope containing improved lenses that could magnify objects almost 300-fold or 270x. Leeuwenhoek studied the tiny organisms under the microscope and named them 'animalcules', which included protozoa and other unicellular organisms, like bacteria. Antonie van Leeuwenhoek is regarded as the 'Father of Microbiology'. He is known for the discovery of bacteria.

In 1838, M.J. Schleiden and Theodore Schwann formulated the 'cell theory'. The cell theory is based on the following three principles:

- All organisms are composed of cells.
- Cell is the structural and functional unit of life.
- Cells arise from pre-existing cells.

Fundamentals of Cell

The cell is the fundamental unit of life. All living organisms on planet Earth are composed unicellular (single cell) or multicellular (many cells). Cells range in its size from a millimeter to microns and generally varies in their shapes. Few cells are flat, oval, rod, curved, spherical, concave, rectangular, and various other shapes are also found. Most of the cells are microscopic in size and can only be seen under the microscope. Some cells are fairly long and large. For example, a neuron in the human body is approximately 100 microns or 1 meter long and the ostrich egg is the largest cell which ranges from 14-15 cm long and 12-13 cm wide (Refer Figures1.1 and 1.2).

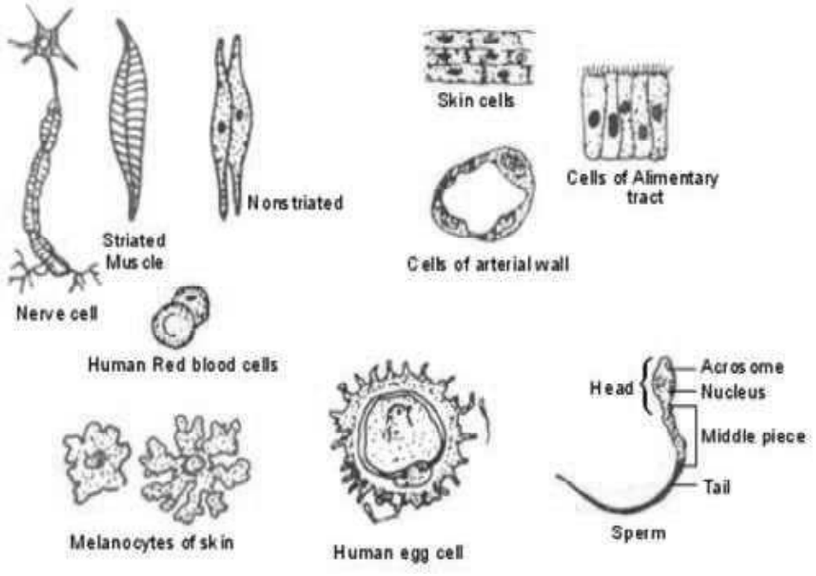


Fig. 1.1 *Different Shapes of Cells*

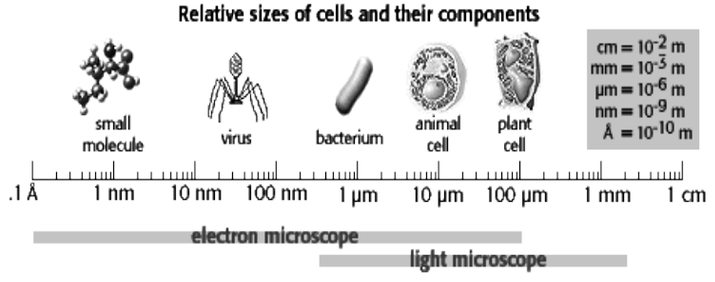


Fig. 1.2 *Different Sizes of Cells and Their Components*

NOTES

Animal cells are a typical eukaryotic cell with a membrane-bound nucleus with the presence of DNA inside the nucleus. They also comprise of other organelles and cellular structures which carry out specific functions necessary for the cell to function properly.

A cell may be defined as a unit of protoplasm bounded by a plasma or cell membrane and possessing a nucleus. Protoplasm is the life giving substance and includes the cytoplasm and the nucleus. The cytoplasm has in it organelles, such as Ribosomes, Mitochondria, Golgi Bodies, Plastids, Lysosomes and Endoplasmic Reticulum. Plant cells have in their cytoplasm large vacuoles containing non-living inclusions like crystals, pigments, etc. The bacteria have neither organelles nor a well formed nucleus. But every cell has following three major components:

- Plasma Membrane
- Cytoplasm
- DNA (naked in bacteria and covered by a membrane in all other organisms)

1.2.1 Structure of Prokaryotic and Eukaryotic Cells

There are many different types of cells in the organism, such as blood cells, skin cells, bone cells and even bacteria. All cells, whether from bacteria, human, or any other organism will be one of two general types. There is another basic cell structure that is present in many but not all living cells, termed as 'nucleus'. The nucleus of a cell is a structure in the cytoplasm that is surrounded by a membrane (the nuclear membrane) and contains, and protects, most of the cell's DNA. Based on whether they have a nucleus, the cytologists recognize two basic types of cells as Eukaryotic cell and Prokaryotic cell (Refer Figure 1.3). Their differences have been tabulated below in Table 1.1. Organisms which do not possess a well formed nucleus are prokaryotes, such as the bacteria. All others possess a well-defined nucleus, covered by a nuclear membrane and are termed as eukaryotes.

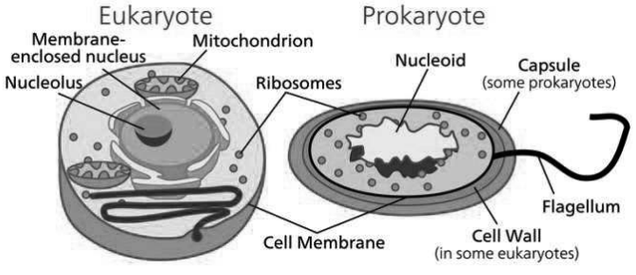


Fig. 1.3 *Two Basic Types of Cells*

Higher eukaryotes have multiple organs to perform specific functions, such as liver, kidney and heart. Each organ has specific tissue and each tissue is composed of cells. Hence, cell is the structural and functional unit of life as it contains all necessary infrastructure to perform all functions. Based on cellular structure, cells are classified as Prokaryotic and Eukaryotic cells. In most of the cases, prokaryotes are single cells whereas eukaryotes are either single cells or part of multicellular tissues system. Besides this, both types of cells have several structural and metabolic differences as given in Table 1.1.

NOTES

Table 1.1 *Difference between Eukaryotic Cell and Prokaryotic Cell*

Feature	Prokaryote	Eukaryote
Size	Small, in µm range	Variable size, upto 40µm in diameter.
Genetic material	Circular DNA present in cytosol as free material	DNA in the form of linear chromosome present in well defined double membrane nucleus, no direct connection with cytosol
Replication	Single origin of replication	Multiple origin of replication.
Genes	No Intron	Presence of Intron
Organelles	No membrane bound organelles	Membrane bound orgelles with well defined function.
Cell walls	Very complex cell wall	Except Fungi and plant, eukaryotic cells are devoid of a thick cell wall.
Ribosome	70S	80S
Trancription and translation	Occurs together	Transcription in nucleus and translation in cytosol

Structure of Prokaryotic Cells

A prokaryotic cell is much simpler and smaller than eukaryotic cells. It lacks membrane bound organelles including nucleus. The structure and components of a typical prokaryotic cell is shown in Figure 1.4(A). The description of different structural feature of prokaryotic cells is as follows:

1. **Outer Flagella:** A flagellum attached to the bacterial capsule is a central feature of most of the prokaryotic cell especially of the motile bacteria. It provides motion or locomotion to the bacteria and be responsible for chemotaxsis of bacteria. Movement of bacteria towards a chemical gradient, such as glucose, is known as chemotaxsis. Flagellum is a part of cell wall and its motion is regulated by motor proteins present inside the cell. Flagellar motion is an energy consuming process and it is governed by an ATPase present at the bottom of the shaft. It is made up of protein flagellin and reduction or suppression of flagellar protein reduces bacterial infectivity or pathogenicity and ability to grow.
2. **Bacterial Surface Layers:** Bacteria possess three anatomical barriers to protect the cells from external damage. Bacterial capsule is the outer most layer and made up of high molecular weight polysaccharides. It is impermeable to the water or other aqueous solvent and it is responsible for antigenicity of bacterial cells. Cell wall in bacteria and its response to gram staining is the basis of classification of bacterial species. Gram staining is developed by a Danish scientist Hans Christian Gram. This technique differentiates bacterial strains based on their cell wall composition, especially thickness of the Peptidoglycan Layer. During the staining procedure bacterial sample is stained with two dyes, Crystal Violet and Safarnin. During a washing step with non-polar solvents, such as alcohol or acetone (decolorization), 'Gram –ve' bacteria leave the Blue Stain due to a thin Peptidoglycan Layer in cell wall whereas 'Gram +ve' bacteria retains both stains and appear as Pink. Cell wall composition in Gram –ve and Gram +ve bacteria is different. Bacterial cell wall has different

NOTES

constituents and be responsible for their reactivity towards Gram stain (Refer Figure 1.4(B)).

A. Peptidoglycan Layer: The peptidoglycan layer is thick in Gram +ve bacteria and thin in Gram –ve bacteria. Peptidoglycan is a polymer of NAG (N-Acetyl-Glucosamine) and NAM (N-Acetyl-Muramic Acid) linked by a ^{2}-(1,4) linkage. Sugar polymer are attached to peptide chain composed of amino acids, L-Alanine, D-Glutamic Acid, L-Lysine and D-Alanine. Peptide chain present in one layer cross linked to the next layer to form a mesh work and be responsible for physical strength of the cell wall. Peptidoglycan synthesis is targeted by antibiotics, such as, penicillin, whereas, lysozyme (present in human saliva or tears) degrades the peptidoglycan layer by cleaving glycosidic bond connecting NAGNAM to form polymer.

B. Lipoteichoic Acids: LipoTeichoic Acid (LTA) are only found in Gram +ve bacteria cell wall and it is an important antigenic determinant.

C. Lipopolysaccharides: The LipoPolySaccharides (LPS) are found only in Gram –ve bacterial cell wall and it is an important antigenic determinant.

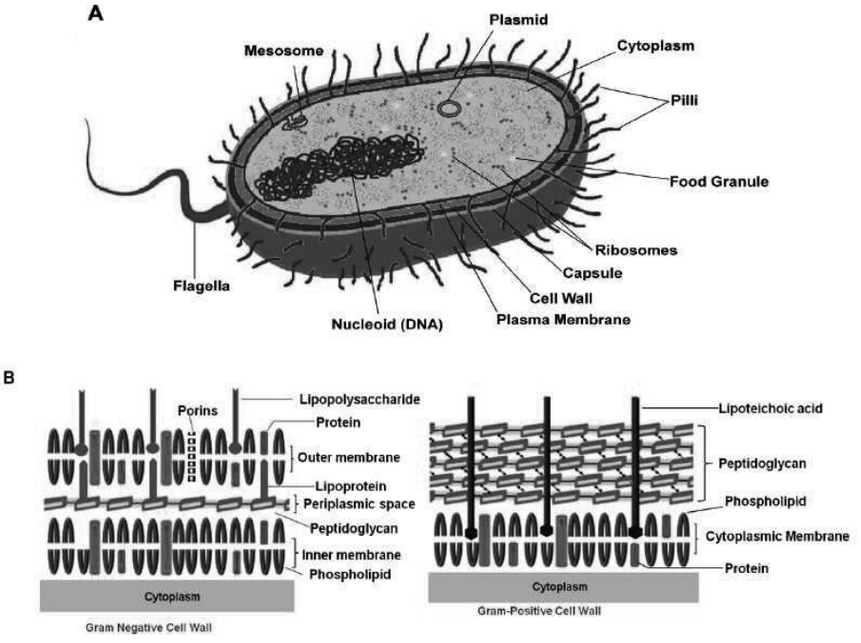


Fig. 1.4 *Structural Details of a Typical Prokaryotic Cell (A); Composition of Cell Wall of Gram Negative and Gram Positive Bacteria (B)*

3. **Cytosol and Other Organelles:** Prokaryotic cells do not contain any membrane bound organelle. The organelles are present in cytosol, such as ribosome (70S), genetic material whereas electron transport chain complexes are embedded within the plasma membrane.
4. **Chromosome and Extra Chromosomal DNA:** Prokaryote cell contains genetic material in the form of circular DNA, known as 'bacterial chromosome'. It contains genetic elements for replication, transcription and translation. Bacterial chromosome follows a rolling circle mode of DNA replication. The genes present on chromosome does not contains non coding

region (introns) and it is co-translated to protein. Besides main circle DNA, bacteria also contains extra chromosomal circular DNA known as 'plasmid'. Presence of plasmid containing resistance gene confers resistance towards known antibiotics. Exchange of extra-chromosomal DNA between different bacterial strains is one of the mechanisms responsible for spread of antibiotic resistance across the bacterial population.

Structure of Eukaryotic Cell

The eukaryotic cell is much more complex and it contains many membrane bound organelles to perform specific functions. It contains a nucleus isolated from cytosol and enclosed in a well defined double membrane. Eukaryotic cells are usually larger than prokaryotic cells, and they are found mainly in multicellular organisms. Organisms with eukaryotic cells are called eukaryotes, and they range from fungi to humans.

Eukaryotic cells also contain other organelles besides the nucleus. An organelle is a structure within the cytoplasm that performs a specific job in the cell. Organelles called mitochondria, for example, provide energy to the cell, and organelles called vacuoles store substances in the cell. Organelles allow eukaryotic cells to carry out more functions than prokaryotic cells can. This allows eukaryotic cells to have greater cell specificity than prokaryotic cells. Ribosomes, the organelle where Proteins are made, are the only organelles in prokaryotic cells. A typical eukaryotic animal and plant cell is shown in Figure 1.5.

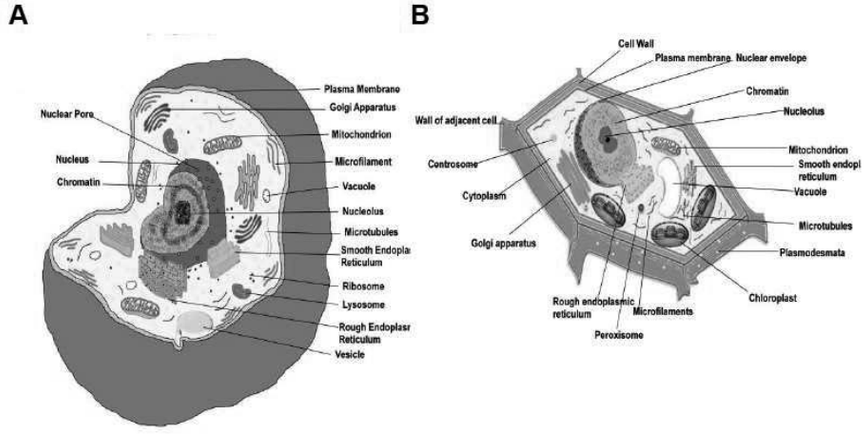


Fig. 1.5 *Structure of Eukaryotic Cell - Animal Cell (A), Plant Cell (B)*

Different Organelles of Eukaryotic Cells (Animal)

Following is the description of different structural features of eukaryotic cell:

1. **Cytosol:** Cytosol is the liquid part filled inside the cell and it contains water, salt, macromolecules (Protein, Lipid, RNA). It has an array of microtubule fiber running throughout the cytosol to give vesicular structure to its destination. Besides this, cytosol exhibits 'Sol' to 'Gel' transition and such transition regulates multiple biochemical and cellular processes.
2. **Nucleus:** Nucleus is the central processing unit of eukaryotic cell and homologous to the processor in a typical computer (Refer Figure 1.6(A)).

The liquid filled inside nucleus is called as 'nucleoplasm'. It is a viscous liquid containing nucleotides and enzymes to perform replication, transcription, DNA damage repair, etc. It contains genetic material (DNA) in a complex fashion involving several proteins (histones) to pack into nuclear bodies or chromosomes. The chromatin in eukaryotic nucleus is divided into Euchromatin and Heterochromatin. Euchromatin is a part of chromatin where DNA is loosely packed and it is transcriptionally active to form mRNA, whereas, heterochromatin is more densely packed and it is transcriptionally inactive. Nuclei in eukaryotic cells are present in a double layer of membrane known as nuclear envelope (Refer Figure 1.6(B)). Outer membrane of nuclear envelope is continuous with the rough endoplasmic reticulum and has ribosome attached to it. The space between these two membranes is called as perinuclear space. Nuclear envelope often has nuclear pore and as per calculation an average nucleus has 3000-4000 pores per nuclear envelope.

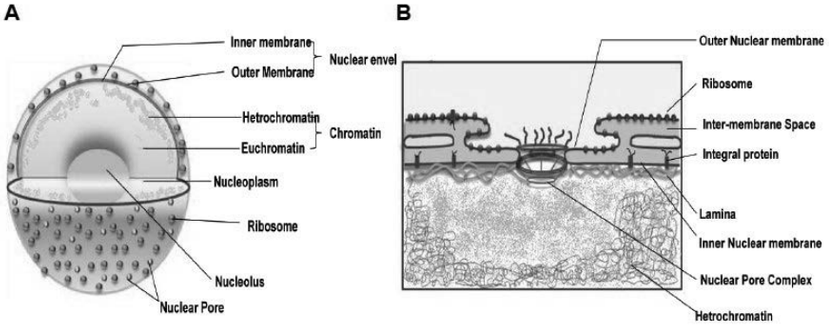


Fig. 1.6 *Structural Details of Eukaryotic Nucleus - Entire Nucleus (A) and Enlarged View of Nuclear Pore (B)*

Nuclear pore is 100nm is diameter and consists of several proteins. It is a gateway for transfer of material between nucleus and cytosol. RNA formed after transcription from DNA within the nucleus and move out of the nucleus into the cytosol through nuclear pore. Similarly protein from cytosol crosses nuclear pore to initiate replication, transcription and other processes.

3. **Mitochondria:** Mitochondria is popularly known as 'power house of the cell' as the organelle is actively involved in the generation of ATP to run the cellular activities. Mitochondria is a double layered membrane-bound organelle with different structural properties (Refer Figure 1.7(A)). Outer membrane is smooth and cover the complete organelle with large number of integral proteins, known as porins. Porin allows free movement of molecules less than 5000da within and outside mitochondria. Whereas, larger molecules or proteins moves into the mitochondria through transporters involving signal peptides known as 'mitochondrial targeting sequence'. Inner membrane is folded into membrane projections to form cristae. Cristae occupies major area of membrane surface and house machinery for anaerobic oxidation and electron transport chain to produce ATP. Due to presence of inner and outer membrane, mitochondria can be divided into 2 compartments: first in between the inner and outer membrane, known as intermembrane

space and second inside the inner membrane known as matrix. The proteins present in intermembrane space have a role in executing 'programmed cell death' or 'apoptosis'. Matrix is the liquid part present in the inner most compartment of the mitochondria and it contains ribosome, DNA, RNA, enzymes to run Krebs's cycle and other proteins.

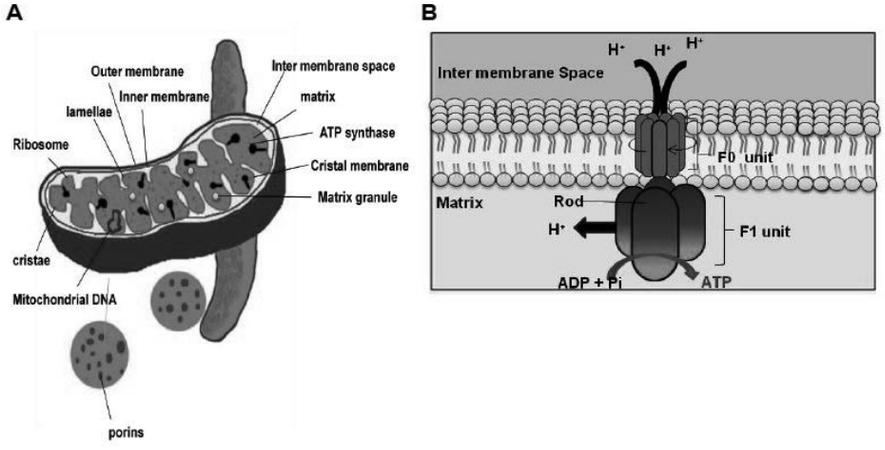


Fig. 1.7 *Mitochondria - Structure of Mitochondria (A) and Enlarged View of ATP Synthase (B)*

Mitochondrial DNA is circular and it has full machinery to synthesize its own RNA (mRNA, rRNA and t-RNA) and proteins. Marked differences exist between mitochondrial DNA and DNA present in nucleus and these differences are not discussed here due to space constrain. Electron transport chain components (Complex I to Complex V) are integral proteins, present in the inner membrane of mitochondria. During metabolic reactions, such as, glycolysis. Krebs's cycle produces large amount of reducing equivalent in the form of NADH2 and FADH2. Electron transport chain process reducing equivalent and flow of the electron through different complexes (Complex I to Complex IV) causes generation of proton gradient across the membrane. Proton expelled in the intermembrane space returned back to the matrix through Complex V (ATP Synthase) to generate ATP. ATP Synthase (Refer Figure 1.7(B)) is a mushroom shaped multimeric protein complex, mainly composed of two proteins F0 and F1. F0 is a membrane bound portion whereas F1 is the complex present into the lumen towards matrix. F0F1 complex of mitochondria harvest the proton motive force to catalyze phosphorylation reaction involving ADP and phosphate to generate ATP. The functions of mitochondria include:

- Production of ATP
- Generation of Reactive Oxygen Species (ROS) in immune cells to kill infectious agents
- Used to track tree of a family
- Role in programmed cell death or 'apoptosis'

4. **Chloroplast:** Chloroplasts are found in plant, algae and other lower invertebrates, such as euglena. Contrasting to mitochondria, chloroplast has

NOTES

outer membrane, an inner membrane and then light pigment containing inner most thylakoid membrane (Refer Figure 1.8(A)). Outer membrane is porous to the small molecules but protein or large molecules are transported by TOC (translocon on the outer chloroplast membrane) complex. Movement of material passed through outer membrane gets into the inner membrane through TIC (translocon on the inner chloroplast membrane) complex. In between outer and inner membrane is intermembrane space filled with aqueous liquid.

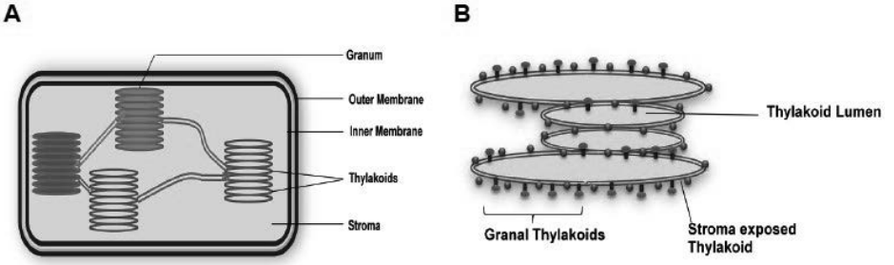


Fig. 1.8 *Structure of Chloroplast (A) and Arrangement of Thylakoid Membrane in Chloroplast (B)*

The inner membrane of the chloroplast further folds to a flattened membrane system known as thylakoids. The photosynthesis machinery, such as light absorbing pigments, electron carriers and ATP synthesizing machinery is present on inner membrane as integral protein complex. Thylakoid membranes are arranged like stack of coin to form granum (Refer Figure 1.8(B)). The granum throughout the chloroplast are connected by tubule to share the material. Overall structure of chloroplast is similar to mitochondria but it has few significant structural and biochemical differences. Thylakoid membrane contains photosynthetic green colored pigment chlorophyll and undergoes the following reaction to produce Glucose ($C_6H_{12}O_6$) and release Oxygen (O_2).

$$6CO_2+6H_2O+\text{Solar Energy}\rightarrow C_6H_{12}O_6+6O_2.$$

Photosynthesis is an assimilation reaction involving CO_2 and water to produce sugar in the presence of solar energy (photons) that catalyzes fusion reaction as shown in the above Equation. The photo system present on thylakoid membrane consists of two photo system, Photo System-I (PS-I) and Photo System complex II (PS-II). PS-II absorbs the photon from solar energy to excite the electron to the higher energy state, and catalyze water break down into the proton and oxygen. The electron pass through multiple electron carrier and during this proton are exported out of the thylakoid membrane into the lumen. The proton passes through ATP synthase and returns back into the stroma to generate ATP (Refer Figure 1.9). The electron from PS-II is eventually been received by PS-I and been excited after absorbing photon from sun light to high energy state. The energy associated with these electrons are used to generate NADPH in the stroma. Hence as a result of photosynthesis, solar energy is trapped by photo synthesis apparatus to generate ATP and NADPH into the lumen. Both of them are used to run Calvin cycle to assimilate environmental CO_2 to form sugar.

NOTES

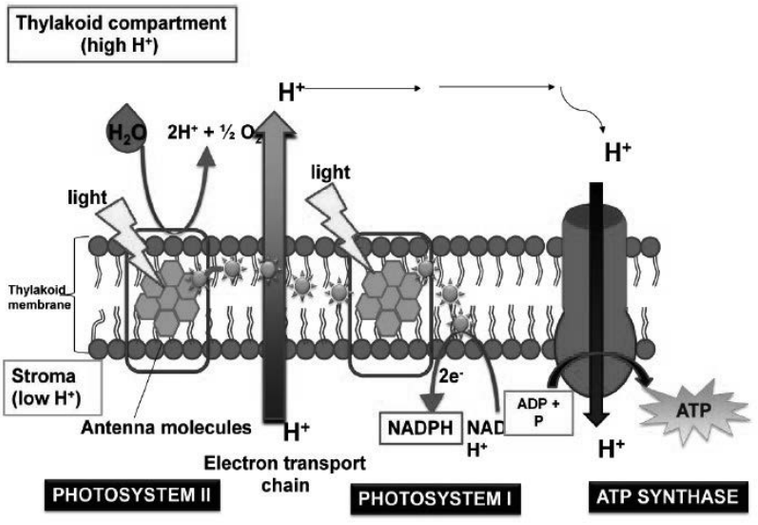


Fig. 1.9 *Different Steps of Photosynthesis*

5. **Organelles of Vesicular Trafficking System:** The main function of these organelles is to manage the distribution of material (food particles or proteins) throughout the cell. Three different organelles, such as Endoplasmic Reticulum (ER), Golgi apparatus and Lysosome, coordinately work together to maintain vesicular transport of material across the cell (Refer Figure 1.10). Eukaryotic cell takes up the solid material from outside through a process called 'endocytosis' whereas uptake of liquid is through a process called as 'pinocytosis'. Similarly material is secreted out of the cells through 'exocytosis'. In addition, intravesicular system delivers protein synthesized in endoplasmic reticulum to different organelles. During endocytosis, material present outside the cells binds to the cells surface through cell surface receptors and trap it in a membranous structure called as endosome. Endosomal vesicles are fused with the lysosomes to form late endosome. In late endosome, with the help of lysosomal enzymes material is digested and then endosome is fused with the golgi bodies and deliver the content for further distribution. In the similar manner, during secretion, vesicles originate from golgi bodies and fuse with the plasma membrane to release the content outside of the cell.

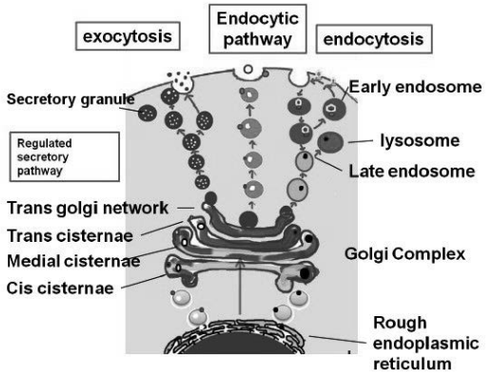


Fig. 1.10 *Intra-Cellular Vesicular Trafficking System of Cell*

NOTES

The vesicular network starts from nuclear membrane and spread throughout the cytosol constitutes endoplasmic reticulum (Refer Figure 1.11). There are two different types of Endoplasmic Reticulum (ER) present in the cell, Rough Endoplasmic Reticulum (RER), and Smooth Endoplasmic Reticulum (SER). RER has ribosome attached to it to give a rough appearance whereas smooth endoplasmic reticulum is devoid of ribosomes. Protein synthesis on ribosome attached to RER are sorted into three different categories, such as integral membrane proteins, proteins for secretion and protein destined for different organelles. Proteins are synthesized with n-signal peptide and these signal peptides are recognized by signal recognition particle on their target organelles. For example, if a protein is synthesized with a signal peptide for mitochondria, it will attach to signal recognition particle and receptor onto the outer mitochondrial membrane to deliver the protein. The proteins without any signal peptide tags are supposed to remain in the cytosol.

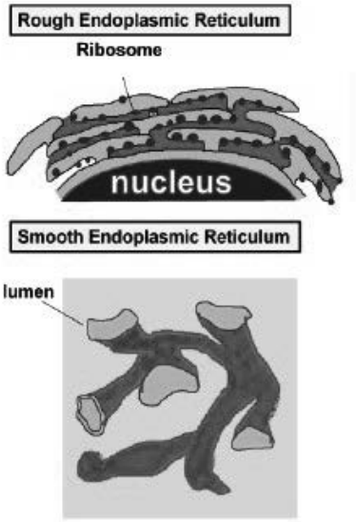


Fig. 1.11 *Endoplasmic Reticulum*

The functions of endoplasmic reticulum are:

- Synthesis of steroid hormone in gonad cells.
- Detoxification.
- Ca^{2+} sequestration.
- Synthesis of protein, phospholipid and carbohydrate.
- Protein sorting to different organelles.
- Protein modifications such as glycosylation, etc.

Golgi bodies were first visualized by a metallic stain invented by Camillo Golgi and are is made of flattened, disk like cisternae arranged in a stacked manner to give three distinct zones (Refer Figure 1.12). Cis-face receives material or vesicles from endoplasmic reticulum, medial golgi is the actual place where protein are covalently modified with the sugar. Trans golgi is

the face of golgi towards plasma membrane and this site sort's vesicle for their destined organelles or plasma membrane.

The functions of golgi bodies include:

- Protein Sorting
- Protein Modifications (Glycosylation)
- Proteolysis

NOTES

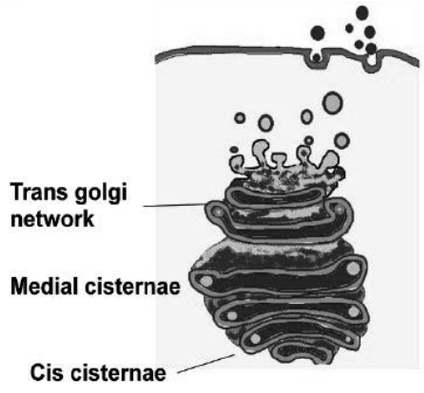


Fig. 1.12 *Schematic Structure - Golgi Bodies*

1.2.2 Intracellular Organelles and Their Functions

All living organisms are made up of cells and new cells are produced when live cells divide. The cell is the smallest unit of life in an organism. The cell lives and, as a result, the organism lives. Whatever an organism does for survival it does for the survival of its cells. All the living matter of a cell is called protoplasm. The cell is surrounded by a cell or plasma membrane. The nucleus is the control centre of the cell. The cytoplasm surrounds the nucleus. The cytoplasm is everything within the cell except for the nucleus. There are many small organelles within the cytoplasm. The fine detail of a cell when seen by an electron microscope is called ultrastructure.

The major components of the cell are (1) Cell Membrane, (2) Cytoplasm and (3) Nucleus.

Ultra Structure of Cell Membrane (Plasma Membrane)

Each cell has a limiting boundary, the cell membrane, plasma membrane or plasmalemma. It is a living membrane, outermost in animal cells but next to cell wall in plant cells. It is flexible and can fold in (as in food vacuoles of amoeba) or fold out (as in the formation of pseudopodia of amoeba). The plasma membrane is made of proteins and lipids and several models were proposed regarding the arrangement of proteins and lipids. The fluid mosaic model proposed by Singer and Nicholson (1972) is widely accepted as shown in Figure 1.13.

NOTES

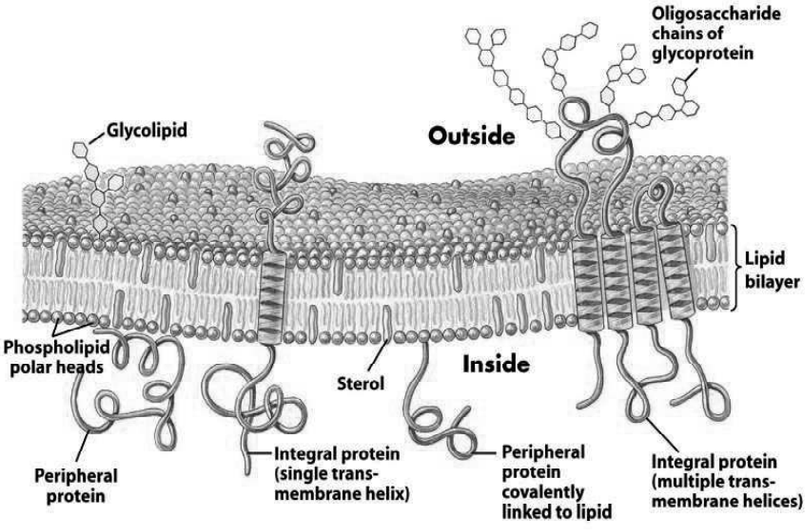


Fig. 1.13 *Fluid Mosaic Model of Cell Membrane*

According to the fluid mosaic model,

(i) The plasma membrane is composed of a lipid bilayer of phospholipid molecules into which a variety of globular proteins are embedded.

(ii) Each phospholipid molecule has two ends, an outer head hydrophilic, i.e., water attracting, and the inner tail pointing centrally hydrophobic, i.e., water repelling.

(iii) The protein molecules are arranged in following two different ways:

- **Peripheral Proteins or Extrinsic Proteins:** These proteins are present on the outer and inner surfaces of lipid bilayer.
- **Integral Proteins or Intrinsic Proteins:** These proteins penetrate lipid bilayer partially or wholly.

The functions of plasma membrane are:

(i) The plasma membrane encloses the cell contents.

(ii) It provides cell shape (in animal cells), for example the characteristic shape of red blood cells, nerve cells, bone cells, etc.

(iii) It allows transport of certain substances into and out of the cell but not all substance, so it is termed selectively permeable.

Transport of small molecules (such as, glucose, amino acids, water, mineral ions, etc.). Small molecules can be transported across the plasma membrane by any one of the following three methods:

- **Diffusion:** Molecules of substances move from their region of higher concentration to their region of lower concentration. This does not require energy. For example, absorption of glucose in a cell.
- **Osmosis:** The movement of water molecules from the region of their higher concentration to the region of their lower concentration through a

semipermeable membrane. There is no expenditure of energy in osmosis. This kind of movement is along concentration gradient.

- **Active Transport:** When the direction of movement of a certain molecules is opposite that of diffusion, i.e., from region of their lower concentration towards the region of their higher concentration, it would require an 'active effort' by the cell for which energy is needed. This energy is provided by ATP (Adenosine TriPhosphate). The active transport may also be through a carrier molecule.

Transport of Large Molecules (Bulk Transport)

During bulk transport the membrane changes its form and shape. It occurs in two ways: (i) Endocytosis (taking the substance in) and (ii) Exocytosis (passing the substance out). Cell membrane regulates movement of substance into and out of the cell. If the cell membrane fails to function normally the cell dies. Endocytosis, as shown in Figure 1.14, is of following two types:

Endocytosis

Phagocytosis	Pinocytosis
1. It involves intake of solid particles.	1. It involves intake of fluid droplets.
2. The membrane folds out going round the particle, forming a cavity and thus engulfing the particle.	2. The membrane folds in and forms a cup like structure sucks in the droplets.

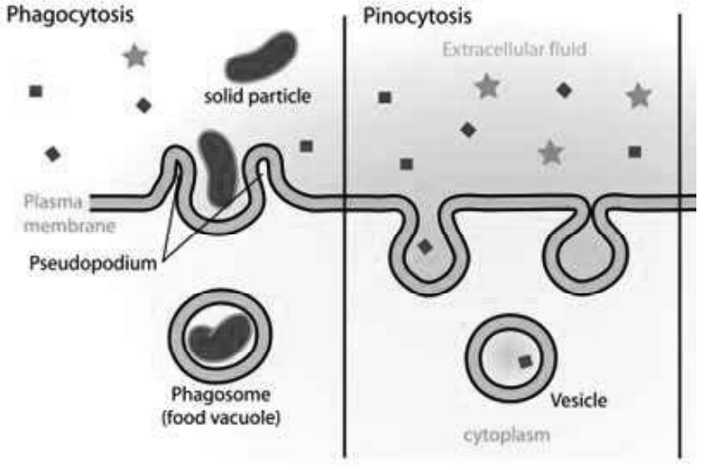


Fig. 1.14 *Diagrammatic Representation of Phagocytosis and Pinocytosis*

Cell Wall

In bacteria and plant cells the outermost cell cover, present outside the plasma membrane is the cell wall. Bacterial cell wall is made of peptidoglycan. Figure 1.15 illustrates the bacterial cell wall.

NOTES

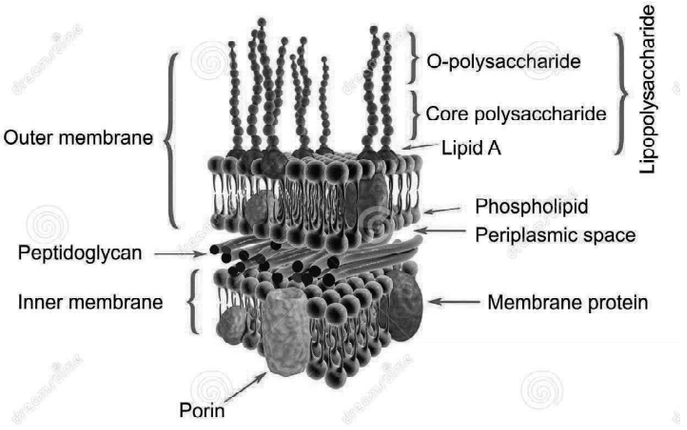


Fig. 1.15 *Bacterial Cell Wall*

Cytosol and Other Organelles

The cytoplasm contains many cell organelles, such as:

1. Those that trap and release energy - mitochondria and chloroplasts.
2. Those that are secretory or involved in synthesis and transport - golgi, ribosomes and endoplasmic reticulum.
3. The organelles for motility - cilia and flagella.
4. The suicidal bags – lysosomes.
5. The nucleus which controls all activities of the cell and carries the hereditary material.

Mitochondria

Mitochondria (found in plant and animal cells) are the energy releasers.

- Appear as tiny thread like structure under light microscope. Approximately 0.5 - 1.00 ¼m (micrometer).
- Number usually a few hundred to a few thousand per cell. Smallest number is just one as in an alga (Micromonas).

The general plan of the internal structure of a mitochondria observed by means of electron microscope is shown in Figure 1.16.

The wall is made of double membrane. The inner membrane is folded inside to form projections called cristae which project into the inner compartment called matrix.

Pyruvic acid oxidises (breakdown product of glucose) to release energy which gets stored in the form of ATP for ready use. This process is also called cellular respiration.

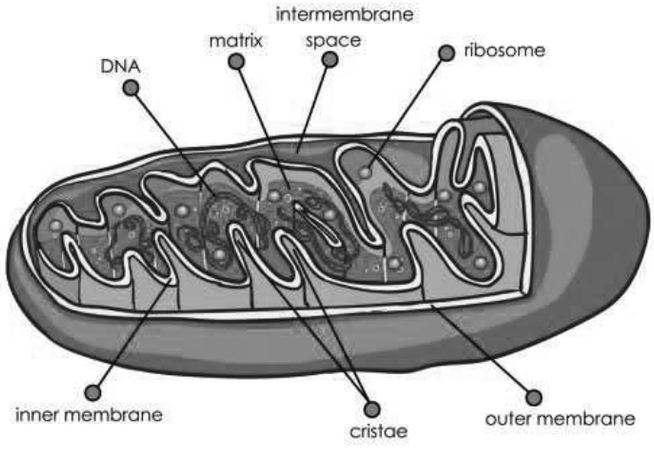


Fig. 1.16 *Structure of a Mitochondrion*

Centriole

It is present in all animal cells (but not in *Amoeba*), located just outside the nucleus. It is cylindrical, 0.5 ¼ m in length and without a membrane. It has nine sets of peripheral tubules but none in the centre. Each set has three tubules arranged at definite angles (Refer Figure 1.17). It has its own DNA and RNA and therefore, it is self duplicating.

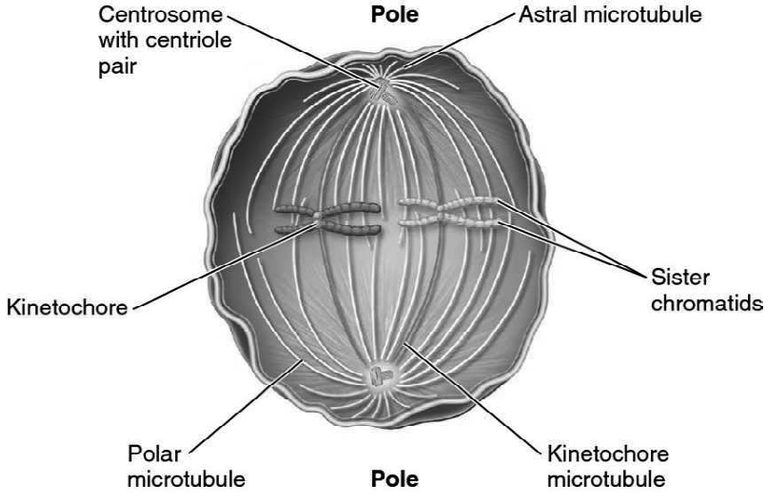


Fig. 1.17 *Structure of Centriole*

Centrioles are involved in cell division. They give orientation to the 'mitotic spindle' which forms during cell division.

Basal Bodies

These are structures similar to centrioles. They have the same nine sets of triplet organization, as in the centrioles. The cilia and flagella appear to arise from the basal bodies.

NOTES

Nucleus: The Hereditary Organelle

General structure of nucleus includes:

(i) It is the largest organelle seen clearly when the cell is not dividing.

(ii) It stains deeply, is mostly spherical, WBC have lobed nuclei.

(iii) It is mostly one in each cell (uninucleate having single nuclei, multinucleate in which some cells have many nuclei).

(v) Double layered nuclear membrane enclosing nucleoplasm which contains chromatin network and a nucleolus (Refer Figure 1.18).

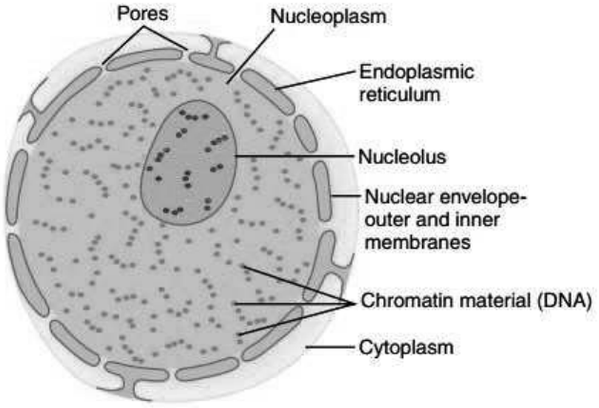


Fig. 1.18 *Structure of Nucleus*

The functions of nucleus include:

- It maintains the cell in a working order.
- It co-ordinates the activities of organelles.
- It takes care of repair work.
- It participates directly in cell division to produce genetically identical daughter cells, this division is called mitosis.
- It participates in production of gametes through another type of cell division called meiosis.

Parts of Nucleus

A nucleus is composed of the following components.

Nuclear Membrane

- Double layered membrane is interrupted by large number of pores.
- Membrane is made up of lipids and proteins (like plasma membrane) and has ribosomes attached on the outer membrane which make the outer membrane rough.
- The pores allow the transport of large molecules in and out of nucleus, and the membranes keep the hereditary material in contact with the rest of the cell.

NOTES

Chromatin

Within the nuclear membrane there is jelly like substance (karyolymph or nucleoplasm) rich in proteins as shown in Figure 1.19. In the karyolymph, fibrillar structures form a network called chromatin fibrils, which gets condensed to form distinct bodies called chromosomes during cell division. On staining the chromosomes, two regions can be identified in the chromatin material heterochromatin dark and autromaticn (light). Heterochromatin has less DNA and genetically less active than euchromatin which has more DNA and genetically more active. Number of chromosomes is fixed in an organism. During cell division chromosomes divide in a manner that the daughter cells receive identical amounts of hereditary matter.

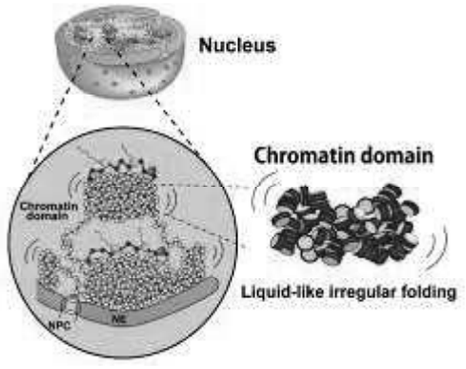


Fig. 1.19 *Structure of Chromatin*

Nucleolus

Membraneless, spheroidal bodies present in all eukaryotic cells except in sperms and in some algae.

Their number varies from one to few, they stain uniformly and deeply.

It has DNA, RNA and proteins.

It is a store house for RNA and proteins; it disappears during cell division and reappears in daughter cells.

It regulates the synthetic activity of the nucleus.

Thus nucleus and cytoplasm are interdependent, and this process is equal to nucleo–cytopalsmic interaction.

Chromosome

The darkly stained, rod shaped bodies visible under light microscope in a cell during metaphase stage of mitosis are referred to as chromosomes. Strasburger was the pioneer man who discovered chromosomes in 1875, and the term chromosome was coined by Waldeyer in 1888. The main features of eukaryotic chromosomes are given below:

- Chromosomes are not visible during interphase under light microscope. During other stages of cell division, they are visible, but are more clearly visible during mitotic metaphase. Hence, they are studied during metaphase.

NOTES

- Chromosomes bear genes in a linear fashion and thus are concerned with transmission of characters from generation to generation.
- Chromosomes of eukaryotes are enclosed by a nuclear membrane, while in prokaryotes, they remain without such envelope free in the cytoplasm.
- Chromosomes vary in shape, size and number in different species of plants and animals.
- Chromosomes have property of self-duplication, segregation and mutation.
- Chromosomes are composed of DNA, RNA and histones. DNA is the major genetic constituent of chromosomes.

Chromosome shape is usually observed during anaphase. The shape of chromosomes is determined by the position of centromere, a part of chromosome on which spindle fibres are attached during metaphase. Chromosomes have generally three different shapes, viz., rod shape, J shape and V shape. These shapes are observed when the centromere occupies terminal, sub-terminal and median (middle) position on the chromosomes, respectively.

Chromosome size is measured with the help of micrometer at mitotic metaphase. It is measured in two ways, viz., in length and in diameter. Plants usually have longer chromosomes than animals. Moreover, species or individuals which have fewer chromosome numbers have larger chromosomes.

The maximum length of chromosome is observed during interphase and minimum during anaphase. Thus chromosome size varies from species to species. Giant chromosomes have length up to 300 μ.

There are three types of chromosome numbers, viz., haploid, diploid and basic number.

- **Haploid:** It represents half of the somatic chromosome number of a species and is denoted by n. Since haploid chromosome number is usually found in the gametes, it is also known as gametic number.
- **Diploid:** It refers to somatic chromosome number of a species and is represented by 2n. Since diploid chromosome number is found in zygotic or somatic cells it is also referred to as zygotic or somatic number.
- **Basic Number:** The gametic chromosome number of a true diploid species is called basic number. It is the minimum haploid chromosome number of any species which is denoted by x. For example, in wheat, the basic number is 7, whereas the haploid number is 7, 14 and 21 for diploid, tetraploid and hexaploid species, respectively. Thus haploid chromosome number differs from basic number. Both are same in case of true diploid species but differ in case of polyploid species. Thus, basic number can be a haploid number but all haploid numbers cannot be basic number. Chromosome number differs from species to species.

Chromosomes are considered as physical basis of inheritance. The first conclusive evidence that chromosomes carry the units of inheritance was put forward by Sutton in 1903. Working with grasshopper, he gave a hypothesis that chromosomes contain genes and their behaviour during meiosis is the physical basis of Mendelian

laws of heredity. Thus his work formed the basis of chromosomal theory of heredity. Now this theory is universally accepted. Various evidences which support that genes are located on the chromosomes which form the physical basis of heredity include the following:

- Two copies of each in somatic cells one copy in gametic cell
- Self-duplication or replication capacity
- Segregation during meiosis
- Mutability

All these parallel features between gene and chromosomes suggest that chromosomes carry genes and represent the physical basis of heredity.

Biochemical studies reveal that hereditary units (genes) are composed of DNA in eukaryotes and RNA in some prokaryotes. The major part of DNA is found in chromosomes which prove beyond doubt that chromosomes are the carriers of hereditary units what we call genes.

1.2.3 Endoplasmic Reticulum and Golgi Apparatus

The Endoplasmic Reticulum (ER) is an important organelle in eukaryotic cells. It plays a major role in the production, processing, and transport of proteins and lipids. The ER produces transmembrane proteins and lipids for its membrane and for many other cell components including lysosomes, secretory vesicles, the Golgi apparatus, the cell membrane, and plant cell vacuoles. It was first reported by Porter in 1945. This continuous membrane system connects the nuclear membrane on one end and the cell membrane on the other.

Structure of Endoplasmic Reticulum (ER)

Endoplasmic Reticulum contains three different types of structure. These are cisternae, vesicles and tubeless (Refer Figure 1.20).

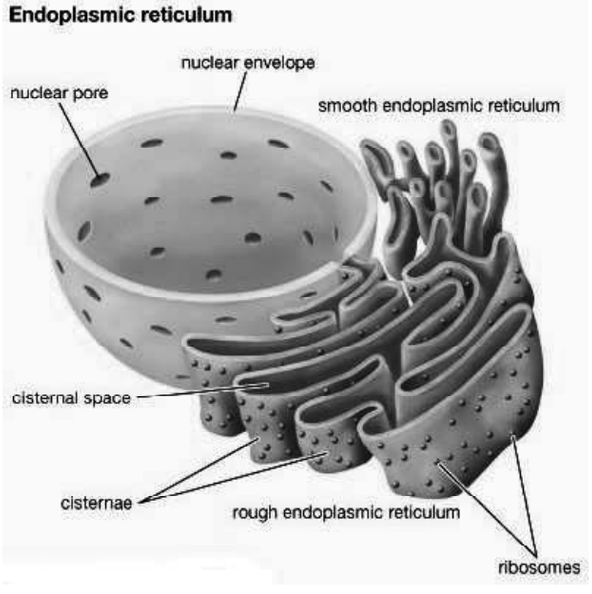


Fig. 1.20 *Structure of Endoplasmic Reticulum*

NOTES

(a) **Cisternae:** These are long, flat and unbranched plates or lamellae arranged in parallel rows.

(b) **Vesicles:** They are usually round or ovoid sacs. They often occur isolated in the cytoplasm.

(c) **Tubules:** They are irregularly branched tube-like structures having a diameter of 50-100n. These are surrounded by this unit membrane of 50-60 thickness and their lumen is filled with the secretary products of the cell.

There are two types of ER, such as smooth walled and rough walled. They may be present in the same or different types of cells namely:

(i) **Smooth Endoplasmic Reticulum (SER):** The surface of this type of reticulum is smooth as ribosome's not attached to it. Smooth ER is actively engaged in steroid synthesis, carbohydrate metabolism, pigment production, etc. in cells.

(ii) **Rough Endoplasmic Reticulum (RER):** The Rough Endoplasmic Reticulum is named so because of its appearance. It is a series of connected flattened sacs that have many ribosomes on their outer surface, hence the name. It synthesizes and secretes proteins in the liver, hormones and other substances in the glands. This type of endoplasmic reticulum is especially prominent in certain kinds of cells like hepatocytes where active protein synthesis occurs.

Functions of Endoplasmic Reticulum (ER)

The basic functions of Endoplasmic Reticulum (ER) are:

- It is mainly responsible for the transportation of proteins and other carbohydrates to another organelle, which includes lysosomes, golgi apparatus, plasma membrane, etc.
- It provide the increased surface area for cellular reactions.
- It helps in the formation of nuclear membrane during cell division.
- It plays a vital role in the formation of the skeletal framework.
- It plays a vital role in the synthesis of proteins, lipids, glycogen and other steroids like cholesterol, progesterone, testosterone, etc.

Golgi Apparatus

The golgi apparatus, like the endoplasmic reticulum, is a canalicular system with sacs that performs some important cellular functions like biosynthesis of polysaccha rides and packaging of cellular products.

In 1898, an Italian biologist, Camillo Golgi discovered a dark yellow network located near the nucleus of nerve cells. This network, which was later identified in other cell types, was named the golgi complex or golgi body. Since originally these were known to be networks, they were also called 'dictyosomes' (Gr., dictyes = net).

The golgi complex occur in all cells except the prokaryotic cells and some eukaryotic cells like mature sieve tubes of plants, mature sperm and red blood

cells of animals. In animal cells, there usually occurs a single golgi apparatus however, its number may vary.

Thus, Paramoeba species has two golgi apparatuses, and nerve cells, liver cells and chordate oocytes have multiple golgi apparatuses. In animal cells, the Golgi apparatus is a localized organelle. Usually it remains polar and occurs in-between the nucleus and the periphery (for example, thyroid cells and goblet cells).

In nerve cells, it occupies a circum-nuclear position. However, in the cells of higher plants, the golgi bodies or dictyosomes are usually found scattered through out the cytoplasm.

Morphology of Golgi Complex

The Golgi apparatus is highly pleomorphic. In some cells it is compact and limited, in others it is reticular. Typically it consists of flattened disk like cisternae with dilated rims and associated vesicles and tubules (Refer Figure 1.21).

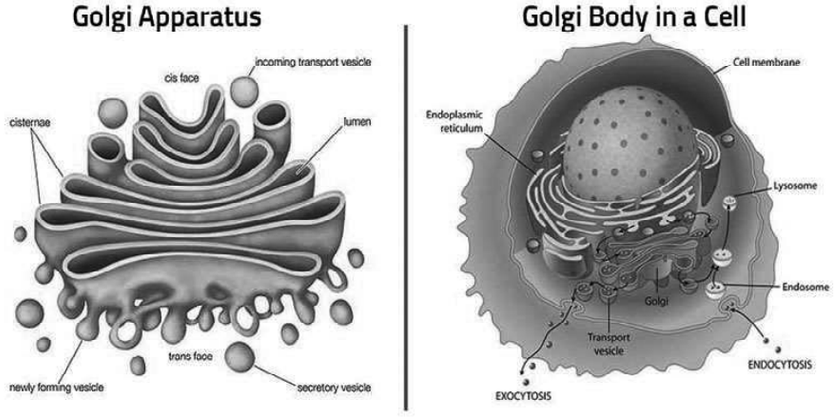


Fig. 1.21 *Structure of Golgi Apparatus*

The detailed structures of the basic components of the golgi apparatus are as follows:

- **Cisternae:** The cisternae are flattened, plate or saucer-like closed compartments. These are arranged in an orderly stack, much like a stack of pancakes. Typically, a golgi stack contains fewer than eight cisternae. Depending on the cell type, an individual cell may contain from a few to several thousand stacks per cell. The diameter of cisternae varies from 0.5 to 1.0 nm and in each stack, cisternae are separated by a space of 20 to 30 nm, which may contain rod-like elements or fibres. Each cisterna is bounded by a smooth unit membrane and is curved in a manner resembling a shallow cup. The cisterna closest to the Endoplasmic Reticulum (ER) is usually convex and is said to be at the cis face or proximal face or forming face, while the cisterna at the opposite end of the stack is of concave shape and said to be at the trans or distal or maturing face. This polarization is called cis-trans axis of the golgi apparatus. Functionally the golgi complex is also divided into four distinct compartments, the cis, medial, trans cisternal, and

NOTES

the Trans Golgi Network (TGN). Newly synthesized membrane, secretory and lysosomal proteins leave the ER and enter the golgi through its cis face and then pass across the stack to the trans face (Refer Figure 1.22).

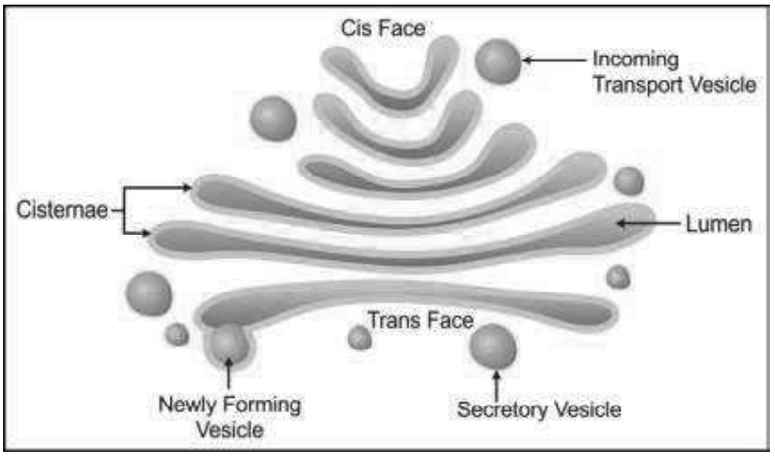


Fig. 1.22 *Cis and Trans Face of Golgi Apparatus*

- **Tubules:** The cis or forming face is characterized by the presence of small transi-tion vesicles or tubules that converge upon the golgi cisternae, forming a kind of fenestrated plate.
- **Vesicles:** Vesicles are those that bud from the ER or from the cis and medial cisternae of the golgi complex and are seen in the electron microscope to be covered by a indistinct fuzzy coat that contains a GTP- binding protgin, called ARF (Adenosylation Ribose Factor).

Functions of Golgi complex

Golgi complex represents a special mem-branous compartment interposed between the ER and the extracellular space, through which there is a continuous traffic of substances. During transportation through different compartments of golgi, substances are modified, transformed and then are directed to their proper destinations. The overall functions of golgi apparatus are as follows:

a. **Synthesis of Glycosphingolipids and Glycoproteins:** The Golgi plays a major role in the glycosidation of lipids and proteins to produce glycosphingolipids and glycoproteins. Following synthesis on the membrane-bound ribosomes of the RER, the polypeptides reach the golgi complex, where the terminal side chains of galactose, fucose, and sialic acid are added by the corresponding transferases present in the golgi apparatus. The golgi also appears to be involved in the addition of sulfate to the carbohydrate moiety of the glycoproteins in cartilage cells.

b. **Secretion:** The main function of golgi complex is cell secretion, not only of exportable proteins but also of the enzymes present in lysosomes and peroxisomes. Secretion may be continuous where the product is discharged as soon as it is elaborated, as found in liver cells and plasma cells. In other cells the secretory cycle is discontinuous, with storage in secretory or zymogen granules,for example pancreas, parotid gland.

In the pancreas, following six steps can be recognized:

(i) **The Ribosomal Stage:** The synthesis of proteins by polysomes attached to the RER.

(ii) **The Cisternal Stage:** The protein is processed and stored within the ER.

(iii) **Intracellular Transport:** The secreted proteins enter the transitional tubules and vesicles that lead to the golgi complex, in which they fuse with large condensing vacuoles present at the maturing face of the golgi. This transport requires the use of energy (ATP) and is same in all protein-producing cells.

(iv) **Concentration of the Secretion:** By the processes of progressive filling, and concentration, the condensing vacu-oles are converted into the zymogen granules that have characteristic electron-opaque content. The conver-sion does not require energy and is probably due to the formation of osmotically inactive aggregates of sulfated peptidoglycans. It involves movement of water from vacuoles to the cytosol.

(v) **Intracellular Storage:** The previous step culminates with the storage of the secretory product into secretory gra-nules, which are released following appropriate stimulus.

(vi) **Exocytosis:** The discharge of secre-tory granules involves its movement towards the apical region and fusion between its membrane and the lumi-nal plasma membrane. Exocytosis requires energy (ATP) and Ca++ and this requirement is related to the process of fusion-fission of membranes. Thus, Golgi complex acts as a centre of reception, finishing, packaging and dispatching for a variety of cellular products.

c. Recycling of Membranes: Membranes do not arise de novo. Actually, membrane components flow by means of vesicles from ER through golgi to the plasma membrane. During exocytosis, the secretory granules become fused with the plasma membrane. To remove the excess membranes from the api-cal region of the cell, it has been suggested that patches of membranes are invaginated from the surface as small vesicles that move back into the golgi, to be reutilized in the packing of more secretion. Thus a dual func-tion at the cis and trans faces has been postulated.

d. Formation of Primary Lysosome: The golgi complex is involved in the formation of primary lysosomes and glycosylation of many lysosomal enzymes that are mostly glycoprotein in nature.

e. Formation of Melanin: According to Novikoff, the GERL of the golgi complex acts as the site for the formation of melanin granules; processing and packaging of secretory materials of many endo and exocrine cells.

f. Molecular Processing of Secretion: Many proteins are synthesized as biologically inactive precursors, which are activated later by the removal of a portion of the polypeptide chain by enzymes present in the Golgi apparatus. For example, the hormone insulin has two precursors – preproinsulin and pro-insulin.

NOTES

The former is activated in the ER into pro-insulin which is then activated in the Golgi through the removal of the C pep-tide by the converting enzyme and is pack-aged as active hormone insulin within the secretory granules.

g. **Acrosome Formation:** Electron micro-scopic studies have showed that acrosomes of spermatozoa are developed from golgi cisternae and vacuoles.

h. **Other Cellular Functions:** In addition, golgi apparatus is also involved in many other cellular functions, such as secretion of materials of primary and secondary cell walls in plants, secretion of cortical granules of a variety of oocytes, formation of yolk and vitelline membrane of growing primary oocytes.

Lysosomes

In the cytoplasmic matrix of the cells, there occur variously shaped bodies usually bounded by a single surface membrane and containing hydrolytic enzymes. These are called lysosomes. These enzyme containing bodies play important role in the digestion or lysis of intracellular substances, so they are called lysosomes. Lysosomes were first reported by Christian de Duve and co-workers in Belgium in 1955 following their extensive work on the biochemical identification of certain hydrolytic enzymes in the liver cells of rats. During that period pioneering works were carried out by Novikoff in America when he was able to complement the biochemical work on Lysosome with observations at light microscopic and ultra-structural levels.

Occurrence of Lysosomes

Lysosomes appear to be absent in prokaryotes. Cytochemical and electron microscopic studies have revealed the presence of membrane bound enzyme containing bodies in animal tissues. It is not certain whether the structures equivalent to animal lysosomes are present in plant cells.

The presence of lysosomes has been demonstrated in several slime moulds, fungal hyphae and algae. In 1964, P. Matile reported the occurrence of lysosomes in neumspora. Among the algae acid phosphatase has been located in lysosomes of euglena and a few other species.

Structure of Lysosomes

Lysosomes represent a class of morphologically heterogeneous cytoplasmic particles. The polymorphic nature of lysosomes has been attributed to specific functions, substances they contain and stage of digestion of those substances. Their size ranges from 0.25 to 0.8μ in diameter. In mammalian kidney cells they may be as large as 5μ and they may be even larger in phagocytes.

The density of lysosomes lies between that of mitochondria and microsomes (fragments of endoplasmic reticulum). Initially they were discovered by differential centrifugation and not by electron microscopes. They were originally obtained by sub-fractionation of the classical mitochondrial fraction. The classical mitochondrial fraction was sub-fractionated into the heavier and lighter components (Refer Figure 1.23).

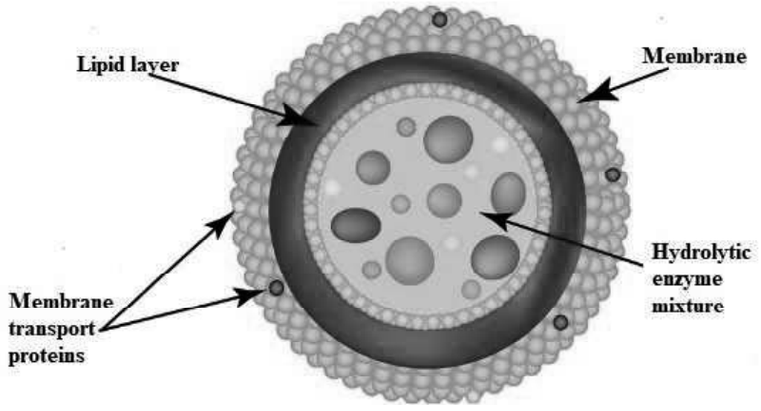


Fig. 1.23 *Structure of Lysosome*

The light mitochondrial sub-fraction or L-fraction was later named lysosome (lytic bodies). The lighter fraction (lysosomes) lacked cytochrome oxidase but had high concentration of acid phosphatase which became active only after severe mechanical disruption. The absence of cytochrome oxidase in lysosomes distinguishes them from mitochondria. Lysosomes are bounded by a single limiting membrane of lipoprotein that is homologous with the unit membrane. Chemically lysosome is a bag packed with a variety of hydrolysing enzymes. The surface membrane is impermeable or very little permeable to substrates of the enzymes contained in the lysosome. The internal organization is quite variable; some lysosomes have solid or very dense content, others have a dense outer zone with a less dense core and still others have a cavity or vacuoles.

On the basis of morphological and functional criteria, a variety of lysosomes can be recognised in different cells as well as within a single cell. Two basic forms of lysosomes have teen distinguished; (i) primary lysosomes; and (ii) secondary lysosomes.

(i) **Primary Lysosomes (Storage Granules):** They originate from the endoplasmic reticulum or are cut off indirectly from the tips of golgi saccules and have not yet been involved in digestive process.

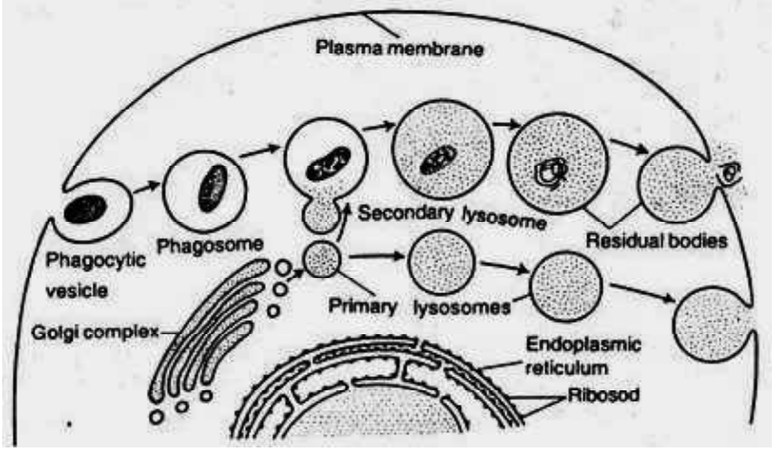


Fig. 1.24 *Formation of Primary and Secondary Lysosomes*

NOTES

(ii) **Secondary Lysosomes:** These are digestive vacuoles which are sites of digestive activity.

Lysosomes belonging to this group may be classified into two separate types:

(a) Heterophagic

(b) Autophagic on the basis of exo or endogenous origin of the material undergoing digestion

Each of these types may be further sub-divided as follows:

(i) Pre-Lysosomes whose enzymes have never been engaged in hydrolysis.

(ii) Lysosomes which are sites of present digestive activity.

(iii) Post-Lysosomes which have lost their enzymes.

Lysosomes start functioning when the cell takes up substances either by phagocytosis or pinocytosis. Undigested substances or residues within the lysosomes usually remain packed with different inclusions. Under certain circumstances lysosomes start digesting the very cell in which they occur by opening their membrane and setting free their digestive acid hydrolases. The lysosomes contain some 40 enzymes. Some of the important acid hydrolases identified from lysosomal fractions.

It is not necessary that all the hydrolytic enzymes should occur in a single lysosome. These enzymes have been located in a variety of tissue types. Lysosomes obtained from one source do not present a homogeneous picture under the electron microscope. This suggests that a particular lysosome may contain only a few or even a single species of enzymes. Some evidences do indicate that various hydrolytic enzymes are associated with different lysosomes in a variety of cells. These hydrolases have an acid optimum of about pH5. The enzymes of lysosomes become active only when the surface membrane is ruptured.

Functions of Lysosomes

The functions of lysosomes are:

1. Extracellular Digestion

Lysosomes are small bags containing digestive enzymes. They behave like tiny time bombs waiting for their explosion in the cytoplasm. When the limiting membrane ruptures, the digestive enzymes are released which take part in the digestion.

Sometimes lysosomal enzymes may be released outside the cell where they digest extracellular substances. Saprophytic fungi and other micro-organisms utilize extracellular digestion of complex substrates in the habitat and degrade them into simpler soluble forms which are then absorbed.

2. Intracellular Digestion

The digestive enzymes released in the cytoplasm may be involved in autophagy or heterophagy.

(a) Autophagy refers to digestion of endogenous materials or breakdown of molecules and pieces of cytoplasmic materials within the cell. The simpler

substances formed after the digestion are then utilized in the synthesis of some other substances. This recycling of cell component is called turnover. Actually the digestive materials are non-functional part of other organelles like mitochondria, endoplasmic tubules, enclosed in a vesicle or digestive vacuole called autophagic vesicle or autophagosome.

Primary lysosomes fuse with the autophagic vesicles as a result of which lysosomal enzymes are discharged into them. In this way secondary Lysosomes are formed. The presence of hydrolytic enzymes in the autophagosomes and progressive disintegration of enclosed organelles give indication of digestive activity. Autophagy may bring about cellular digestion after the death of a cell and so it brings about the self-clearance of dead cells. This is why Christian de-Duve called lysosomes the suicide bags.

(b) Heterophagy refers to intake of extraneous matter into the cell and subsequent break-down of that material by acid hydrolases. The bulk intake of exogenous material is called endocytosis (Refer Figure 1.25). The intake of liquid material is pinocytosis while the intake of solid matter is referred to as phagocytosis. The process of phagocytosis is illustrated by Figure. 1.25 below.

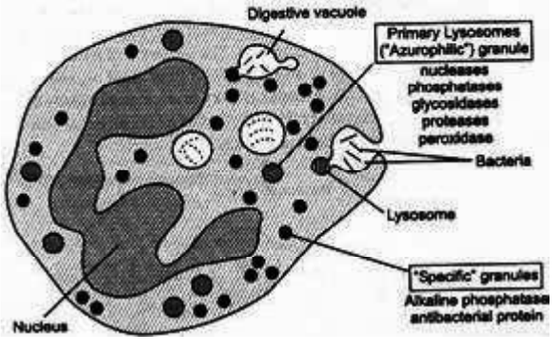


Fig. 1.25 *Phagocytosis in a Cell*

3. Role in the Release of Hormones

There is evidence that lysosomal acid hydrolases are involved in release of certain hormones from secretory cells of certain glands, e.g., thyroid hormones are released by hydrolysis of thyroglobulin.

4. Role in the Penetration of Sperm Nucleus into the Egg

The enzymes released from acrosome vesicle, the giant lysosomes of sperms, dissolve the cortical granules, the structure surrounding the egg nucleus and help in the penetration of sperm nucleus into the egg.

5. Role in Metamorphosis

During the development of embryo, several tissues become functionless which are digested by the enzymes released from lysosomes and the digested materials are then absorbed by the surrounding cells. Rudolph Weber has reported the absorption tail of tadpole in this way.

NOTES

6. Lysosomal Activity in Relation to Pathology

Several recent pathological studies have indicated that irregularities in lysosomal activity may cause fever, congestive heart failure, hepatitis, pylonephritis, hypertension, joint injuries, leucocyte granules and tissue injuries. Release of nucleases (DNAase and RNAase) which attack DNA and RNA respectively may cause chromosomal breaks and rearrangement and in this way it can result in structural abnormalities in chromosomes. The chromosomal abnormality may induce certain harmful mutations in characters. Sometime it may lead to carcinogenesis (cancer formation). Malignant cells are also found to have abnormal chromosomes. Partial deletion of chromosome 21 of man has been associated with chronic myeloid leukemia.

Accumulation of certain indigestible materials such as silica, asbestos particles, crystals of sodium ureate in the cells under certain conditions may result in cell inflammation. The inflammations of cells in such cases results due to release of enzymes after lysosomal break down in the cells containing ingested particles.

The actively dividing cells have been found to contain only a few lysosomes which are located generally near the periphery of the cell rather than near the nucleus. This suggests that the release of hydrolases from lysosomes has something to do with the division of cell.

7. Protection

Lysosomes of leucocytes help in defence against infection by bacteria and other microbes and guard against toxic molecules by digesting them. A mature leucocyte or white blood cell entering the circulation contains many lysosomes. During the life-time a leucocyte may ingest a foreign body such as a bacterium.

In this event, the lysosomes disappear releasing their enzymes into the digestive vacuoles containing the bacterial cell. The lysosomes rupture, bacterium is digested, and the leucocyte ultimately dies having performed its major function in the body.

In animal body, several kinds of cells have short lives such as, outer layer of skin and the mucous membrane lining the body. The short lived cells are being continuously replaced and the lysosomes of dead or degenerating cells release their enzymes into the body of the cells so that the whole cell may be digested. The process of tissue degeneration or necrosis can be attributed partly to lysosomal function.

Ribosomes

Ribosomes are cytoplasmic non- membranous ribonucleoprotein granules of 150 -200 A diameter. They have a typical binary and constricted structure with the two units being unequal in size. The prokaryotic and eukaryotic ribosomes are differentiated on the basis of the sedimentation coefficient. These are of two basic types, 70S and BOS ribosomes found in prokaryotes and eukaryotes respectively. The 'S' (Svedberg units) refers to sedimentation coefficient which shows how fast a cell organelle sediments in an ultracentrifuge. The heavier a structure the more is its sedimentation coefficient.

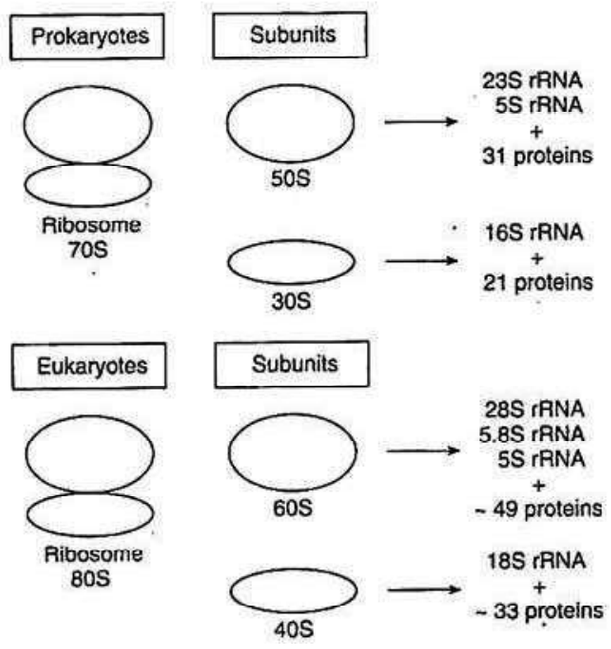


Fig. 1.26 *Composition of Typical Prokaryotic and Eukaryotic*

The 70S ribosome of prokary-otes is relatively smaller and consists of a large 50S subunit and a small 30S subunit. The ribosomes of eukaryotes are heavier and made up of large 60S subunit and small 40S subunit (Refer Figure 1.27).

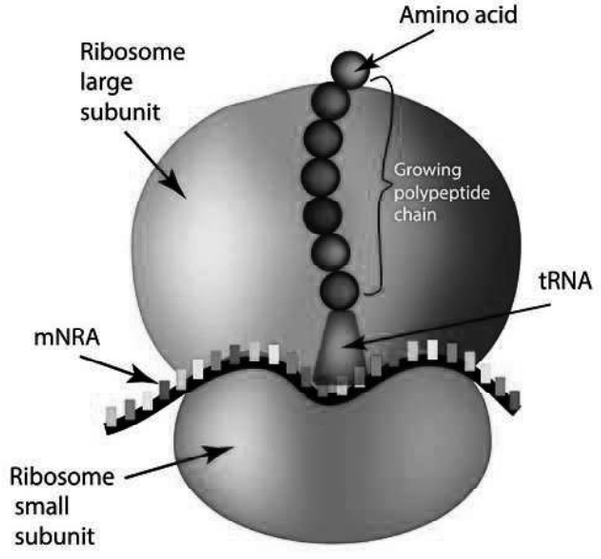


Fig. 1.27 *Structure of Ribosome*

Chemically ribosome is made up of rRNA, proteins and some divalent metallic ions. The subunits of ribosome can dissociate and asso-ciate on the basis of Mg++ concentration. Ribosomes may occur in the free form in prokaryotes,

NOTES

and are then called monosomes, or may be associated with mRNA to form poly-somes as in eukaryote. Ribosomes reported in plastids and mitochondria have a sedimentation coefficient of 70S and similar to prokaryotes in size and are different from cytoribosomes.

The structural model of ribosomes, as proposed by Stoffier and Wittmann, has the frontal face of the 30S subunit with its hollow facing the vaulted seat of the SOS subunit. The long axis of the BOS subunit is oriented transversely to the central protuberance of the SOS subunit. A tunnel is formed between the hollow of the small sub unit and the vaulted seat of the large subunit (Refer Figure 1.28).

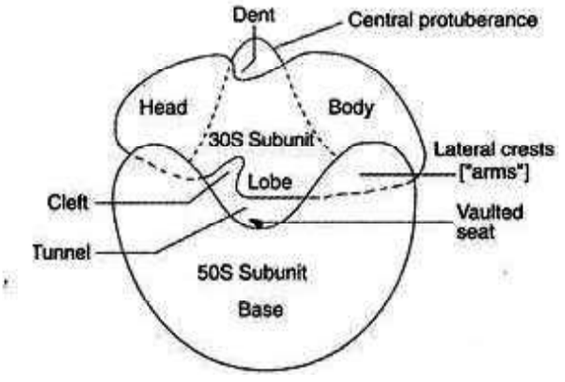


Fig. 1.28 *Stoffier and Wittmann Model of Ribosome*

Functions of Ribosomes in Cell

Ribosomes are the site of protein synthesis and are string together by mRNA to form polysomes or polyribosomes. Interaction of the tRNA-amino acid complex with mRNA, which brings about translation of the genetic code, is coordinated by the ribosomes.

During protein synthesis, the messenger RNA moves through a channel between two subunits of ribo-some. Each ribosome has two functional sites – amino acyl or acceptor (A) site and the peptidyl or donor (P) site (Refer Figure 1.29). The acceptor sites receive the tRNA amino acid complex, and the donor site binds the growing polypeptide tRNA. As such, they perform most important function in protein synthesis.

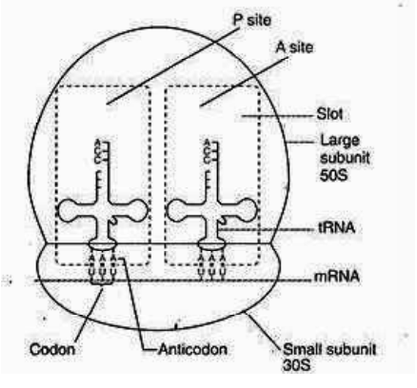


Fig 1.29 *Amino Acyl or Acceptor (A) Site and the Peptidyl or Donor (P) Site*

NOTES

Peroxisomes

In addition to lysosomes, a group of smaller particles than mitochondria and lysosomes are found in liver cells.These particles are rich in the enzymes peroxidase, catalase, D-amino acid oxidase and to a lesser extent, urate oxidase and are called peroxisomes (Refer Figure 1.30).

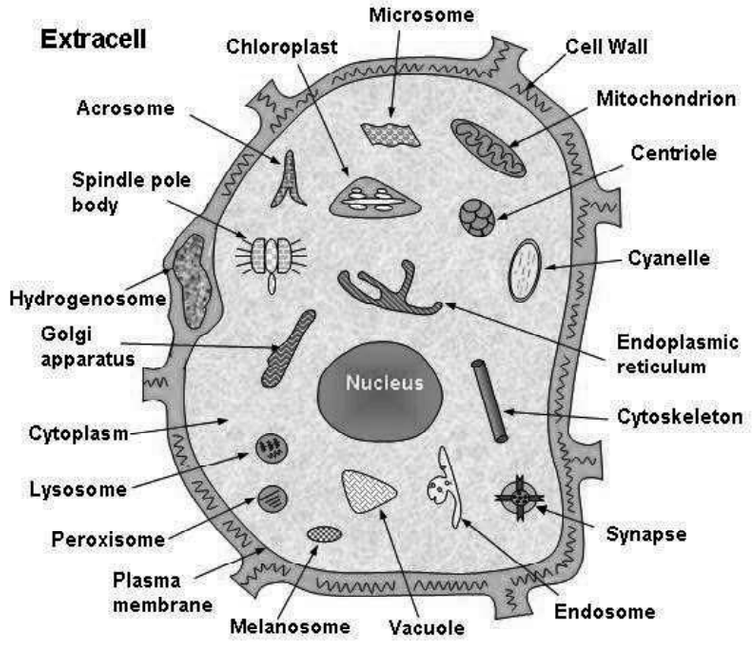


Fig 1.30 *Peroxisome*

Peroxisomes are found in liver, kidney, protozoa, yeast and many cell types of higher plants. Peroxisomes present in plant cells show some morphological similarities to the peroxisomes in animal cells. But plant peroxisomes have different enzymes including the enzymes of the glyoxylate cycle. Hence, their name is glyoxysomes.

Structure

Peroxisomes are ovoid granules limited by a single membrane. They contain a fine, granular substance which may condense in the centre, forming an opaque and homogeneous core or nucleoid.The average size of the peroxisomes in rat liver cells was shown to be 0.6 to 0.7 μm. The number of peroxisomes per cell varied between 70 and 100, whereas 15 to 20 lysosomes were found per liver cell. In many tissues peroxisomes show a crystal-like body made of tubular subunits.

In contrast to the nucleoid-containing peroxisomes found in liver and kidney, there are others which are smaller and lack a nucleoid. These are called microperoxisomes found in all cells and are related to endoplasmic reticulum. These may be considered as regions of ER in which catalase and other enzymes are found.

NOTES

Origin

Both types of peroxisomes are formed in the endoplasmic reticulum, and the enzymes they contain are synthesized by ribosomes bound to the granular ER. It is assumed that the peroxisome grows and is destroyed, probably by autophage after four or five days.

Peroxisomal Enzymes

The enzymes of peroxisomes are synthesized in the ribosomes attached to the rough endoplasmic reticulum. Liver peroxisomes contain four enzymes related to the metabolism of H_2O_2. Three of them are urate oxidase, D-amino oxidase, and oc-hydroxylic acid oxidase, which produce peroxide (H_2O_2), and fourth is catalase which destroys peroxide. Catalase is found in the matrix of peroxisomes and represents upto 40% of the total protein. H_2O2 is toxic so catalase destroys it and it probably plays a protective role. The enzyme urate oxidase, D-amino oxidase and hydroxylic acid oxidase present in amphibian and avian peroxisomes are related to the catabolism of purines.

Functions of Peroxisomes

Peroxisomes derive their name from their use of molecular oxygen for metabolic processes. These organelles are largely associated with lipid metabolism and the processing of reactive oxygen species. Within lipid metabolism, peroxisomes mostly deal with α–oxidation of fatty acids, the mobilization of lipid stores in seeds, cholesterol biosynthesis and steroid hormone synthesis.

β–Oxidation

The main reason for the high energy density of fats is the low proportion of oxygen atoms in every fatty acid molecule. For instance, palmitic acid, a fatty acid containing 16 carbon atoms and having a molecular mass of over 250 gms/mole, has only two oxygen atoms. While this makes lipids good storage molecules, they cannot be directly burnet as fuel or quickly catabolized in the cytoplasm through glycolysis. They need to be processed before they can be shunted into the mitochondria for complete oxidation through the citric acid cycle and oxidative phosphorylation.

When these molecules need to be oxidized to release ATP, they need to be first broken down into smaller molecules before they can be processed in the mitochondria. Within peroxisomes, long chain fatty acids are progressively broken down to generate acetyl coenzyme A (acetyl CoA) in a process called β–oxidation. Acetyl CoA then combines with oxaloacetate to form citrate. While most carbohydrates enter the citric acid cycle as a three-carbon molecule called pyruvate which is then decarboxylated to form acetyl CoA, peroxisomal β–oxidation allows fatty acids to access the citric acid cycle directly.

One of the main byproducts of β–oxidation is hydrogen peroxide which can be harmful for the cell. This molecule is also carefully detoxified by the enzyme catalase within peroxisomes.

Lipid Biosynthesis and Detoxification

In animal cells, peroxisomes are the sites for some amount of lipid biogenesis, especially of special phospholipids called plasmalogens that form the myelin sheath in nerve fibers. Peroxisomes are also necessary for the synthesis of bile salts. About 25% of the alcohol we consume is oxidized to acetaldehyde in these organelles. Their role in detoxifying and oxidizing a number of molecules, metabolic byproducts and drugs makes them a prominent part of kidney and liver cells.

Disorders Relating to Peroxisome Function

Disorders arising from deficient peroxisome function could arise from defects in peroxisome biogenesis, mutated peroxisomal enzymes, or non-functional transporters that recognize PTS1 and PTS2 in cytoplasmic proteins. The most severe of these are rare genetic disorders that result in impaired brain development and neuronal migration, along with myelin deficiency. Other organs affected include the skeletal system, liver, kidney, eyes, heart and lungs.

These disorders are usually caused by mutations in PEX genes, which are necessary for organelle biogenesis – from the formation of the subcellular membrane, to the recognition of cytoplasmic proteins and their import into the matrix of the organelle. For instance, PEX16 is involved in the synthesis of peroxisomal membranes, while PEX2 forms the translocation channel for the import of matrix proteins. PEX5, on the other hand is the receptor for recognizing the PTS1 signal sequence. Defects in these proteins can cause the accumulation of long chain fatty acids in blood plasma or urine as well as the inappropriate presence of phospholipids like plasmalogens in red blood cells.

1.2.4 Comparison of Plant and Animal Cells

Let's discuss the attributes of plant and animal cells in order to understand them.

Below is the table that explains the difference between a plant and animal cell. Figures 1.31 and 1.32 illustrates the structure of plant cell and animal cell.

Table 1.2 *Difference between Plant Cell and Animal Cell*

PLANT CELL	ANIMAL CELL
• A cell wall is present on the outside to provide shape and rigidity	• A cell wall is absent
• Plastids are present	• Plastids are absent
• Food reserve is starch	• Food reserve is glycogen
• Golgi apparatus consists of units called dictyosomes	• Golgi apparatus consists of single complex called Golgi body
• A mature plant cell possess large central vacuole	• A central vacuole is absent but small sap vacuoles may be present

NOTES

• Centrioles are absent except in lower forms	• Centrioles are present
• Spindle apparatus is anastral	• Spindle apparatus is astral
• Lysosomes are less common and their function is performed by vacuoles	• Lysosomes are abundant
• It does not take part in phagocytosis	• It may take part in phagocytosis
• Cells are fused in the region of middle lamella	• Cells are fused by means of junctions
• Plant cells don't burst if placed in hypotonic solution	• Animal cells placed in hypotonic solution will burst
• Glyoxysomes may be present in fat rich tissues	• Glyoxysomes are absent
• Cytokinesis occurs by cell plate method	• Cytokinesis occurs by cleavage

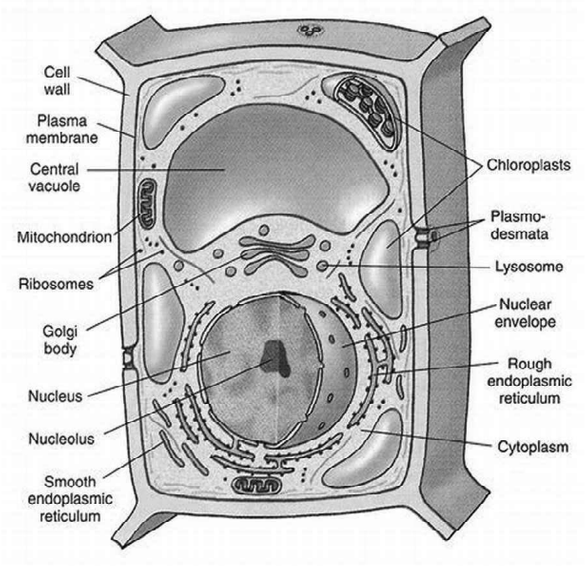


Fig 1.31 *Generalized Ultrastructure of Plant Cell with Cell Wall*

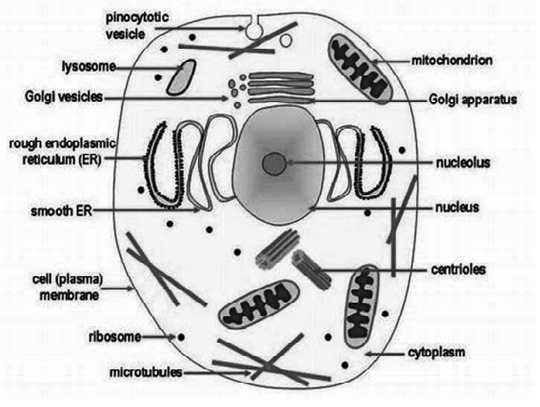


Fig 1.32 *Ultrastructure of Animal Cell under Electron Microscope*

Check Your Progress

1. Define cell.
2. Give the cell theory formulated by M.J. Schleiden and Theodore Schwann.
3. What are the different types of cells in an organism?
4. What is cytosol?
5. Define nucleus of a eukaryotic cell.

NOTES

1.3 OVERVIEW OF METABOLIC PROCESSES: CATABOLISM AND ANABOLISM

The sum of the physical and chemical processes occurring within a living organism is known as metabolism. Some other important terms related to metabolism are as follows:

- **Metabolic Pathway:** It is also known as a metabolic map and constitutes a series of enzymatic reactions to produce specific products.
- **Metabolite:** This term is applied to a substrate or an intermediate or a product in the metabolic reaction.

Metabolism can be further divided into the following categories:

1. **Catabolism:** It is also known as the process of the breakdown of molecules. The degradation processes concerned with the breakdown of complex molecules to simpler ones, with the release of energy, is known as catabolism. Catabolism occurs in three steps which are as follows:
   (i) **Conversion of Complex Molecules into Building Blocks:** This process occurs in the gastro- intestinal tract where complex molecules (polysaccharide, protein, lipids) are broken down into a simpler form (monosaccharides, amino acid, fatty acid, glycerol).
   (ii) **Formation of Simple Intermediate:** The simple molecules are absorbed into the intestine and are taken by the blood to the cell. In the cell, these are then converted into simple intermediates such as acetyl CoA.
   (iii) **Formation of Adenosine Triphosphate (ATP):** Actyl CoA is finally oxidised in mitochondria to release ATP.
2. **Anabolism:** Also called the process of building up molecules, this is the biosynthetic reaction that involves the formation of complex molecules from simple precursors in the presence of energy. The two steps involved in anabolism are the following:
   (i) This step precursors molecule like pyruvate and acetyl coenzyme (CoA) and intermediates the citric acid cycle and assembles them together.

NOTES

(ii) Then energy is utilized [as ATP or guanosine triphosphate (GTP)] and there is a reducing equivalent (as NADH).

3. **Amphibolism:** It is a term that is used for those reactions which are both catabolic and anabolic in nature.

Figure 1.33 illustrates catabolism and anabolism.

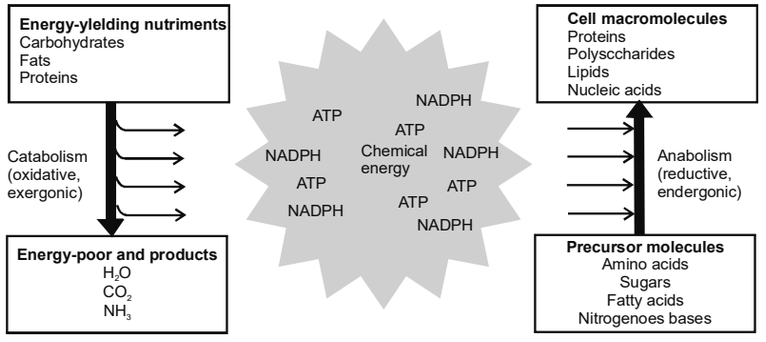


Fig. 1.33 *Catabolism and Anabolism*

1.3.1 Metabolic Diversity in Living Organisms

Most of the cells in living organisms have the similar sets of central metabolic pathways. However, this may also differ and different cells in different organisms may also possess alternative pathways. For example, organisms can be classified on the basis of the major metabolic pathways that they follow to gain carbon or energy. The classification that is based on carbon requirements identifies two major groups, which are as follows:

1. **Autotrophs:** They are the organisms that use only carbon dioxide as their sole source of carbon.
2. **Heterotrophs:** They are those organisms that require an organic form of carbon, like glucose, for synthesizing other essential carbon compounds.

The classification that is based on energy sources also identifies the following two groups:

1. **Phototrophs:** They are photosynthetic organisms that use light as the source of energy.
2. **Chemotrophs:** These organisms make use of organic compounds (like glucose) or oxidizable inorganic substances (like Fe^{2+}, NO_2^-, NH_4^+) or elemental sulphur as sole sources of energy. The energy is gained from oxidation–reduction reactions.

On the basis of these characteristics, every organism falls into one the four given categories, as shown in Table 1.3.

NOTES

Table 1.3 *Major Metabolic Pathways in Living Organisms*

Classification	Carbon Source	Energy Source	Electron Donors	Examples
Photoautotrophs	CO_2	Light	H_2O, H_2S, S, other inorganic compounds	Green plants, algae, cyanobacteria, photosynthetic bacteria
Photoheterotrophs	Organic compounds	Light	Organic compounds	Non-sulfur purple bacteria
Chemoautotrophs	CO_2	Oxidation-reduction reactions	Inorganic compounds: H_2, H_2S, NH_4^+, NO_2^-, Fe^{2+}, Mn^{2+}	Nitrifying bacteria; hydrogen, sulfur and iron bacteria
Chemoheterotrophs	Organic compounds	Oxidation-reduction reactions	Organic compounds, e.g., glucose	All animals, most microorganisms, non-photosynthetic plant tissue such as roots, photosynthetic cells in the dark

1.3.2 ATP: The Biological Energy Currency

The energy provided by the catabolic processes is largely used by anabolic processes in the form of ATP. For example, ATP is generated in energy-producing processes like photosynthesis and oxidative phosphorylation from adenosine diphosphate (ADP) and a phosphate ion.

$$ADP + HPO_4^{2-} \qquad ATP$$

In another way, the energy-consuming processes like biosynthesis, muscle contraction, etc. hydrolyse ATP into ADP and phosphate ion with a large amount of energy. Thus, ATP is the biological energy currency which couples the anabolic and catabolic processes together.

1.3.3 Different Types of Metabolic Reactions

There are four types of biochemical reactions, which are as follows:

1. **Oxidation–Reduction Reactions:** Oxidation refers to the loss of electrons or an increase in oxidation by a molecule, atom or ion whereas, reduction is the gain of electrons or a decrease in the oxidation state by a molecule, atom, or ion.

   Reduction

   $$Oxidant + e^- \rightarrow Product$$

   (gain of electrons) (oxidation number decreases)

   Oxidation

   $$Reductant \rightarrow Product + e^-$$

   (loss of electrons) (oxidation number increases)

   Elements that possess the ability to oxidize others are said to be oxidative and are called oxidizing agents. In these, the oxidant removes electrons from another substance and is itself reduced because it accepts the electrons. Therefore, it is also called an electron acceptor. For example, cellular respiration is the process of oxidation of glucose ($C_6H_{12}O_6$) to CO_2 and the reduction of oxygen to water. Its equation is as follows:

   $$C_6H_{12}O_6 + 6\ O_2 \rightarrow 6\ CO_2 + 6\ H_2O$$

NOTES

Elements that possess the ability to reduce others are said to be reductive and are also called reducing agents or reducers. The reducing agent transfers electrons to another substance and in the process is itself oxidized. Owing to the fact that it donates electrons, it is also known as an electron donor. For example, photosynthesis is a reduction reaction, whose equation is as follows:

$$6\ CO_2 + 6\ H_2O + \text{light energy} \rightarrow C_6H_{12}O_6 + 6\ O_2$$

2. **Rearrangement and Isomerisation Reactions:** A rearrangement reaction is a broad class of organic reactions where the carbon skeleton of a molecule is rearranged to give a structural isomer of the original molecule. It has been seen that a substituent moves from one atom to another in the same molecule. In the following example, the substituent R moves from carbon atom 1 to carbon atom 2:

$$\begin{array}{l} -C-C-C- \\ \quad R \end{array} \qquad\qquad \begin{array}{l} -C-C-C- \\ \quad\quad R \end{array}$$

3. **Group Transfer Reactions:** This refers to the transfer of a group from one compound to another, like the transfer of the acetyl group to oxaloactate, in order to form citrate.
4. **Making and Breaking of Carbon Bonds Reactions:** It refers to a chemical reaction in which carbon bonds are broken in the reactants and new carbon bonds are formed in the products.

Check Your Progress

6. What is a metabolite?
7. What are oxidizing agents?

1.4 ORIGIN OF LIFE

The term 'evolution' stems from the Latin term *evolutio* meaning 'unfolding'. Darwin has described evolution as 'descent with modification.' According to him, the present day complex living beings have evolved from simpler living beings of the past by gradual modifications over millions of years.

Natural environmental conditions always keep on changing. Living organisms have the ability to change themselves according to these changing conditions. This is known as adaptability of organisms. This adaptability leads to 'origin of new species'. Evolution is a very complex and extremely slow process, so it is impossible to see the changes occurring in organisms, but in general, the new species are better adapted to the environment and are more organized than their ancestors.

1.4.1 Theories of Evolution

Several theories have been put forward for the study of evolution. These have been categorized in the following four major categories:

1. Theory of Special Creation

Father Francisco Suarez (1548–1617), a Spanish priest, put forward the idea that the world was created in six days by the God.

2. Greek Theories

Several theories have been put forward by the Greek philosophers for evolution. Some of them are as follows:

i. **Thales (624–548 BC):** He propounded that life is aquatic or marine originated.

ii. **Anaximander (611–547 BC):** He proposed that all living beings have arisen from primordial fluid, to which they will ultimately return.

iii. **Anaximenes (588–524 BC):** He said that air is the source of life and proposed the theory of abiogenesis.

iv. **Empedocles (504–433 BC):** He talked about spontaneous generation and proposed that evolution is a series of attempts of nature to produce the more perfect forms.

v. **Aristotle (384–32 BC):** He proposed that there is a vital force which operates constantly to produce the more perfect forms. He placed the living forms of life in the increasing order of complexity and named the chain, 'ladder of life'.

3. Pre-Modern Theories

In the medieval age, the following evolutionists presented their theories on evolution.

i. **Francis Bacon (1561–1626):** He agreed with Aristotle and suggested that flying fishes come in between fishes and birds and bats between birds and quadrupeds.

ii. **Jan Swammerdum (1637–1680):** He proposed the preformation theory which says that ova contain miniature versions of adults in preformed state. When spermatozoa were discovered, they were described to possess miniature versions of embryos. This theory was discarded in 1759 when Casper Friedrich Wolff discovered that no miniatures were present in chick embryos.

iii. **Demaillet (1656–1738):** He studied the nature and formation of fossils and proposed that the terrestrial forms evolved from the marine forms.

iv. **Maupertius (1698–1759):** He was the first one to propose the general theory of evolution. He proposed that genetic material transfers from parent to offsprings.

v. **Bonnet(1706–1793):** He proposed the encasement theory. According to this, the miniature versions of the next generation of individual organisms pre-exist in the germ cells of mothers.

vi. **Linnaeus (1707–1778):** He is known as the father of taxonomy. He believed that species are fixed and created by the God.

vii. **Buffon (1707–1788):** He believed in the inheritance of acquired characters and the direct effect of environment.

NOTES

NOTES

viii. **James Hutton (1726–1797):** He postulated that the magma from volcanic activity comes on the surface of the earth which on solidification forms the new rocks. The activity of heat, air, cold, rain and ice erode the surface molecules of earth, water transports them and deposits them into layers which again form rocks.

ix. **Erasmus Darwin (1731–1802):** He was the grandfather of Charles Darwin. He gave the first clear statement on the inheritance of acquired characters.

4. Modern Theories of Evolution

There are various modern theories which have been discussed here in detail.

i. **Lamarck's Theory (1744–1829):** He elaborated the idea of Erasmus Darwin and said that the changes produced by the environment are carried by the offspring in successive generations. Though he was building on the work of his mentor, Count George-Louis Leclerc de Buffon, Jean-Baptiste Lamarck is often given credit for taking the first large advance toward modern evolutionary theory as he was the first to suggest a mechanism by which change in species could take place gradually. He explained the change over time, saying that though life began simple, it became complex over time. Also, he said that life started out simple and became more complex. In 1809, in his work, *Philosophie Zoologique,* he described a two- part mechanism using which change could slowly and steadily be introduced into the species and passed down, generation after generation. His theory is also referred to as the theory of transformation or Lamarckism. Though today Lamarck's work is considered a major step forward, in his lifetime he did not receive much recognition. The common example employed to explain the concept of use and disuse is the elongated neck of the giraffe. According to Lamarck, a giraffe could, develop an elongated neck over a lifetime because of continuous straining to reach the higher branches. However, the main downfall of this theory was that he could not explain how this could even though he discussed a 'natural tendency toward perfection.' Lamarck also used the example of the toes of water birds. According to him, due to years of straining their toes in order to swim through water, the toes of these birds became elongated and webbed to facilitate swimming. These two examples demonstrate how traits could be changed due to use. Lamarch also believed that disuse could, in a similar way, lead to reduction of the trait. The wings of penguins, for instances, are shorter than those of other birds because they do not use their small wings to fly.

   Lamarck believed that traits that changed or were acquired over an individual's lifetime could be inherited by its offspring. Giraffes with long necks would have offspring with long necks instead of the short necks that their parents were born with. This type of inheritance, at times called Lamarckian inheritance, has since been disproved by the discovery of hereditary genetics.

ii. **Theory of Catastrophism:** This theory explains that each creation is preceded by a catastrophe due to volcanic activity, upheaveling of earth,

NOTES

etc. This theory modifies the theory of *Special Creation* and states that all the creations of life by God, were preceded by a catastrophe resulting from some kind of geological disturbance. As per this theory, since each catastrophe resulted in complete destruction of existing life, each new creation consisted of life forms that were naturally different from those that existed previously. The main supporters of this theory were French scientists Georges Cuvier (1769-1832) and Orbigney (1802 to 1837).

iii. **Theory of Eternity of Life:** According to this theory, life existed in the past as it is in the present and will remain in the same form in future as well.

iv. **Darwin's Theory (1809-1882):** Darwin formulated the theory of 'origin of species by natural selection'. Charles Robert Darwin was an English naturalist who proved that all species of life have descended over time from common ancestors. He proposed the scientific theory that the branches of evolution resulted from a process known as natural selection. His arguments made the general public accept evolution as a fact in his lifetime. Not until the modern evolutionary synthesis emerged in the 1950s, did a consensus develop that natural selection formed the basis of the mechanism of evolution. The modified version of Darwin's theory explains the diversity of life and forms the unifying theory of life sciences. Darwin studied the transmutation of species and came up with the theory of natural selection in 1838. Although his ideas were discussed with other naturalists, he indulged in extensive research and his geological work had priority. In 1858, Alfred Russel Wallace worked Darwin and the two published their theories together. Darwin's work established evolutionary descent with modification playing the key role in nature's diversification. Darwin went on to analyse human evolution and sexual selection. According to Darwin's theory, all life is related, descending from a common ancestor. He believed that the birds, flowers, fruits, etc. were all related because he presumed that life developed from non-life. His stress was on an undirected and naturalistic descent with modification. As random genetic mutations take place within the genetic code of any organism, the advantageous mutations are naturally preserved because they facilitate survival. This process is referred to as 'natural selection'. These beneficial mutations get accumulated and give rise to a totally different organism, which is not just a variation of the original, but some being that is totally different)

v. **Weismann's Theory of Continuity of Germplasm:** According to this theory, the animal body is composed of germplasm and somatoplasm. The germplasm produces gametes which consist of characters of parents and the rest of the body is made up of somatoplasm. Over 100 years ago, the German biologist Weismann recognized that animals are made up of body cells (somaplasm), which comprise gamete-producing cells (germplasm). According to Weismann, the somaplasm provides the housing for the germplasm, and ensures that the germplasm is protected, nourished, and conveyed to the germplasm of the opposite sex to give rise to the next generation. He believed that the chicken is simply one egg's tool for laying another egg. Weismann also explained his theory for aging saying that once

NOTES

the opportunity to pass germplasm on has gone, there is no need to maintain the integrity of the somaplasm. Therefore, there is a decline in body function with aging .

vi. **De Vries Theory of Mutation:** He suggested that variations are sudden and large. He called these changes 'mutation'. In his theory of mutation, De Vries combines his theory of pangenesis (explaining heredity) and his theory of the possibility of new species arising only from a very large and completely spontaneous variation, which he termed as a 'mutation'. This mutation resulted from a new pangene or several new pangenes. De Vries contrasted his mutation theory with the Darwinian theory of selection. He emphasized that he saw the origin of the species through mutation whereas saw it through the selection of ordinary or fluctuating variation. In 1904 he made a lecture tour of the United States, where he expounded his theory.

vii. **Modern Synthetic Theory:** It accumulates facts and theoretical conclusions from a number of scientists. For example, Dbzhansky in his book '*Genetics and the Origin of Species*' explained the role of genetic changes. Julian Huxley and Ernest Mayr have explained the mechanism of origin of variations in higher animals. The Synthetic Theory of evolution or Modern Synthetic theory emerged around the middle of the 20th century from the ideas of three authors: Theodosius Dobzhandsdy (genetics), Ernst Mayr (species of living beings) and George G. Simpson (the great categories of the organisms).

We can summarize the important differences of the Modern Synthetic Theory of evolution and the Theory of Darwin in the following:

- It accepts the random genetic change as a mechanism of the important evolution, along with the mechanism of natural selection.
- It assumes that the traits are inherited through the genes. The variations of the population are due to the presence of multiple variations of a gene.
- It assumes that the specialization is due, usually, to small random changes in the genetic information.

This theory incorporates the aspects related to the microevolution and it assumes that the macroevolution is simply the accumulation of the microevolution. Concerning the evolutionary leaps and modern synthetic theory, the controversy is currently present because of the fossil registry, about which there is not a clear position within the scientific community.

There are some other theories like Haeckel-Ernst (ontogeny recapitulates phylogeny) and theory of orthogenesis.

1.4.2 Evidences of Evolution

There are several circumstantial proofs in favour of evolution. Some of these include the following:

1. **Morphology and Comparative Anatomy:** The study of various organs gives a clear idea of evolution. It suggests that similarity exists among organs of different animals based on common ancestry or common embryonic origin. This is known as homology; for example, homology in the limb and heart

structure of vertebrates and mouthparts and legs of insects. Homology is of three types:

i. **Phylogenetic Homology:** This refers to similarities between animals and plants of different species.

ii. **Sexual Homology:** This refers to similarities between the two sexes of the same species.

iii. **Serial Homology:** This refers to similarities between segments of the same organism, for example, all arthropods have segmented body with an exoskeleton of chitin.

Evolution can be defined on the bases of analogy. This means that the organs similar in appearance perform the same function, but their origin may be different. For example, the wings of insects, the wings of bats, and the wings of birds are analogous. It can also be defined on the basis of vestigial organs. A few organs are vestigial organs in our body. These do not have any function in our body, but they might have had important functions in our ancestors' bodies. For example, vermiform appendix in man, vestigial tail vertebrae in man, wisdom teeth, mammary glands in males and body hair.

2. **Evidences from Embryology:** Through observation, Haeckel discovered general resemblance between embryos of different groups of animals. All the multicellular organisms show a common pattern of development, for example, they start from zygote and then after repeated division, they form blastula and thereby two layered gastrula. The outer layer represents future ectoderm and the inner layer future endoderm.

3. **Evidences from Palaeontology:** It gives direct evidence of evolution. The study of fossils or ruminants highlights the differences between the shape, size and form of animals of the past and the present. For example, in Eocene period, the eohippus, the ancestor of horse, had five toes.

4. **Evidences from Taxonomy:** Naming plants and animals on the basis of their natural characters is a natural system of classification. This system is based on natural affinities between plants and animals. Thus, it can be concluded that all animals or plants have come from common ancestors.

5. **Evidences from the Connecting Links:** There are several plant and animal species, which have the characters of two different species, like amphibians can live on the water as well as on the ground. So, it can be concluded that they are the connecting link between the water and terrestrial species. Another example is virus; it also possesses the characters of both living beings and non-living beings, and thus can be considered as a connecting link.

6. **Evidences from Cytology:** Cell and cell organelles have almost the same structure and functions, like all cells have the basic pattern of organization and chromatin material is the essential component of nucleus of every living cell. This chromatin material organizes into chromosomes during cell division.

7. **Evidences from Genetics:** It has been confirmed that new generations inherit genes from their parents. Sometimes mutations produce changes in these genes and this leads to the differences in the characters of the old and the new.

8. **Evidences from Sources of Variations:** The differences between organisms in the same environment are called variations. No two organisms are alike. There are several sources of these variations, which are as discussed follows:

   i. **Environment:** The environment directly influences the organisms and induces various changes. For example, changes in the vegetation lead to changes in the environment. With the changing environment, changes in the eating habits, presence and absence of food etc. happen. Tower and Agar conducted several experiments showing the effect of temperature on the growth of rat and mice.

   ii. **Endocrine Glands:** The hormone secreted by the endocrine glands also produces variations in the organisms by changing their physical and mental characters during the development process.

   iii. **Mutation:** The term 'mutation' was introduced by Hugo De Vries. It refers to a hereditary change which occurs due to any change in the genes. New species arise suddenly as a result of mutation. The mutation theory has the following main features:

      (a) The individuals who develop from the result of mutation are called mutants; they are markedly different from their parents.

      (b) Mutation is inheritable.

      (c) Mutation is different from fluctuating variations.

      (d) Mutation can occur in any direction.

      (e) Mutation is subject to natural selection. The mutants, which are found unsuitable are discarded by natural selection.

      There are various bases to classify mutations. These include the following:

      (a) According to the Nature of Tissues: According to the nature of tissues, mutations can be classified into the following two types:

      - **Somatic Mutations:** These happen in the somatic cells of the organisms. They are non inheritable and are restricted to a few cells.
      - **Germinal Mutations:** These happen in the Germplasm or gamete. They are inheritable and thus get expressed in the next generations.

      (b) According to the Nature of Genetic Material: According to the nature of genetic material, mutations can be classified into the following two types:

      - **Point Mutation:** When there is sudden change in the structure or function of genes, it is called point mutation.
      - **Chromosomal Mutation:** When there is a change in the structure of chromosome or number of genes or arrangement, it is called chromosomal mutation.

(c) According to the Stages of Life Cycle: According to the stages of life cycle, mutations can be classified into the following two types:

- **Gametic Mutations:** These are inheritable mutations which either occur at the time of gamete formation or are introduced in the gametes.
- **Zygotic Mutations:** These inheritable mutations occur in the zygote.

(d) According to Mutagenic Effect: According to the mutagenic effect, mutations can be classified into the following two types:

- **Dominant Mutations:** The dominant mutations present themselves in the cells. These introduce dominant changes, like normal genes become recessive and mutant genes become dominant.
- **Recessive Mutations:** The mutant gene is recessive to normal and can exist only in homozygous conditions.

(e) According to their Significance: According to their significance, mutations can be classified into the following types:

- **Beneficial Mutations:** These are useful to organisms.
- **Lethal Mutations:** These types of mutations are lethal in homozygous conditions and are visible in heterozygous conditions.
- **Detrimental Mutations:** These are the recessive mutations which affect viability in homozygous conditions. The heterozygous are normal.
- **Biochemical Mutations:** These affect metabolic reactions and cause changes in the reaction intermediates and end products.

(f) According to their Size: According to their sizes, mutations can be classified into the following types:

- **Micro Mutations:** These do not produce significant changes.
- **Macro Mutations:** These have major phenotypic effects.

(g) According to their Method of Introduction: According to the method of introduction, mutations can be classified into the following two types:

- **Spontaneous Mutations:** These are the naturally occurring mutations.
- **Induced Mutations:** These mutations are caused by mutagenic agents like X-rays.

Mutations can occur either naturally or through some agents called mutagens. When the genes are reproducing themselves at that time there can be mutation. Mutations can also be induced by some agents like X-rays, UV-rays or other ionizing or non-ionizing radiation. These radiations can be natural or man made.

9. **Recombination:** Recombination is a process of physical breaking and eventual rejoining of DNA. Homologous recombination produces new combinations of DNA sequences during meiosis. These new combinations generally lead to positive effects and a changed effect on the phenotype. The advantageous combinations are favoured by the nature through natural selection.
10. **Genetic Drift:** Random fluctuation in gene frequencies, in a small population is called genetic drift. It occurs in nature by chance. It is an evolutionary force that changes genetic frequency by chance. As a result of genetic drift, any gene may get lost. The present change in the gene or mutation may get lost due to adaptability. Genetic drift can preserve or eliminate any gene without distinction. For example, assume that there is a small population of 100 individuals. Out of which 90 is AA and 10 is aa. If due to any reason 10 aa individuals die without producing any offspring, then AA constitutes 100% of the population because of complete loss of gene of aa.
11. **Natural Selection:** It is a natural phenomenon, which by promoting favourable features enables organisms to survive. Natural selection favours those characters in the genes which make them more efficient to communicate into their environment. Due to this, the next generation is always much more adaptive to the environment and possesses some new characters which may or may not be recognized in the generation at that time. For natural selection, the following must be true:
    i. The organism or the species should be capable of reproduction.
    ii. The organisms should possess variations in inheritable characters.
    iii. More than one unit should be present in the same environment.

    The main categories of natural selection include the following:
    - **i. Stabilizing Selection:** It occurs when the environment is nearly stable and the species are already adaptive to it. Any mutation in this condition is bound to get discarded. As a result, the population is genetically constant in this case.
    - **ii. Directional Selection:** This occurs when the environment changes itself in a particular direction and the population also changes in that direction so as to adapt itself to the new environment.
    - **iii. Disruptive Selection:** The disruptive selection involves breaking up a homogeneous population into different adaptive forms. It breaks up the homogeneous population into several phenotypes.
    - **iv. Cyclic Selection:** When the environment is not stable, generations may show fluctuation in phenotype and genotype because the selection occurs in one direction in one generation and season and in the opposite direction in the next season.
12. **Migration:** It is a natural phenomenon, in which the organisms (mainly birds and mammals) migrate from one place to another, in search of right living and breeding conditions and food. The migration occurs through a well-defined route either in groups or individually.

Migration can be on daily or seasonal basis either in search of food or to perform other daily activities, like sea birds migrate to the shore to breed and African antelopes migrate seasonally to avoid draught.

NOTES

1.4.3 Concept of Species

Ernst Mayr defined species as 'groups of interbreeding natural populations that are reproductively isolated from other such groups'. A species is derived from its own lineage and has its own history of development. Members of a single species can only interbreed with one another and produce offspring.

Specification and Isolation

Isolation refers to segregation of similar individuals into two different species. This is done to prevent them from interbreeding. Dobzhansky classifies isolation into two types:

1. **Geographic Isolation:** In this type of isolation, individuals of the same species are isolated by physical or geographical barriers, which may include mountains, deserts, forests, water, etc. As different individuals occupy different geographic regions, there is no chance of interbreeding. Moreover, as the environmental conditions are different in different places, the single gene pool gradually bifurcates into two gene pools.
2. **Reproductive Isolation:** For the reproductive isolation there are two views; according to Muller, when the population occupies different geographical regions, it isolates into species and subspecies and originates into a different population. According to Dobzhansky, reproductive isolation occurs due to changes in the complementary gene complexes of the population by natural selection.

1.4.4 Chemical Evolution and Rise of Living Systems

There are various theories about the origin of life on earth, like life originating on earth rather than coming from some heavenly bodies, or from non-living matter, or from some pre-existing life. Various such theories were formulated, but none was able to describe where the first life came from. Charles Darwin provided an important hint to this problem in a letter to Joseph Hooker in 1871, in which he imagined that if we could provide all ammonia, phosphoric acids, light, heat, electricity, and so on to a warm pond, it could lead to the chemical production of some protein compounds, which could then lead to some complex changes. This kind of setup is not possible now after living creatures have formed. This hint gave an idea that life could have arisen from non-living matter as the environmental conditions present at that time were very much different from the ones now. This idea was expanded into the chemical theory independently by two workers: Russian biochemist Aleksandr Ivanovich Oparin in 1923 and English biologist J. B. S. Haldane in 1928. This was summarised by Oparin in his book, *Origin of Life* in 1936 as 'abiogenesis first, but biogenesis ever since'. Oparin and Haldane held that:

- Present environmental conditions cannot give rise to spontaneous generation. There is a huge difference between the Earth's surface and atmosphere now and a billion years ago.

NOTES

- The previous chemical and physical conditions of the Earth explain the chemical evolution of organic chemicals.
- The initial atmosphere of Earth was a reducing one.
- A progressive series of chemical reactions occurred, which led to the production of a collection of chemicals and gave rise to life. Atoms combined into inorganic molecules, which turned into simple organic compounds (monomers) that formed complex organic compounds (polymers). They accumulated and finally organised into living matter. This abiogenesis occurred before 4 billion years ago.
- The earth's solar heat radiation and lightning provided the energy for molecule evolution.

Evolution of Chemicals

The four major events that occurred in chemical evolution are:

- **Formation of Inorganic Matter:** Early Earth had a very high temperature (4000–8000 °C).As the earth's temperature cooled, it favoured the formation of molecules by atoms. So, the atoms of various different elements came together and formed different inorganic molecules. The first molecules were formed by the atoms of the lighter elements like hydrogen, nitrogen, oxygen, and carbon, which were present in abundant quantity in the primitive atmosphere. So, from this explanation it can be understood that the first molecules must be the gases like, hydrogen (H_2), nitrogen (N_2), ammonia (NH_3), carbon dioxide (CO_2), methane (CH_4), and water vapour (H_2O). Volcanic eruptions, sunlight, and lightning provided the energy needed for this reaction to occur.
- **Spontaneous Formation of Simple Organic Molecules (Monomers):** The inorganic molecules formed later interacted with one another to form inorganic molecules like sugars, fatty acids, glycerol, amino acids, and organic bases (purines and pyrimidines). Energy for such a creation was provided by solar radiation like ultraviolet rays and electric discharges like lightning. Such conditions of chemical synthesis cannot be repeated as oxygen in the environment now creates an ozone layer which shields the earth from UV rays. Simple organic molecules occurred partially in the primitive atmosphere, but they mainly resided in the oceans. When atmospheric water condensed on further cooling of the earth, these inorganic precursors were brought down with rain. These precursors were also dissolved from the rocks by rain water.These inorganic molecules and their organic molecules get dissolved and then mixed with the sea water. The main factors which are responsible for the accumulation of these organic molecules are the lack of life and the reducing atmosphere of the atmosphere. If life had existed, it would have led to the consumption and alteration of these compounds by microorganisms. But, as there was no one to consume them, the compounds continued to react. Also, reducing the environment preserved these compounds. Stanley L. Miller, a biochemist, and Harold Urey, an astronomer, experimentally demonstrated in 1953 that simple organic compounds can be formed in nature, as was explained. The Miller and

Urey experiment was performed with a glass apparatus having a gas flask, a condenser, and a liquid interconnected with tubes and fitted with sources of energy. Then the apparatus was given a condition like that of the primitive earth, reducing the environment and ocean. The apparatus circulated methane (CH_4), ammonia (NH_3), and hydrogen (H_2) in a 2:4:1 ratio, as well as water vapour (H_2O) at 800 °C. The gases were given the same atmosphere as that of the primitive atmosphere. Energy from the interaction of the gases was also provided in the form of electric pulses of 75,000 volts from electrodes in the gas flask. The sparks stimulated lightning. The gases were then condensed through a condenser in a narrow tube and passed through a liquid flask. Here, with an electric heater, heat energy was provided, which stimulated volcano-like conditions. As a result, primitive earth conditions were created in the apparatus. The experiment was continued for 18 days. It was expected that small organic molecules would be created in the atmosphere (gas flask) and, with the rain (condensation), be carried to the ocean (liquid flask). The chemical compositions of the reaction products from the mixture were analysed by chromatographic and calorimetric methods. Various simpler organic compounds like amino acids, such as glycine, alanine, and aspartic acid, adenine, and simple sugars like ribose were found. Various similar experiments were performed and the reaction products containing amino acids found in proteins, nitrogenous bases of nucleic acids, nucleotides like ATP, fatty acids and glycerol were found. In a similar experiment, Orgel successfully found six nucleotides containing nucleic acid.

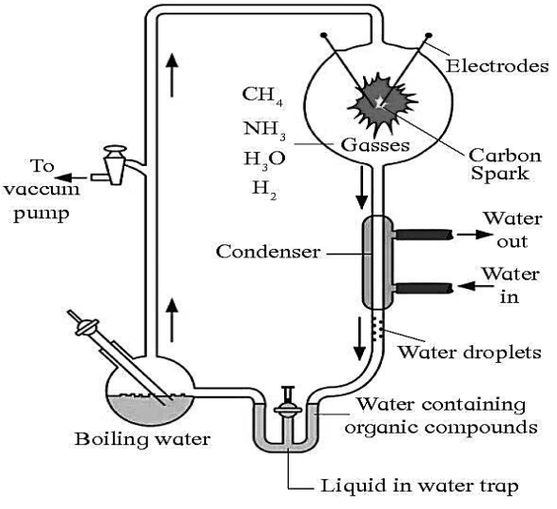


Fig. 1.34 *Miller's Experiment*

- **Spontaneous Formation of Complex Organic Compounds (Polymers):** Small, simple organic compounds came together and bonded to form large, complex organic compounds. Amino acids bonded to form polypeptides and protein molecules, simple sugars combined to from large sugar or starch molecules, fatty acids and glycerol connected to form fats and nitrogenous based with sugars and phosphate group form nucleotides which polymerized to form nucleic acids. Sun provided the heat energy for the formation of

complex organic compounds. The accumulation of this compounds in sea led sea to become a hot, dilute broth in which the molecules further collided to form new molecules of increasing size and complexity. In the underwater rocks, complex charged clay and other minerals surface are also present which attracted these organic molecules. The charged surface held molecules together and also, provided the catalytic surface for the chemical reactions and selection of the best bond. This is preferred to as the first selection process of the biological suitable molecules. Today also, mineral catalyst is used by chemical industries for the production of organic molecules and enzyme also use mineral ions as their cofactors.

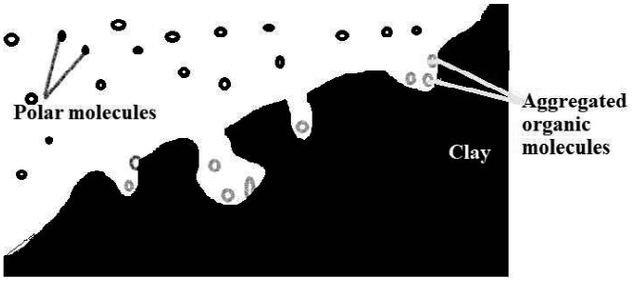


Fig. 1.35 *Concentration of Organic Molecules by Clay*

The experimental evidence that simple molecules form complex organic molecule came from the Fox's experiment. S.W. fox in 1950, took the mixture of amino acids and heated them to 139°C to 180°C for several hours and then cooling them in water which led to the synthesis of polypeptides. These molecules were termed by him as proteinoid microspheres which were 1-2 μm in diameter. They can also be induced to constrict. When inorganic phosphate and hydrogen cyanide reacted, they formed nucleotides and ATP. He also showed suggested that waves and rain in primitive environment splashed organic monomers on fresh lava or hot rocks, which would have favored the formation of polymers abiotically. When he tried to do this in his laboratory, nucleic acids were produced by heating the mixtures of nucleotides and phosphate to about 68°C for some time.

- **Spontaneous Formation of Molecular Aggregates:** Large organic molecules according to Oparin and Sydney fox, came together spontaneously on the primitive earth due to intramolecular interactions and formed large colloidal aggregates which were termed as coacervates.

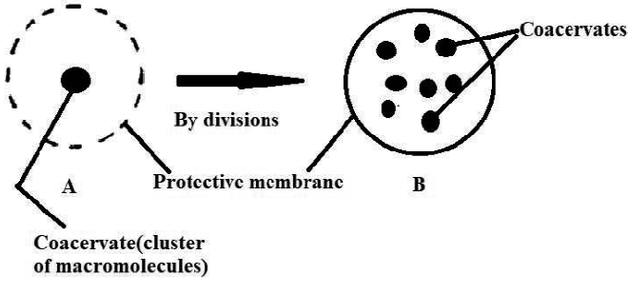


Fig. 1.36 *Coacervate Formation*

An envelope of water molecule was formed around each of these aggregates because of the zwitterionic nature of protein (Zwitter ion is the ion carrying both positive and negative charge). The envelope protected the aggregates and also increased the chances of chemical reactions between them by keeping them closer. Continuous decomposition and combination reaction started occurred in which the former provided the energy for the latter. The proteins and other materials were selectively absorbed from the ocean by these aggregates. This led to the increase in both the size and internal complexity of the coacervates. Due to the favorable internal organization of some aggregates, they were able to absorb the materials more easily and were enlarged more quickly as compared to others became dominant. So, it an be said that natural selection occurred in chemical evolution too had led to the formation of chemical molecule or aggregates which were more similar. Further, increase in size divided the large coacervates into small ones. The ones that multiplied and grew more quickly got the more availability of organic materials from the environment. Rapid growth required efficient reactions with the absorbed compounds for increasing the colloidal material to the dividing level.

Abiogenesis

Origin of life from these non-living chemical compounds is thought to have occurred in two stages:

- From coacervates to first primitive living systems called protocells or eobionts or protobionts
- From eobionts to the first living cells

Formation of Eobionts

As elaborated in previous sections, coacervates are in continuous process of breaking down and building up or it can be said that they are in the state if dynamic equilibrium. Degraded materials inside them needs to be removed and they require the constant intake of new materials for their growth. Coacervates have all the basic properties of living cells like metabolism, growth and reproduction but lack the complex molecular organization, catalytic proteins (like enzymes), electron transport chain and control of nucleic acids. Then, with the growth nucleic acids took the control of coacervates. After it, duplication become detailed and genetic material got its existence. These led to the conversion of coacervates into the first primitive living systems called eobionts. This first noncellular form of life could have originated about 3 billion years ago. The non-cellular forms would have been the huge molecules (RNA, proteins, polysaccharides etc.). As now it is known that RNA molecules have catalytic properties too, so before these enzymes would have evolved, RNA molecules must be used for the catalytic property. RNA could catalyze its own replication too. It is also held that coacervates cannot be the precursors for the eobionts as they lack lipid membrane. Mixture of some artificially synthesized organic compounds form microspheres. If this mixture contained lipids, the microspheres could acquire the lipid membrane around them. Such microspheres are the best suitable precursors for the development of eobionts.

Formation of First Cells

Once the complex protein structures have been formed just by some chance, this may have the property for catalyzing the chemical reaction or can act like enzymes. Not all proteins can act like enzymes but if by some molecular organization the protein got the catalytic centres, it can act like an enzyme. Enzymes can speed up the reaction by many folds. So, the presence of catalytic proteins can led to the synthesis of various molecules in the eobionts. This might have made eobionts capable of producing more food and energy than others. Continuous selection of eobionts must have made them more perfect with time.

Nucleic Acids

From the nucleotide precursors, polynucleotides were formed which on polymerization gave nucleic acids like Deoxyribonucleic acid (DNA) and Ribonucleic acid (RNA). First RNA developed and then DNA. The two then started to control the protein synthesis process. Then, DNA acquired the ability to mutate. This mutation was influenced by the high temperature on the early earth. Thermal mutations continued to affect the structure of proteins which changed their activity and function. Some changes were favourable and increased the ability of eobionts to compete with the organic molecules and led to the production of more and more organic compounds from them.

Bio Membrane

As the protein and lipid molecules were synthesized in the eobionts, the lipid-protein, membrane was formed around the eobionts. This membrane allowed the accumulation, infusion and outflow of the molecules in them. This aided in their ability to survive, grow and compete. So, as the growth of eobionts proceeded, growth and division is them also become regulated and precise. Due to this, the first living cell or organism was produced. It was about 2000 million years ago. These were probably the single cells. So, it can be concluded that the life arose from the inorganic molecules of the primitive earth (abiogenesis) and then it led to biogenesis.

Free atoms (hydrogen (H), carbon (C) and nitrogen(N))

↓

Inorganic molecules (hydrogen (H_2), nitrogen(N_2), ammonia (NH_3), carbon dioxide (CO_2), methane (CH_4), and water vapour (H_2O))

↓

Simple organic molecules

$CH_4, H_2O \rightarrow$ Sugars, Fatty acids, Glycerol

$CH_4, H_2O, NH_3 \rightarrow$ Amino acids

$CH_4, H_2O, NH_3, HCN \rightarrow$ Nitrogenous bases (Purines, Pyrimidines)

↓

Complex organic molecules

Sugar+ Sugar $\rightarrow$ Polysaccharides

Fatty acids + Glycerol $\rightarrow$ Fats

Amino acid + Amino acid ⟶ Proteins

Nitrogenous bases + Sugars + Phosphates ⟶ Nucleotides, Adenosine phosphates

Nucleotides + Nucleotides ⟶ Nucleic acids

↓

Coacervates or microspheres

(Aggregates of large complex organic molecules capable of growth and division)

↓

Eobionts

(Nucleic acid- Controlled coacervates)

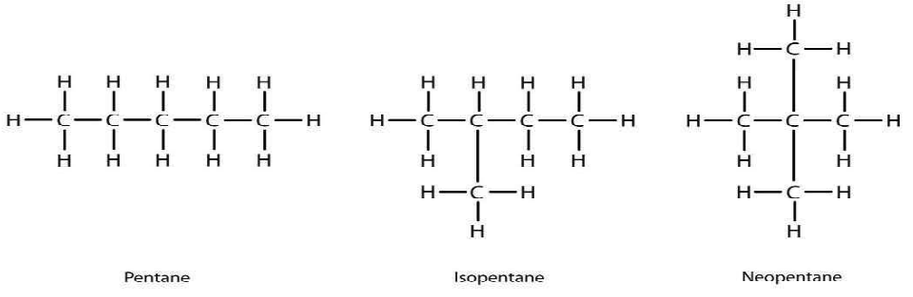
↓

First primitive cells

(Lipid- protein membrane bound units with enzyme – controlled metabolism and nucleic acid regulation without organized nucleus.)

1.5 UNIQUE PROPERTIES OF CARBON

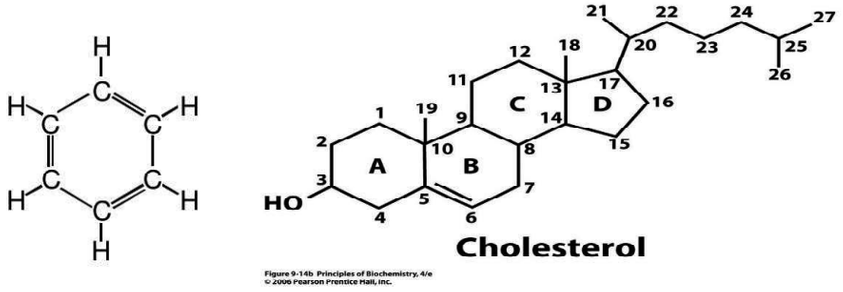
Carbon is one of the six elements that form the major framework of living matter. It is the most prime and versatile element of life. Its presence is found in all the known living systems, which means that no life can exist without it. Carbon has four electrons in its outer valence shell, and thus it forms four covalent bonds. A bond can be formed between carbon atoms through a property called catenation, and also with many other elements and molecules. Carbon-carbon bonds are very strong and stable bonds. Due to this bonding property of carbon, it can form an infinite number of molecules, which are the fundamental components of living things. The molecules can be bonded in various forms, for example, in the form of a ring structure or various rings bonded with one another, which presents the diversity of molecular forms. These molecular forms make a diverse range of macromolecules of different functions. Hydrocarbons are organic compounds composed entirely of carbon and hydrogen atoms. They are the components of various important macromolecules found in living systems. Figure 1.37 shows examples of a variety of molecules with different structures formed due to the different bonding patterns shown by carbon.



Pentane Isopentane Neopentane

(A)

NOTES



(B) (C)

Fig. 1.37 *Variety of Molecules*

A. Hydrocarbons, n-pentane (C_5H_{12}) with a straight chain structure, and isomers of n-pentane, isopentane (C_5H_{12}) and neopentane (C_5H_{12}) with branched chain structures, all with carbon bonded with a single covalent bond.

B. The hexagonal planar ring structure of the hydrocarbon Benzene (C_6H_6) has six carbons bonded by three single and three double bonds, and six hydrogen atoms all bonded with covalent bonds.

C. The fused ring structure of cholesterol ($C_{27}H_{45}OH$), a steroid with three regions: 1. A hydroxyl group (attached to C_3)$_2$. A four-ring structure containing hydrocarbons (C1-C20). and 3. C21-C26 hydrocarbon tail.

Carbon is contained in almost 90% of the compounds found in living systems. It forms the backbone of all organic compounds, which are the building blocks of living organisms. Important carbon compounds like nucleotides, nucleic acids, lipids, carbohydrates, amino acids, proteins, play various important roles in living beings like metabolism, respiration, building of structural parts of cells, which in turn form tissues and organs, providing energy for cellular functions, storing energy and much more. Carbon dioxide, the vital element of respiration, is also a form of carbon. Thus, carbon is an integral part of biological processes that play a very important role in the survival of living organisms. Carbon cycles through different biochemical pathways between the atmosphere and the biomass of living things. This shows how carbon atoms migrate from the atmosphere to species and back to the atmosphere over and over again by the carbon cycle.

Check Your Progress

8. State the theory of evolution by Empedocles.
9. What do you mean by species?
10. Who expanded the chemical theory of evolution of life?
11. What are the stages of origin of life from the non-living chemical compounds?
12. Why is carbon able to form infinite bonds?

1.6 INTRODUCTION TO BIOMOLECULES

NOTES

The term 'biomolecule' refers to any organic molecule that is formed by a living organism. This includes large polymeric molecules like proteins, polysaccharides and nucleic acids and small molecules like primary metabolites, secondary metabolites and natural products. Any living system is primarily composed of six elements, which are as follows:

- Carbon
- Hydrogen
- Oxygen
- Nitrogen
- Phosphorus
- Sulphur

There are some other functional elements such as calcium (Ca), sodium (Na), magnesium (Mg), copper (Cu), zinc (Zn), etc. as shown in Table 1.4. Among these, carbon (C) is the most important element.

Table 1.4 *Chemical Compositions of Important Biomolecules*

Element	Per cent by mass
Oxygen	65
Carbon	18
Hydrogen	10
Nitrogen	3
Calcium	1.5
Phosphorus	1.2
Potassium	0.2
Sulfur	0.2
Chlorine	0.2
Sodium	0.1
Magnesium	0.05
Iron	0.004
Cobalt, Copper, Zinc, Iodine	< 0.05 each
Selenium, Fluorine	< 0.01 each

Carbon (C) is the most important owing to its property to form an infinite number of compounds. This helps carbon to form stable covalent bonds and C–C chains of unlimited length. Carbon has a unique property called catenation, that is, the property of forming bonds with the same element. It is tetravalent in nature and can form single, double and triple covalent bonds to combine with hydrogen, oxygen, sulphur, nitrogen and chlorine to form various types of compounds.

Every living system is built up in a hierarchical manner. Multicelllular organisms are made up of organ systems, which are constituted by tissues. These tissues are a collection of cells, which contain organelles. It is after this that the world of biomolecules appears (Table 1.5). Macromolecules (proteins, lipids, nucleic acids and polysaccharides) are formed from monomeric units or building blocks (amino acids, fatty acids, glycerol, nucleotides and monosaccharides).

NOTES

Table 1.5 Chemical Compositions of Macromolecules

Macromolecule	Monomers	Major Functions
Proteins	Amino acids	Fundamental basis of structure and function of cell (static and dynamic function)
Polysaccharide	Monosaccharides	Form of storage of energy to meet short-term demands
Lipid	Fatty acids, glycerol	Form of storage of energy to meet long-term demands; structural components of membranes
Nucleic acids	Nucleotides	Carrier of genetic information

1.6.1 Building Blocks of Bio-Macromolecules

A macromolecule is defined as very large molecule, for example a protein molecule which is composed of thousands of covalently bonded atoms. Most of the macromolecules are considered as polymers which are made of smaller molecules termed as monomers. In biochemistry, the commonly known macromolecules are biopolymers, namely nucleic acids, proteins, and carbohydrates. The building blocks in biology are referred as the foundation of all living organisms. These are the basic and essential molecules that constitute the macromolecules, cells, tissues, organs and hence the organ systems. The smallest element that contributes to the formation and structure of a living organism is scientifically termed as a biological building block. Several such biological building blocks can be specifically combined simultaneously in various forms or patterns in order to obtain the necessary results. For example, a complex bond of nucleic acid, phosphate and sugar (considered as the various biological building blocks) are responsible for the construction of a DNA molecule. Characteristically, the biological building blocks are typically categorized into four main categories, based on their specific properties and functions. These four main biological building blocks are Lipids, Carbohydrates, Proteins and Nucleic Acids:

- Lipids are classified as a significant group of naturally occurring organic compounds, which are related by their insolubility in water, and solubility in non-polar solvents, such as alcohols and ether. Lipids are utilised for the storage of energy, protection of internal organs, and in the formation of structural components of cells.
- A carbohydrate is a biomolecule consisting of Carbon (C), Hydrogen (H) and Oxygen (O) atoms. Basically, the carbohydrates are the sugars, starches and fibres commonly found in fruits, grains, vegetables and milk products. Carbohydrates are considered as the energy source essential for muscles.
- Proteins are one of the most vital biological building blocks. Proteins are formed by linking the amino acids. It has a significant role in the formation of tissues and in defence.
- Nucleic acids are large and complex molecules that are present in cells and viruses. Basically, the nucleic acids are of two types, DNA (Deoxyribo Nucleic Acid) which is responsible for the storage of hereditary information and RNA (Ribo Nucleic Acid) which is responsible for expressing the stated

information by means of protein synthesis. The nucleic acids are made up of very small or tiny units termed as 'Nucleotides', which are made of 5 carbon sugar attached to a nitrogen containing aromatic base and phosphate groups. Each nucleic acid contains 4 nitrogen-containing bases out of 5 namely Adenine(A), Guanine(G), Thymine(T), Cytosine(C) and Uracil(U). All nucleic acids contain A, G and C, whereas the presence of T and U depends on the type of nucleic acid.

Check Your Progress

13. What are the components of macromolecules?
14. What is the smallest element that contributes to the formation and structure of a living organism called?

1.7 ANSWERS TO 'CHECK YOUR PROGRESS'

1. By definition, a 'cell' is the fundamental and structural unit of all living organisms. It is the smallest biological, structural and functional unit of all plants and animals. Therefore, cells are called the 'building blocks of life' or the 'basic units of life'.
2. In 1838, M.J. Schleiden and Theodore Schwann formulated the 'cell theory'. The cell theory formulates the following three principles:
   a. All organisms are composed of cells.
   b. Cell is the structural and functional unit of life.
   c. Cells arise from pre-existing cells.
3. There are many different types of cells in the organism, such as blood cells, skin cells, bone cells and even bacteria. Based on whether they have a nucleus, the cytologists recognize two basic types of cells as Eukaryotic cell and Prokaryotic cell.
4. Cytosol is the liquid part filled inside the cell and it contains water, salt, and macromolecules (Protein, Lipid, RNA). It has an array of microtubule fibres running throughout the cytosol to give vesicular structure to its destination. Besides this, cytosol exhibits 'Sol' to 'Gel' transition and such transition regulates multiple biochemical and cellular processes.
5. Nucleus is the central processing unit of Eukaryotic cell and homologous to the processor in a typical computer. The liquid filled inside nucleus is called as 'nucleoplasm'. It is a viscous liquid containing nucleotides and enzymes to perform replication, transcription, DNA damage repair, etc. It contains genetic material (DNA) in a complex fashion involving several proteins (histones) to pack into nuclear bodies or chromosomes.
6. A metabolite is a substrate or an intermediate or a product in the metabolic reaction.
7. Elements that possess the ability to oxidize other elements are called oxidizing agents.

NOTES

8. Empedocles. (504–433 BC) talked about spontaneous generation and proposed that evolution is a series of attempts of nature to produce the more perfect forms.
9. Ernst Mayr defined species as 'groups of interbreeding natural populations that are reproductively isolated from other such groups'. A species is derived from its own lineage and has its own history of development. Members of a single species can only interbreed with one another and produce offspring.
10. The chemical theory of evolution of life was expanded independently by two workers: Russian biochemist Aleksandr Ivanovich Oparin in 1923 and English biologist J. B. S. Haldane in 1928.
11. The origin of life from the non-living chemical compounds is thought to have occurred in two stages:
    a. From coacervates to first primitive living systems called protocells or eobionts or protobionts.
    b. From eobionts to the first living cells.
12. Carbon-carbon bonds are very strong and stable bonds. Due to this bonding property of carbon, it can form an infinite number of molecules, which are the fundamental components of living things.
13. Macromolecules (proteins, lipids, nucleic acids and polysaccharides) are formed from monomeric units or building blocks (amino acids, fatty acids, glycerol, nucleotides and monosaccharides).
14. The smallest element that contributes to the formation and structure of a living organism is scientifically termed as a biological building block.

1.8 SUMMARY

- By definition, a 'cell' is the fundamental and structural unit of all living organisms. It is the smallest biological, structural and functional unit of all plants and animals. Therefore, cells are called the 'building blocks of life' or the 'basic units of life'.
- Organisms made up of a single cell are termed as 'unicellular' whereas organisms made up of many cells are termed as 'multicellular'.
- The cells form tissues and multiple tissues make up an organ, different organs create an organ system, such as digestive system, respiratory system, circulatory system, nervous system, etc., to perform specific functions in the human body and any other living organism.
- Cells range in its size from a millimetre to microns and generally varies in their shapes. Few cells are flat, oval, rod, curved, spherical, concave, rectangular, and various other shapes are also found.
- Animal cells are a typical eukaryotic cell with a membrane-bound nucleus with the presence of DNA inside the nucleus. They also comprise of other organelles and cellular structures which carry out specific functions necessary for the cell to function properly.

NOTES

- There is another basic cell structure that is present in many but not all living cells, termed as 'nucleus'. It is a structure in the cytoplasm that is surrounded by a membrane (the nuclear membrane) and contains, and protects, most of the cell's DNA.
- Based on whether they have a nucleus, the cytologists recognize two basic types of cells as Eukaryotic cell and Prokaryotic cell.
- A prokaryotic cell is much simpler and smaller than eukaryotic cells. It lacks membrane bound organelles including nucleus.
- The eukaryotic cell is much more complex and it contains many membrane bound organelles to perform specific functions. It contains a nucleus isolated from cytosol and enclosed in a well-defined double membrane.
- Cytosol is the liquid part filled inside the cell and it contains water, salt, macromolecules (Protein, Lipid, RNA). It has an array of microtubule fibres running throughout the cytosol to give vesicular structure to its destination.
- Nucleus is the central processing unit of Eukaryotic cell and homologous to the processor in a typical computer. It contains genetic material (DNA) in a complex fashion involving several proteins (histones) to pack into nuclear bodies or chromosomes.
- Mitochondria is popularly known as 'power house of the cell' as the organelle is actively involved in the generation of ATP to run the cellular activities.
- There are two different types of Endoplasmic Reticulum (ER) present in the cell, Rough Endoplasmic Reticulum (RER), and Smooth Endoplasmic Reticulum (SER).
- The major components of the cell are (1) Cell Membrane, (2) Cytoplasm and (3) Nucleus.
- Each cell has a limiting boundary, the cell membrane, plasma membrane or plasmalemma. It is a living membrane, outermost in animal cells but next to cell wall in plant cells.
- The plasma membrane is composed of a lipid bilayer of phospholipid molecules into which a variety of globular proteins are embedded.
- In mitochondria, the wall is made of double membrane. The inner membrane is folded inside to form projections called cristae which project into the inner compartment called matrix.
- Centrioles are involved in cell division. They give orientation to the 'mitotic spindle' which forms during cell division.
- Number of chromosomes is fixed in an organism. During cell division chromosomes divide in a manner that the daughter cells receive identical amounts of hereditary matter.
- Chromosomes bear genes in a linear fashion and thus are concerned with transmission of characters from generation to generation.
- Chromosomes vary in shape, size and number in different species of plants and animals.

NOTES

- Chromosomes are composed of DNA, RNA and histones. DNA is the major genetic constituent of chromosomes.
- The Endoplasmic Reticulum (ER) is an important organelle in eukaryotic cells. It plays a major role in the production, processing, and transport of proteins and lipids.
- The ER produces transmembrane proteins and lipids for its membrane and for many other cell components including lysosomes, secretory vesicles, the golgi apparatus, the cell membrane, and plant cell vacuoles. It was first reported by Porter in 1945.
- The golgi apparatus, like the endoplasmic reticulum, is a canalicular system with sacs that performs some important cellular functions like biosynthesis of polysaccha-rides and packaging of cellular products.
- The golgi complex occurs in all cells except the prokaryotic cells and some eukaryotic cells like mature sieve tubes of plants, mature sperm and red blood cells of animals.
- The cisternae are flattened, plate or saucer-like closed compartments. These are arranged in an orderly stack, much like a stack of pancakes. Typically, a golgi stack contains fewer than eight cisternae.
- The main function of golgi complex is cell secretion, not only of exportable proteins but also of the enzymes present in lysosomes and peroxisomes.
- Lysosomes were first reported by Christian de Duve and co-workers in Belgium in 1955 following their extensive work on the biochemical identification of certain hydrolytic enzymes in the liver cells of rats.
- The sum of the physical and chemical processes occurring within a living organism is known as metabolism.
- There are various theories about the origin of life on earth, but none was able to describe where the first life came from.
- Natural environmental conditions always keep on changing. Living organisms have the ability to change themselves according to these changing conditions. This is known as adaptability of organisms. This adaptability leads to 'origin of new species'.
- Several theories have been put forward for the study of evolution. These have been categorized in the following four major categories: Theory of special creation, Greek theories, Pre-modern theories, and Modern theories.
- Though he was building on the work of his mentor, Count George-Louis Leclerc de Buffon, Jean-Baptiste Lamarck is often given credit for taking the first large advance toward modern evolutionary theory.
- Isolation refers to segregation of similar individuals into two different species. This is done to prevent them from interbreeding.
- Charles Darwin provided an important hint to this problem in a letter to Joseph Hooker in 1871, in which he imagined that if we could provide all ammonia, phosphoric acids, light, heat, electricity, and so on to a warm pond, it could lead to the chemical production of some protein compounds, which could then lead to some complex changes.

NOTES

- Coacervates are in continuous process of breaking down and building up or it can be said that they are in the state if dynamic equilibrium.
- From the nucleotide precursors, polynucleotides were formed which on polymerization gave nucleic acids like Deoxyribonucleic Acid (DNA) and Ribonucleic Acid (RNA). First RNA developed and then DNA.
- As the protein and lipid molecules were synthesized in the eobionts, the lipid-protein, membrane was formed around the eobionts. This membrane allowed the accumulation, infusion and outflow of the molecules in them.
- Carbon is one of the six elements that form the major framework of living matter. It is the most prime and versatile element of life. Its presence is found in all the known living systems, which means that no life can exist without it.
- Important carbon compounds like nucleotides, nucleic acids, lipids, carbohydrates, amino acids, proteins, play various important roles in living beings like metabolism, respiration, building of structural parts of cells.
- The term 'biomolecule' refers to any organic molecule that is formed by a living organism. This includes large polymeric molecules like proteins, polysaccharides and nucleic acids and small molecules.
- A macromolecule is defined as very large molecule, for example a protein molecule which is composed of thousands of covalently bonded atoms.
- The smallest element that contributes to the formation and structure of a living organism is scientifically termed as a biological building block.

1.9 KEY TERMS

- **Unicellular:** Organisms made up of a single cell are termed as unicellular.
- **Multicellular:** Organisms made up of many cells are termed as multicellular.
- **Mitochondria:** It is popularly known as 'power house of the cell' as the organelle is actively involved in the generation of ATP to run the cellular activities.
- **Prokaryotic Cell:** It is a cell without a well-defined nucleus or membrane-bound organelles such as mitochondria or lysosomes.
- **Eukaryotic Cell:** It is a cell that has a nucleus within the nuclear membrane forming a large complex organism.
- **Endoplasmic Reticulum:** It is a transportation system of the eukaryotic cell. It is made of two subunits, namely rough endoplasmic reticulum (RER), and smooth endoplasmic reticulum (SER).
- **Golgi Body:** It is also known as a Golgi apparatus. It is a cell organelle that helps process and package proteins and lipid molecules, especially proteins destined to be exported from the cell.
- **Adenosine Triphosphate (ATP):** It is an organic compound and hydrotrope that provides energy for the functioning of several processes in living cells.

NOTES

1.10 SELF ASSESSMENT QUESTIONS AND EXERCISES

Short-Answer Questions

1. Differentiate between eukaryotic and prokaryotic cells.
2. What is the significance of nucleus?
3. Write short notes on nucleus.
4. Write a brief note on the structure of lysosomes.
5. What are the differences between plant cell and animal cell?
6. What are the different types of metabolic reactions?
7. What are the pre-modern theories of evolution?
8. What are the main biological building blocks?

Long-Answer Questions

1. Discuss the significance of cell with reference to cell theories and cell structure giving appropriate examples.
2. Explain the structural organization of prokaryotic and eukaryotic cells.
3. Discuss the different cell components with the help of diagrams.
4. Describe the functions and structure of ER in detail.
5. Describe about peroxisomes, its structure, origin, functions and disorders related to peroxisomes functions in detail.
6. Evaluate the four major events that occurred in chemical evolution of life.
7. Discuss the modern theories of evolution in detail.
8. Analyze the evidences of evolution in detail.

1.11 FURTHER READING

Hunter, L. E. 2009. *The Process of Life: An Introduction to Molecular Biology*. USA: The MIT Press.

Weaver, R. F. 2008. *Molecular Biology*. New York: McGraw Hill Higher Education.

Becker, W.; L. Kleinsmith; J. Hardin; and G. Bertoni. 2008. *The World of the Cell*. London: Pearson Education.

Alberts, B.; A. Johnson; J. Lewis; M. Raff; K. Roberts; and P. Walter. 2007. *Molecular Biology of the Cell*. New York: Garland Publishing Inc.

Karp, G. 2007. *Cell and Molecular Biology: Concepts and Experiments*. New York: John Wiley and Sons Inc.

Gupta, P. K. 1999. *Cell and Molecular Biology*. Meerut: Rastogi Publications.

UNIT 2 CARBOHYDRATES

NOTES

Structure

2.0 INTRODUCTION

A carbohydrate is a biomolecule consisting of carbon (C), hydrogen (H) and oxygen (O) atoms, usually with a hydrogen–oxygen atom ratio of 2:1. Carbohydrates perform numerous roles in living organisms. Carbohydrates are central to nutrition and are found in a wide variety of natural and processed foods. The term is most common in biochemistry, where it is a synonym of saccharide, a group that includes sugars, starch, and cellulose. In this unit, we will discuss the structure, types, functions, and metabolism of carbohydrates including monosaccharides and its derivatives, along with the structure and biological functions of glucosaminoglycans. It will also focus on the carbohydrates of glycoproteins and glycolipids, along with the role of sugar in biological recognition, and concepts of blood group substances and ascorbic acid.

2.1 OBJECTIVES

NOTES

After going through this unit, you will be able to:

- Describe the structure, types, functions, and metabolism of carbohydrates including monosaccharides and its derivatives
- Explain the structure and biological functions of glucosaminoglycans
- Discuss the carbohydrates of glycoproteins and glycolipids
- Evaluate role of sugar in biological recognition, and concepts of blood group substances and ascorbic acid

2.2 CARBOHYDRATES: STRUCTURE AND FUNCTIONS

Carbohydrates are polyhydroxy compounds that contain a carbonyl (C=O) group of the general formula $C_n(H_2O)_n$. It is an organic compound which consist of the elements carbon (C), hydrogen (H) and oxygen (O) with a ratio of hydrogen twice that of carbon and oxygen i.e. 1:2:1. Sugars, starches, cellulose and many other compounds are some examples of carbohydrates found in living organisms. The term Carbohydrate is a synonym of saccharides. The word '*saccharides'* come from the Greek word sákkharon, meaning 'sugar'.

Carbohydrates are widely distributed in plants and animals. They have important structural and metabolic roles. In plants, glucose is synthesized from carbon dioxide and water by photosynthesis and stored as starch or used to synthesize cellulose for plant framework. Animal can also synthesize carbohydrate from lipid and amino acids but most of animal carbohydrates are derived ultimately from plants. Glucose is the most important carbohydrate as it is the most dietary carbohydrate absorbed in bloodstream. The other sugars are converted into glucose in liver. Glucose is the precursors for the synthesis of all other carbohydrate in the body including glycogen (storage carbohydrate in animal cells), ribose and deoxy-ribose in nucleic acids, lactose in milk, etc.

2.2.1 Functions of Carbohydrates

The functions of carbohydrates are given as follows:

- Carbohydrates give the primary energy molecules to the body and leave protein to be used up as a molecule for energy production, instead of this protein concentrates in the body and repairing and maintain the body.
- They are the only energy source of energy for the brain. It is necessary for the nervous tissue regulation. They are necessary for the complete metabolization of fats.
- Some carbohydrates encourage the growth of healthy bacteria in the intestine.
- They are sometimes used to add flavors and as sweeteners.
- They prevent ketosis that is caused due to high metabolism of fat.
- They regulate the sugar in the blood.

NOTES

- They help in the fertilization, growth and development of cells.
- Complex carbohydrates are a good source of fiber.
- In excess amount, carbohydrates are stored in the body as fats. Hence, it is recommended to take a minimum daily amount of carbohydrates in the body so that they are not converted to fats and lead to weight gain.
- Cellulose and other carbohydrates form the cell wall of plant cells along with hemicelluloses and pectin.
- Chitin and other carbohydrates form the cell wall of fungal cells and the exoskeleton of arthropods.
- Peptidoglycans and other carbohydrates form the cell wall of bacteria and cyanobacteria.
- Heparin is a form of a carbohydrate that prevents intravascular clotting of blood.
- They function as hormones, such as Follicular Stimulating Hormone (FSH) in ovulation in females) and Luteinizing Hormone (LH).
- Carbohydrates are also used in industries like paper, textile and breweries.
- Hyaluronic acid is a glycoprotein, which acts as a lubricant between joints to provide frictionless movement.
- Many antigens are glycoproteins that give immunological properties to blood.
- Agar is a polysaccharide which is used in the culture media.
- Different forms of carbohydrates are stored in living organisms as food, such as polysaccharides, starch in plants and glycogen in liver and muscles in animals.

Carbohydrates are classified on the basis of number of monomers they contain. They are categorized mainly in four categories as follows:

1) **Monosaccharides:** $C_n(H_2O)_n$ where n = 3-9 and n denotes the number of carbon atom.
2) **Oligosaccharides:** Polymers from 2-20 molecules of monosaccharides.
3) **Polysaccharides:** Polymers containing more than 20 sugar residues.
4) **Glycoconjugates:** Derivatives of carbohydrates containing proteins, lipids.

2.2.2 Role of Sugars in Biological Recognition

Glycocalyx is the major component of mammalian cell surface. It is composed of complex layer of carbohydrates which determines the interaction of cells with its surroundings. Glycans mediates the interaction cellular and subcellular components extending from molecular and subcellular which a cell makes with its environment on micro-, systemic, and—referring to the ever-topical threat of influenza pandemic—sometimes quite literally, global scales. The simplest unit of glycosylation present on a molecular level is O-linked N-acetylglucosamine (O-GlcNAc) resulted from modification of hundreds, and most likely thousands, of nuclear and cytosolic proteins.

NOTES

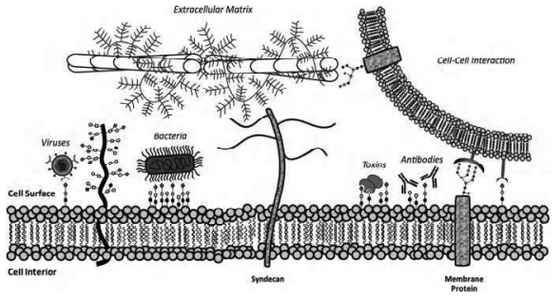


Fig. 2.1 *Structure of Mammalian Cell Surface*

With the increasing complexity, the GlcNAc2Man9Glc3 N-glycan structure contributes in quality control during the folding of membrane-displayed and secreted glycoproteins, and once these and other classes of glycans reach the cell surface they control both nano- and microscale properties of the plasma membrane. An example of the this is 'glycosynapse', while one of the latter is the galectin lattice specified by N-glycan branching status.

Specific surface glycans, or the united cellular complement of these molecules, reflect the internal workings and status of a cell and thus serve as available biomarkers for, amongst other conditions, cancer and stem cell status. The term 'biomarker' suggests about the multiple functions glycans of cellular carbohydrates have within cells (e.g., O-GlcNAc in signaling and N-glycans in protein folding) and as well as in the interactions they make with their surroundings.

Glycan-mediated interactions that a cell makes with its environment begin on a close scale as these molecules modulate adhesion to neighbouring cells. Moving to the systemic and organism-wide levels, glycans are the veritable workhorses of multicellular life, allowing the complex mix of cell types to grow in association with each other. Their contributions begin with fertilization at the very start of life and extend to choreographing the activity of certain hormones, including growth hormone during development and in the adult, as well as orchestrating the immune system.

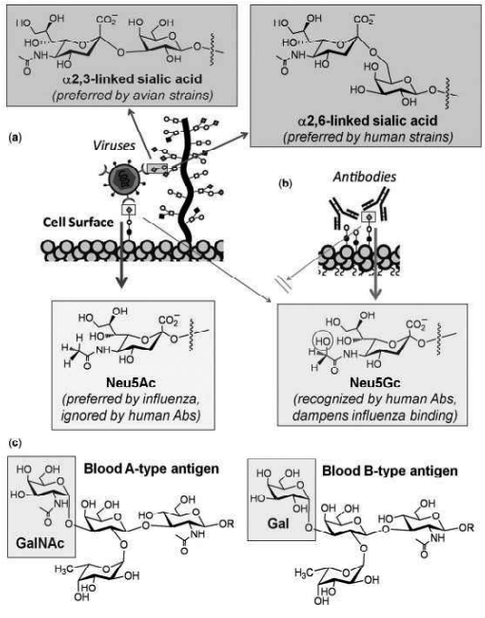


Fig. 2.2 *Glycan-Mediated Interactions*

In the given figure:

a) The level of discrimination of molecules that 'decode' the glycocalyx is illustrated by the ability of viruses such as strains of influenza to discriminate between a 2,3- and a2,6-linked sialic acid and at an even more nanoscopic scale, between the Neu5Ac and Neu5Gc forms of sialic acid that vary by a single oxygen atom.

b) Glycans can also be potent antigens, with the human immune system recognizing the non-human 'Neu5Gc' sialoside.

c) The immune system is also able to discriminate between the single added N-acetyl group in A-type blood compared to B-type (in O-type blood, the entire GalNAc or Gal residue is absent).

The plasma membrane of the animal cell contains asymmetrical distributed glycoproteins and glycolipids which extend their carbohydrate-bearing portions directly into the extracellular environment, and are the major attention getting elements because they have vital role in cell recognition.

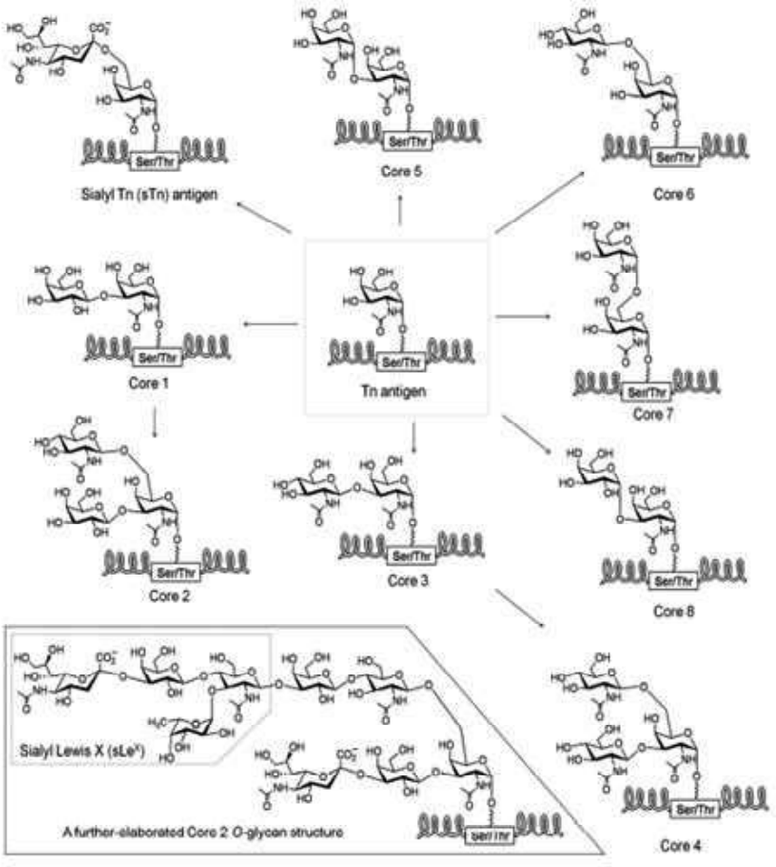


Fig. 2.3 *Overview of Mammalian Glycan Biosynthesis*

In the above figure:

a) Common dietary sugars such as glucose (Glc), glucosamine (GlcN), galactose (Gal), and mannose (Man) are taken into a cell by a family of

NOTES

transporters and diversified into fucose (Fuc), N-acetylglucosamine (GlcNAc), N-acetylgalactosamine (GalNAc), xylose (Xyl), glucuronic acid (GlcA), Nacetylmannosamine (ManNAc), N-acetylneuraminic acid (Neu5Ac), and N-glycolylneuraminic acid (Neu5Gc).

b) Both sets of sugars, with the exception of GlcN and ManNAc, are then converted into nucleotide sugars represented by CMP-Neu5Ac/Gc, UDP-GlcNAc (Glc, GlcA, Gal, GalNAc, and Xyl also utilize UDP), and GDP-Man (Fuc also is linked to GDP), which are used as building blocks for glycan assembly by families of glycosyltransferases.

c) An example of these enzymes is provided by the suite of sialyltransferases that construct a2,3-, a2,6-, and a2,8-linked sialosides; this latter category is represented by a single a2,8-linked residue on GD3 or the homopolymer of dozens of residues resident on the neural cell adhesion molecule (NCAM).

The cell membranes of bacteria, fungi and higher plants are, in general, in contact with a complex carbohydrate-rich cell wall which complicates recognition of sugar-containing plasma membrane components. The cell wall, moreover, shields the underlying membrane from direct interaction with the cell surroundings.

2.3 CONFORMATION OF MONOSACCHARIDES: STRUCTURE AND FUNCTIONS OF IMPORTANT DERIVATIVES GLYCOSIDIC, DEOXY SUGARS, MYO INOSITAL, AMINO SUGARS, N-ACETYLMUAMIC ACID AND SIALIC ACID

Monosaccharide or simple sugars are those carbohydrates that cannot be hydrolyzed into simpler carbohydrate containing a single polyhydroxy aldehyde or ketone unit. Six-carbon sugar D-glucose (sometimes referred as Dextrose) is the abundant monosaccharide found in nature. More than four carbons presented in monosaccharide tend to have cyclic structures. Their general formula is $C_n(H_2O)_n$ or $C_nH_{2n}O_n$ where n is the number of carbon atom which is from 3 to 9 and cannot be less than 3.They are colorless, crystalline solids that are freely soluble in water but insoluble in non-polar solvents such as ether, benzene etc. Most have a sweet taste. The simplest monosaccharide has either aldehydes or ketones with two or more hydroxyl groups. The most common six-carbon monosaccharide glucose and fructose have five hydroxyl groups. Most of the carbon atoms presented in monosaccharide remains attached to hydroxyl groups called as chiral centers or chiral carbon (the carbon atom having four different functional groups) , which give rise to many stereoisomers of sugar found in nature.

The framework of common monosaccharide molecules are un-branched carbon chains in which all the carbon atoms are linked by single bonds. In the open-chain structure of monosaccharide, one of the carbon atoms is double-bonded

NOTES

to an oxygen atom to form a carbonyl group; each of the other carbon atoms has a hydroxyl group. If the carbonyl group is at an end of the carbon chain that is, in an aldehyde group, the monosaccharide is referred as an aldose sugar and if the carbonyl group is at any other position (that is in a ketone group, the monosaccharide is referred as a ketose sugar. The three-carbon trioses sugar includes: glyceraldehyde, an aldotriose, and dihydroxyacetone, a ketotriose which are simplest monosaccharide as shown in Figure 2.4.

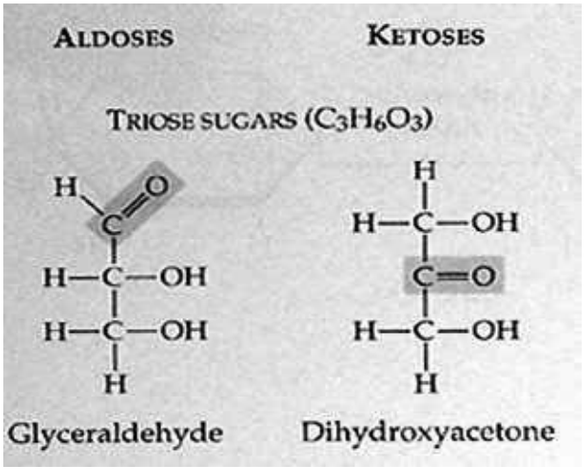


Fig. 2.4 *Showing Structure of Triose Sugar*

Monosaccharides having four, five, six, and seven carbon atoms in their framework are named respectively as tetroses, pentoses, hexoses, and heptoses. There are aldoses and ketoses of each of these chain lengths: aldotetroses and ketotetroses, aldopentoses and ketopentoses, and so on as represented in Table-2.1. Aldohexose D-glucose and the ketohexose D-fructose are the most common monosaccharides found in nature. The aldopentoses D-ribose and ketopentose 2-deoxy-D-ribose are major components of nucleotides and nucleic acids (Refer Figure 2.5)

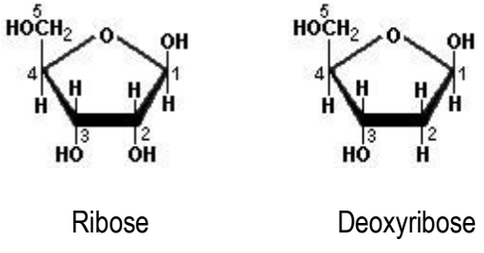


Fig. 2.5 *Showing the Ring form Structure of Ribose and Deoxyribose*

Sometimes, a distinction in naming between aldoses and ketoses is also maintained. The suffix-**oses** is kept reserved for the aldoses and the suffix - uloses is used for ketoses. Thus, glucose is a hexose and fructose is a hexulose.

NOTES

Table 2.1 *Showing the Nomenclature of Aldose and Ketose Sugar*

No. of Carbon atom	Category name	Formula	Aldose	Ketose
3	Trioses	$C_3H_6O_3$	Glycerose (Glyceraldehyde)	Dihydroxyacetone
4	Tetroses	$C_4H_8O_4$	Erythrose	Erythrulose
5	Pentoses	$C_5H_{10}O_5$	Ribose	Ribulose
6	Hexoses	$C_6H_{12}O_6$	Glucose	Fructose
7	Heptoses	$C_7H_{14}O_7$	Glucoheptose	Sedoheptulose

A molecule with n chiral centers can have 2n stereoisomers commonly. Stereoisomers are those isomeric molecules that have the same molecular formula and have constituent sequence of bonded atoms, but differ only in the three-dimensional orientations in space of their atoms. Structural isomers have same molecular formula but differs in the bond order between different atoms/groups.The monosaccharide stereoisomers of each carbon-chain length can be divided into two groups which differ in the configuration about the chiral center most distant from the carbonyl carbon.

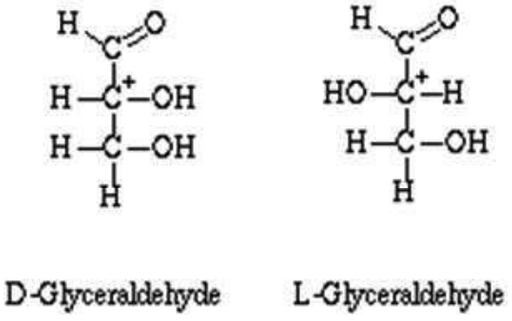


Fig. 2.6 *Showing the Open Chain Structure of D and L-Isomer of Glyceraldehyde*

When the hydroxyl group on the reference carbon is on the right in the projection formula, the sugar is the Dextrorotatory (D) isomer; when on the left, it is the Levorotatory (L) isomer. Those in which the configuration at this reference carbon is the same as that of D-glyceraldehyde are designated D isomers, and those with the same configuration as L-glyceraldehyde are L isomers. Of the 16 possible aldohexose-glucose configuration, eight occurs in D- isomeric form and the other eight occurs in L- isomeric form. Most of the hexoses found in living organisms have D-isomeric form. The numbering of carbon atom of a sugar starts at the end of the carbon chain nearest the carbonyl group. Stereoisomeric form of each of the eight D-aldohexoses differing at C-2, C-3, or C-4, has its own name denoted as- D-glucose, D-galactose, D-mannose, and so forth. The naming of four- and five-carbon ketoses are designated by inserting suffix 'ul' into the name of a corresponding aldose; for example, D-ribulose is the ketopentose corresponding to the aldopentose D-ribose.

NOTES

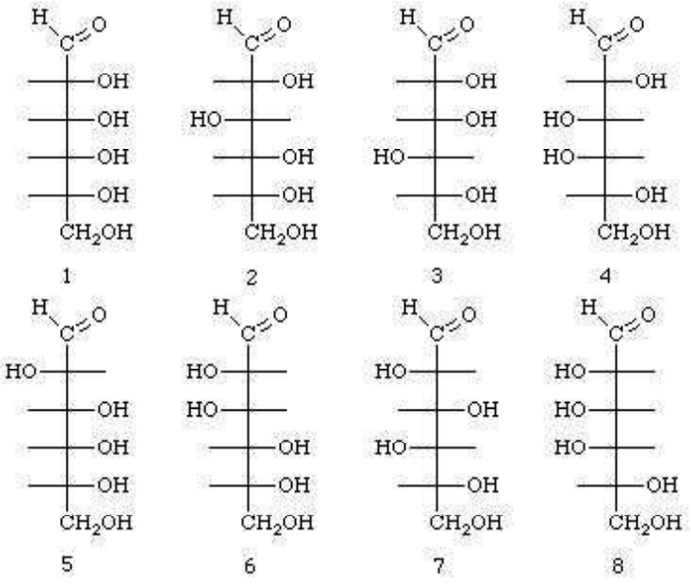


Fig. 2.7 *Showing Stereoisomeric Possibilities of Glucose*

In aqueous solution, aldotetroses and all monosaccharides with framework of carbon atoms in five or more carbon atoms occur predominantly as cyclic (ring) structures in which the carbonyl group has created a covalent bond through oxygen atom of a hydroxyl group alongside the chain. The creation of these ring structures is the outcome of a common reaction occurring between alcohols and aldehydes or ketones to form derivative called hemiacetals or hemiketals, which contain an extra asymmetric carbon atom and thus can exist in two stereoisomeric forms. For example, D-glucose exists in solution as an intramolecular hemiacetal in which the free hydroxyl group at C-5 has reacted with the aldehydic C-1, account the latter carbon asymmetric and producing two stereoisomers, designated α and β. These six-membered ring compounds are called pyranoses because they resemble the six membered ring compound pyran. The methodical names for the two ring forms of D-glucose are α-D-glucopyranose and β-D-glucopyranose. Aldohexoses also exist in cyclic forms having five membered rings, which, because they resemble the five membered ring compound furan known as furanoses. However, the six-membered aldopyranose ring is much steadier than the aldofuranose ring and predominates in aldohexose solutions. Aldoses having five or more carbon atoms can form pyranose rings only.

NOTES

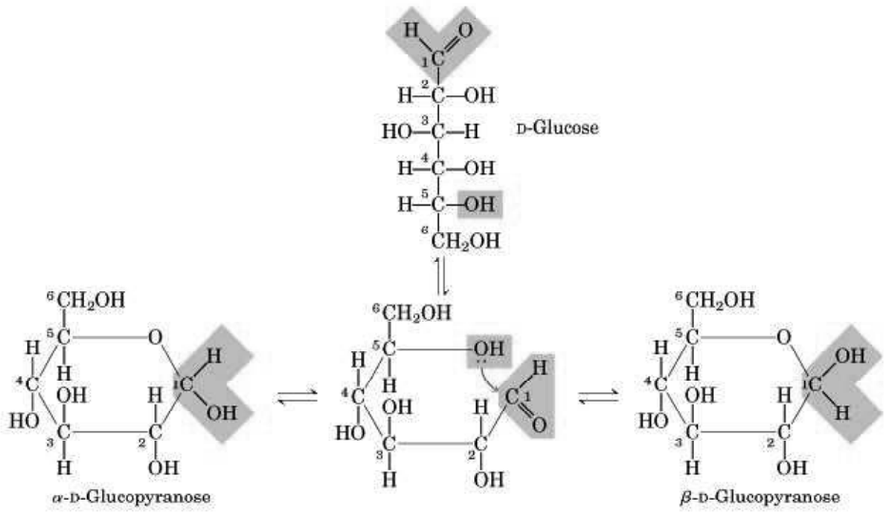
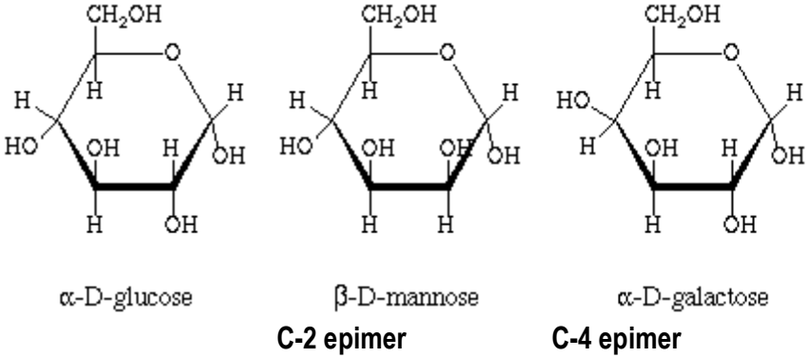


Fig. 2.8 *Showing Furan and Pyran Ring Structure of Glucose*

The monosaccharides isomeric forms that differ only in their configuration about the hemiacetal or hemiketal carbon atom are called anomers. The hemiacetal (or carbonyl) carbon atom is called the anomeric carbon. In aqueous solution α and β anomers of D-glucose gets interconvert by a process called mutarotation. Thus, a solution of α-D-glucose and β-D-glucose eventually form the same equilibrium mixtures having indistinguishable optical properties. This mixture consists of about one-third α -D-glucose, two-thirds β -D-glucose, and very small amounts of the linear and five-membered ring (glucofuranose) forms. Ketohexoses also occur in α and β anomeric forms. In these compounds the hydroxyl group at C-5 (or C-6) reacts with the keto group at C-2, forming a furanose (or pyranose) ring containing a hemiketal linkage (Figure 2.8). D-Fructose eagerly forms the furanose ring. The more ordinary anomer of this sugar in combined forms or in derivatives is β -D-fructofuranose.Two sugars that varies only in the configuration around one carbon atom are called as epimers. For example-D-glucose and D-mannose, D-glucose and which varies only in the stereochemistry at C-4 are D-galactose (Figure 2.6). Most of the sugars occur naturally in their D-isomeric form, however, a few sugars may occur naturally in their L-isomeric form.



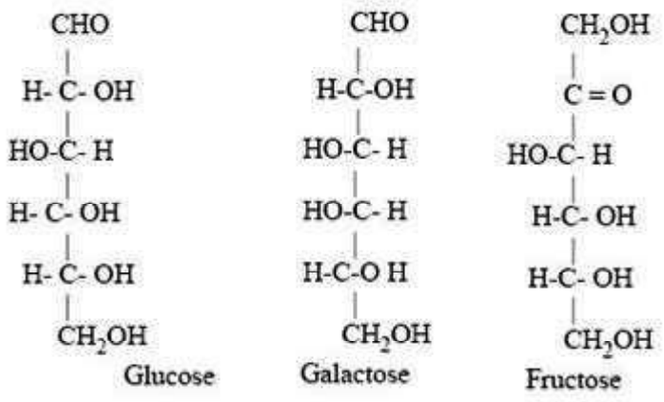


Fig. 2.9 *Showing the Open and Ring Form Structure of Epimeric Form of Glucose*

Comparatively mild oxidizing agents such as ferric (Fe^{3+}) or cupric (Cu^{2+}) ion converts carbonyl carbon presented in monosaccharides to a carboxyl group by oxidation. On the basis of oxidizing activity, monosaccharides are categorized in two classes. Monosaccharides (such as, glucose) which are capable of reducing ferric or cupric ion are known as reducing sugars and those sugars which are incapable of reducing ferric or cupric ion are known as Non-reducing sugars. The differences between reducing and non-reducing sugars are described in Table-2.2. Fehling's reaction, a qualitative test for the presence of reducing sugar is based on property of reduction. So, the concentration of sugar can be estimated by measuring the amount of oxidizing agent reduced by a solution of sugar. For numerous years this test was used to detect and measure elevated glucose levels in blood and urine in the diagnosis of diabetes mellitus. Now days, more sensitive and reliable methods are available for measuring blood glucose, for example, glucose oxidase enzyme assay method.

Table 2.2 *Showing Differences between Reducing and Non-Reducing Sugar*

Reducing sugar	Non-reducing sugar
1. Carbohydrates with a free aldehyde (at C-1)	1. Aldehyde or ketone group is not free but instead utilized in bond formation or a free ketone (at C-2) group.
2. They are in hemiacetal or hemiketal form.	2. They are in acetal or ketal form
3. Do exhibit mutarotation	3. Do not exhibit mutarotation
4. Do form osazones with phenyl hydrazine	4. Do not form osazones.
5. Do form oximes with hydroxylamine.	5. Do not form oximes
Examples - Glucose, Fructose, Lactose, Maltose, Cellobiose etc.	Examples - Sucrose, Glycogen, Inulin etc.

Check Your Progress

1. What is the composition of carbohydrates?
2. What are monosaccharide?
3. What are reducing sugars?
4. What is the simplest unit of glycosylation present on a molecular level?

NOTES

2.4 OLIGOSACCHARIDES

These are compound sugars that yield 2 to 10 molecules of the identical or different monosaccharides on hydrolysis (removal of water molecules). Accordingly, an oligosaccharide yielding 2 molecules of monosaccharide on hydrolysis is designated as a disaccharide, and the one yielding 3 molecules of monosaccharide as a trisaccharide and so on. The general formula of disaccharides is $C_n(H_2O)_{n-1}$ and that of trisaccharides is $C_n(H_2O)_{n-2}$ and so on.

2.4.1 Disaccharides

The condensation of two molecules of monosaccharides via glycosidic bond between the C-1 of one sugar and the -OH of another carbon leads to the formation of Disaccharide. They are water soluble, sweet in taste and used to carry monosaccharides. Maltose, lactose, and sucrose are few examples of Disaccharide which consist of two molecules of monosaccharide joined covalently by an *O*-glycosidic bond which get formed when a hydroxyl group of one sugar reacts with the anomeric carbon of the other sugar (Figure 2.10). Glycosidic bonds presented in disaccharides can be hydrolyzed by dilute acid to yield their free monosaccharide components but resist cleavage by base.

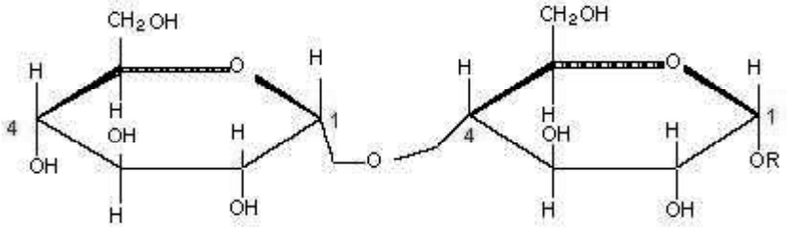


Fig. 2.10 *Showing the Formation of Glycosidic Bond*

The examples of disaccharides include:

1) **Maltose** – It is formed by the condensation of 2 molecules of α-D glucose joined at C-1 of one residue and C-4 of second residue by glycosidic bond (Figure 2.11).

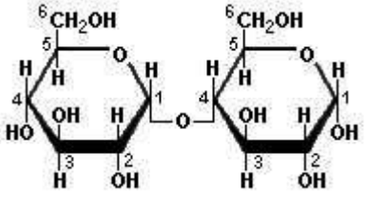


Fig. 2.11 *Showing the Structure of Maltose*

2) **Lactose** - It is formed by the condensation of one molecule of glucose and galactose (Figure 2.12). It is the chief constituent found in the milk. The intestinal distress caused by a deficiency of lactase, an intestinal enzyme needed to digest and absorb lactose in milk. Abdominal pain, bloating gas and diarrhea are caused due to undigested lactose ferments in the colon.

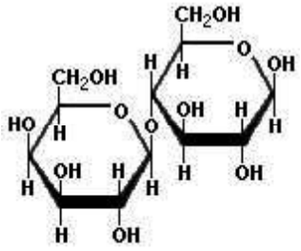


Fig. 2.12 Showing the structure of Lactose

3) **Sucrose –** It is formed by the condensation of one molecule of glucose and fructose (Figure 2.13). It is most commonly known as table sugar or saccharose. Sucrose is the main ingredient found in dried cane juice and brown sugar.

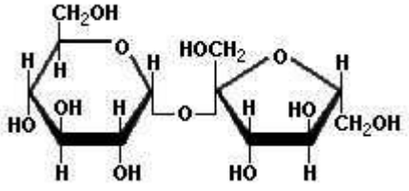


Fig. 2.13 Showing the Structure of Sucrose

2.4.2 Trisaccharide

Trisaccharide such as Raffinose which is broadly found in legumes and cruciferous vegetables including beans, peas, cabbage, Brussels sprouts and broccoli. Raffinose is also called as melitose and are consists of galactose connected to sucrose via (Figure 2.14). Due to lack of enzyme in humans, this tri-saccharide cannot be digested. So, Raffinose is fermented in large intestine of humans by gas-producing bacteria.

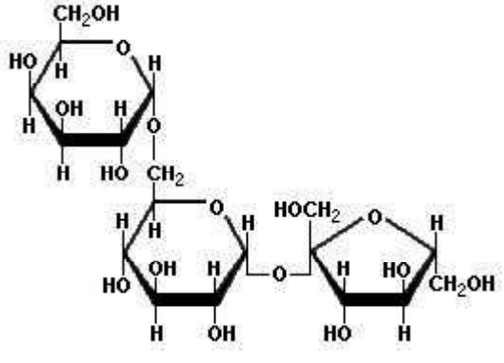


Fig. 2.14 Showing the Structure of Raffinose

2.5 POLYSACCHARIDES

Polysaccharides are condensation yield of more than 20 or more so monosaccharide unit and few may have hundreds or thousands of units. They are also known as Glycans. Most carbohydrates found in nature occur as polysaccharides, polymers of medium to high molecular weight. Depending upon

NOTES

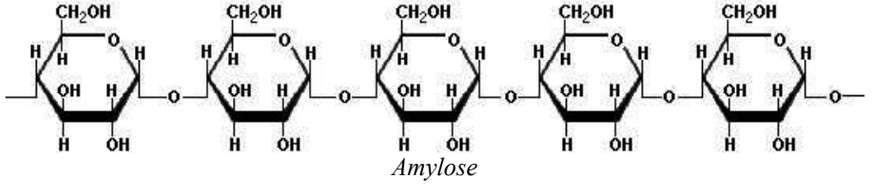
the type of monosaccharide involved in formation, polysaccharides are categorized into two categories, which include Homopolysaccharides and Heteropolysaccharides. The polysaccharides which single type of monomer of monosaccharide are referred as homopolysaccharide whereas those which contain two or more different types of monosaccharide are referred as heteropolysaccharides. Some homopolysaccharide serve as storage purpose that may be used as fuels such as starch and glycogen which occur intracellularly as granules or in large clusters form. Heteropolysaccharide forms a mixture of monosaccharides on hydrolysis. They are found in many number both in plants and as well as in animals. The strain of pneumococcus type III containing soluble sugar is one of the simplest heteropolysaccharide which is the mixture of repeating units of a disaccharide consisting of glucose and glucuronic acid. It yields equimolar concentration of glucose and glucuronic acid on hydrolysis.

2.5.1 Storage Polysaccharides (Starch and Glycogen)

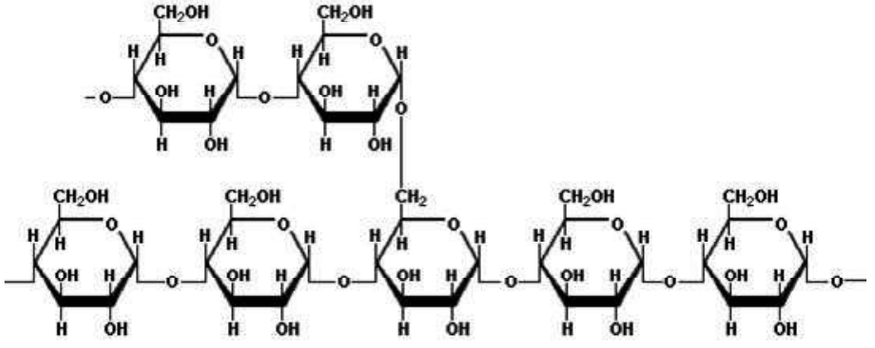
Some of the storage polysaccharides are described below:

Starch

Starch is a homopolymer of glucose forming an α-glycosidic chain called as glucosan or glucan.It is the most abundant dietary carbohydrate in cereals, potatoes, legumes and other vegetables. Plant cells have the capability to form starch. Starch contains two types of glucose polymer, amylose (15-20%) and amylopectin (80-85%) (Figure 2.12). Amylose is consists of long, unbranched chains of D-glucose residues connected by α-1→4 glycosidic linkages having variable molecular weight from a few thousand to more than a million. Amylopectin is highly branched and has a high molecular weight up to 100 million. The glycosidic linkages joining successive glucose residues in amylopectin chains are α-1→4 glycosidic bond and in the branch points which occurs in every 24 to 30 residues are linked by α-1→6 glycosidic linkage. Dextrins are intermediates in the hydrolysis of starch.



Amylose



Amylopectin

Fig. 2.15 *Showing the Structure of Amylase and Aylopectin*

Glycogen

The main storage polysaccharide of animal cells is glycogen. Glycogen is a polymer of subunits of glucose with α-1→4 glycosidic linkage at the linear site and α-1→6 glycosidic linkage at the branching site (Figure 2.16). Glycogen is more extensively branched on average at every 8 to 12 residues and more compact than starch. Glycogen is chiefly abundant in the liver, where it may constitute as much as 7% of the wet weight. In hepatocytes, glycogen is found in large granules which are themselves clusters of smaller granules composed of single, highly branched glycogen molecules with an average molecular weight of several million. It is also present in skeletal muscle. When glycogen is used as an energy source, glucose units are removed one at a time from the nonreducing ends. The conversion of glycogen polymer to its monosaccharides is accomplished by degradative enzymes that act only at non-reducing ends and work simultaneously on the many branch sites.

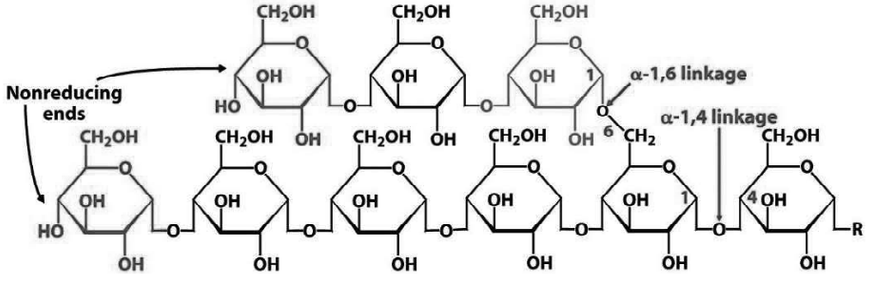


Fig. 2.16 *Showing the Structure of Glycogen*

Glycogen and starch molecules are profoundly hydrated because they have many exposed hydroxyl groups available to hydrogen-bond with water. Glycogen and starch ingested in the diet are hydrolyzed by α-amylases enzymes present in saliva and intestinal secretions that break α-1→4 glycosidic bonds between glucose units. Other homopolysaccharide serve as structural elements in plant cell walls (cellulose) and animal exoskeletons (chitin).

2.5.2 Structural Polysaccharides (Cellulose and Chitin)

Some of the structural polysaccharides are described below:

Cellulose

Cellulose is water-insoluble, fibrous, tough substance found in mainly in the cell walls of plants and predominantly in stalks, stems, trunks, and all the woody portions of the plant body. It constitutes much of the mass of wood, and cotton. The cellulose molecule is a linear, unbranched homopolysaccharide, consisting of 10,000 to 15,000 D-glucose units as resembles in starch and glycogen. The very important difference in cellulose, starch (amylose and Amylopectin) and glycogen is that the glucose repeating units in cellulose have the â configuration (Figure 2.14), whereas in amylose, amylopectin, and glycogen the glucose repeating units is in the â configuration. The glucose residues in cellulose are linked by α-1→4 glycosidic bonds in contrast to the α-1→4 glycosidic bonds of amylose in starch and glycogen. This dissimilarity creates very different structures and physical properties of cellulose

NOTES

and amylose. Due to lack an enzyme to hydrolyze the α-1→4 glycosidic linkages, most animals cannot use cellulose as a fuel source. Termites can readily digest cellulose because their intestinal tract harbors a symbiotic microorganism with Trichonympha that secretes cellulase enzyme which hydrolyzes the α-1→4 linkages. Wood-rot fungi and bacteria also produce cellulase.

Hemicelluloses

Hemicellulose is comprised of a variety of branched-chain polymers containing a mixture of various hexose such as glucose, mannose, and galactose and pentose sugars such as xylose, arabinose, which might also be substituted with uronic and acetic acids. The main hemicelluloses found in plants are xylans (1→4-linked polymers of the pentose sugar xylose), but arabans (polyarabinose), galactans (polygalactose), mannans and copolymers (e.g., glucomannans and galactoglucomannans) are also encountered. They mostly serve as storage and supporting substances in plants. The chief sources of hemicelluloses include whole wheat, some fruits and vegetables and unrefined cereals. They adsorb water and are partially digestible.Enzymes responsible for hemicellulose degradation are named according to their substrate specificity; for example, mannanases degrade mannans, xylanases degrade xylans, etc. As xylans predominate in plant walls, more is known about xylanases. The term 'hemicelluloses' has been discontinued in use as a large number of analogous polysaccharides were revealed in fungi and bacteria.

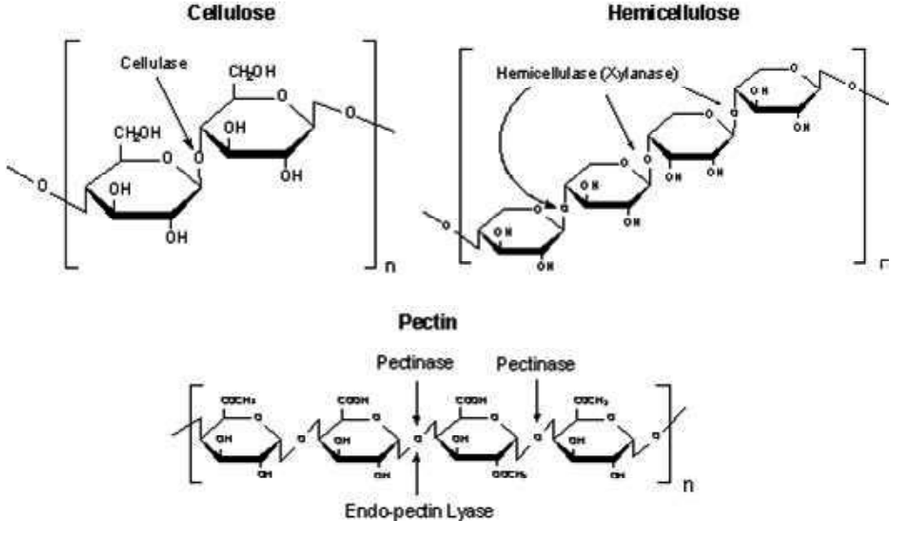


Fig. 2.17 *Showing Cellulose, Pectin and Hemicellulose Sstructure*

Xylan is the most abundant and widely spread carbohydrate following after cellulose. Xylan is a linear homopolymer of an aldopentose sugar called D-xylose having α-1→4 bond bearing side chains of 4–O–methylglucuronic acid or arabinose. Xylan can be derived from a cellulose chain by replacing hydrogen atoms for the —CH_2OH groups, but the number of units per polymer is considerably lower and in the range of 30–100. Straw and bark contain of upto 30% xylan, the residues of sugarcane also comprise about 30% xylan. It also founds in conifer wood comprising between 7–12% and deciduous wood constitute nearby 20–25% xylan.

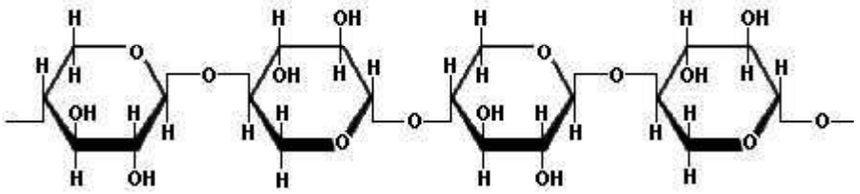


Fig. 2.18 *Showing structure of Xylan*

Xylan is more quickly get digested by a huge number of microorganisms as related to cellulose. Numerous cellulose-degrading organisms such as sporocytophaga, myxococcoides etc. also produce xylanase, an enzyme accountable for the digestion of xylan. The capacity to utilize Xylan is very frequent to be found among fungi and is admirable substrate for the nurturing of mushrooms. The enzyme Xylanase is designed constitutively in some bacteria e.g., Clostridia and in others bacterial cell, it is inducible by xylan.

Chitin

Chitin is a linear homopolysaccharide formed in â linkage of *N*-acetyl glucosamine residues (Figure 2.19). Chitin differs from cellulose only chemically replacement of the hydroxyl group at C-2 with an acetylated amino group. Chitin forms extended fibers analogous to those of cellulose, and it are also cannot be digested by vertebrates like cellulose. Chitin is the most essential component of the rigid exoskeletons of nearly a million species of arthropods including lobsters, crabs, and insects and is almost indeed the second most plentiful polysaccharide, next to cellulose in natural world.

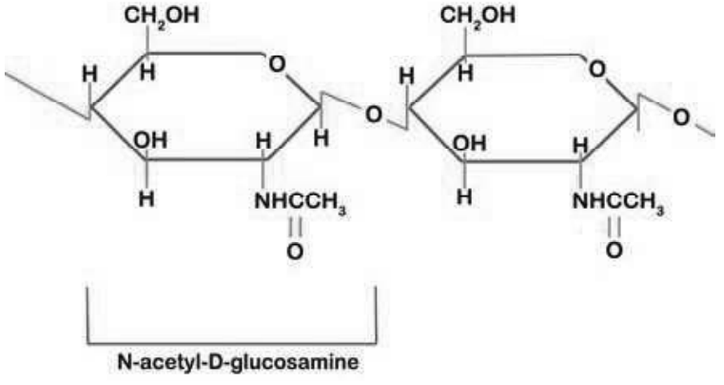


Fig. 2.19 *Showing the Structure of Chitin*

Peptidoglycan

The rigid layer of the bacterial cell wall is comprised of peptidoglycan. Two alternating monosaccharide units which are heteropolymer of *N*-Acetyl-Glucosamine (NAG) and *N*-AcetylMuramic acid (NAM) residues linked by alternating α 1-4 glycosidic bond forms peptidoglycans (Figure 2.20).The exact structure of peptidoglycan depends on the bacterial species and linear polymers stretch out side by side in the cell wall, cross-linked by short peptides. The peptide cross-links connects the polysaccharide chains into a tough sheath that envelops

NOTES

the whole cell and prevents cellular swelling and lysis due to the osmotic entry of water. The enzyme lysozyme destroys bacteria by hydrolyzing the α 1-4 glycosidic bond between N-Acetyl Glucosamine (NAG) and N-Acetyl Muramic acid (NAM). Lysozyme is remarkably enzyme present in eye tears, apparently acts as defensive agent against bacterial infections of the eye. It is also synthesized by certain bacterial viruses to make certain their release from the host bacterial cell, a necessary step of the viral infection cycle. Penicillin and related antibiotics destroys bacteria by preventing synthesis of the cross-links, leaving the cell wall too weak to oppose osmotic lysis.

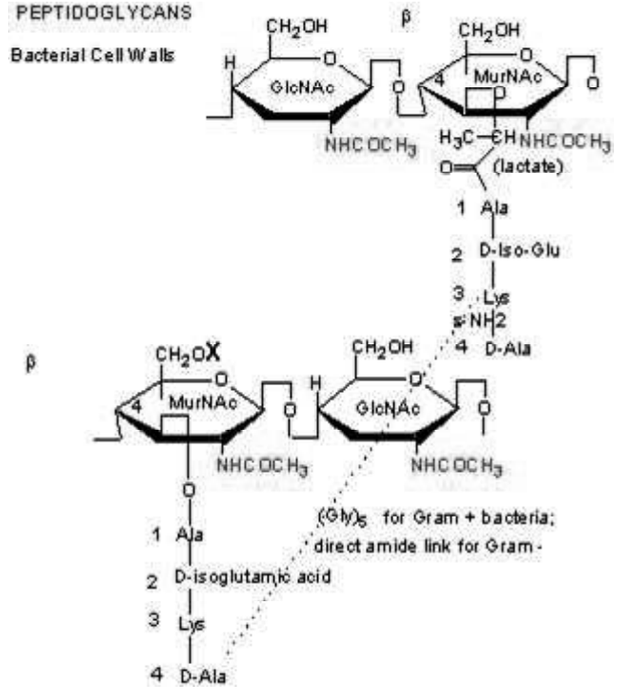


Fig. 2.20 *Showing Structure of Peptidoglycans*

The extracellular space in animal tissues is engaged by several types of heteropolysaccharide, which form a matrix that holds individual cells together and provides shape, support and protection to cells, tissues, and organs.

2.5.3 Structure and Biological Functions of Glycosaminoglycans or Mucopolysaccharides

Polysaccharides which are composed of not only of a mixture of simple sugars but also of derivatives of sugars such as uronic sugars and amino sugars and they are termed as mucopolysaccharides. They mostly acts as structural support material for mucous substances and connective tissue of the body. They have high molecular weights in the range of up to 5×10^6 and are gelatinous in nature. Mucopolysaccharides serve both as a cementing substance and a lubricant. They mostly consist of disaccharide units in which a uronic acid is bound by a glycosidic bond to the C-3 of an acetylated amino acid in $1 \rightarrow 3$ glycosidic linkage. These disaccharide residues are polymerized by $1 \rightarrow 4$ linkages to give a linear macromolecule structure. The sulfuric and uronic acid residues report a strong

acidic character to these substances. The structural features of a few common mucopolysaccharides are illustrated in Table 2.3.

Table 2.3 *Showing the list of Common Mucopolysaccharide*

Mucopolysaccharides	**Two components of the disaccharide units**
Hyaluronic acid	D-glucuronic acid +N-acetyl-D-glucosamine
Chondroitin sulfate A	D-glucuronic acid + N-acetyl-D-galactosamine-4-sulfate
Chondroitin sulfate C	D-glucuronic acid + N-acetyl-D-galactosamine-6-sulfate
Dermatan sulfate (Chondroitin sulfate B)	L-iduronic acid + N-acetyl-D-galactosamine-4-sulfate
Keratosulfate	D-galactose + N-acetyl-D-glucosamine-6-sulfate

Hyaluronic Acid

Hyaluronic acid is the straight-chain polymer of D-glucuronic acid and N-acetyl-D-glucosamine (NAG) alternating in the chain linked by α-1 $\rightarrow$ 3 and α-1 $\rightarrow$ 4 linkage (Figure 2.21). Its molecular weight is approximately 5,000,000 and is an acidic substance at cellular pH due to presence of largely ionized carboxyl groups. The most abundant member of mucopolyssacharides is hyaluronic acid which is chiefly found in the higher complex animals as a component of numerous tissues such as synovial joint fluids, vitreous part of eye and umbilical cord. Hyaluronic acid plays an important role in acting as biological lubricant through its high viscous nature in synovial fluid.

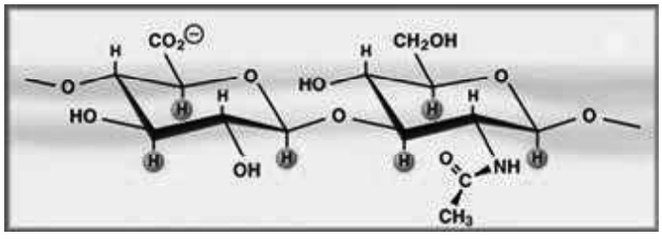


Fig. 2.21 *Showing Structure of Hyaluronic Acid*

Hyaluronic acid upon hydrolysis reaction its produces an equimolar volume of mixture of D-glucuronic acid, D-glucosamine and acetic acid. Hyaluronidase enzyme is involved in the depolymerization of hyaluronic acid and cleavage it to smaller fragments. This enzyme is also has a physiologic role in fertilization as they can move forward better in the cervical canal and finally can able to fertilize the ovum.

Chondroitin

Chondroitin is a polymer of α-D-glucuronido- 1, 3-N-acetyl D-galactosamine joined by α-1 $\rightarrow$ 4 glycosidic linkages (Figure 2.22). The structure of chondroitin is analogous to hyaluronic acid except that it comprises galactosamine in spite of glucosamine. It is a parent substance for two more broadly distributed mucopolyssacharides, namely chondroitin sulfate A and B. It is mainly found in cartilage with limited distribution and also serves as a component of cell coats.

NOTES

NOTES

Upon hydrolysis of Chondroitin, it produces an equimolar mixture of D-glucuronic acid, D-galactosamine and acetic acid.

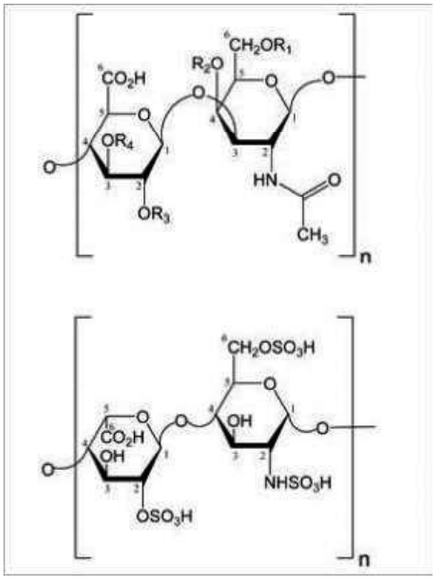


Fig. 2.22 Showing Structure of Chondroitin

Chondroitin Sulfates

Chondroitin sulfates may be considered as byproducts of chondroitin where, in the galactosamine moiety, a sulfate group is esterified either at 4th carbon as in chondroitin sulfate A or at 6th carbon as in chondroitin sulfate C (Figure 2.23). The two linkages includes α-1 → 3 and α-1 → 4 are involved in both types of chondroitin sulfate A and C. Upon hydrolysis, Chondroitin sulfates A and C both produces approximately equal amounts of D-galactosamine, D-glucuronic acid, sulfuric acid and acetic acid. They are more frequently associated with collagen and probably with other proteins too. Two chondroitin sulfate A and C are broadly distributed in the body and form chief structural components of bones, cartilage and tendons.

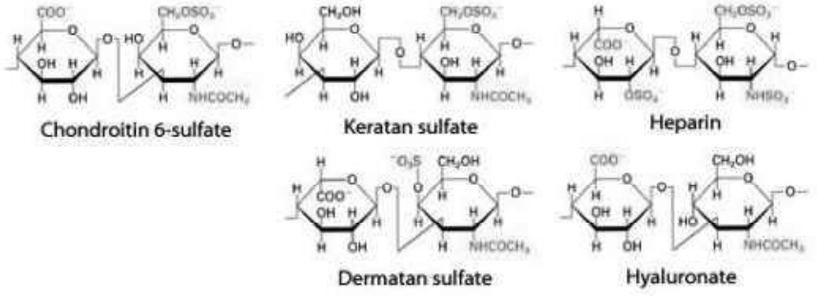


Fig. 2.23 Structure Comparison between different Mucopolysaccharide

Dermatan Sulfate

Dermatan sulfate is mucopolysaccharides structurally analogous to chondroitin sulfate A eliminating that the D-glucuronic acid is substituted by L-iduronic acid at C-5. The linkage involved in dermatan sulfate structure formation includes α-1 →

3 and α-1 → 4 glycosidic linkage (Figure 2.24). Dermatan sulfate is moreover known by its usual name, chondroitin sulfate B. Dermatan sulfate varies from composition of their repeating disaccharide unit both in the chondroitin sulfate A and C.

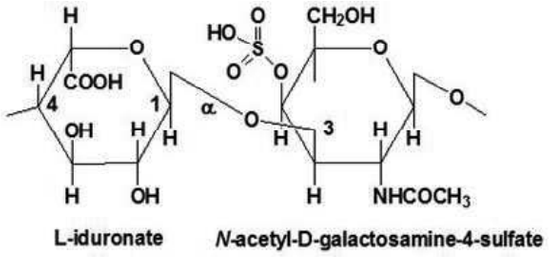


Fig. 2.24 *Structure of Dermatan Sulfate*

Keratosulfate

Keratosulfate contains D-galactose in place of uronic acid and second acetylated amino sugar component N-acetyl-D-glucosamine is esterified by a sulfate group at carbon number 6 by two alternating linkages involved are α-1 → 4 and α-1 → 3, in this case the linkage between the repeating disaccharide units is α-1 → 3 rather than α-1 → 4 (Figure 2.25).

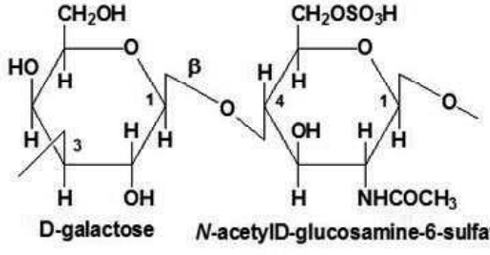


Fig. 2.25 *Structure of Keratosulfate*

Heparin

Heparin is a heteropolysaccharide comprised of D-glucuronic acid units, most of 7 out of every 8 are esterified at C2 and D-glucosamine-N-sulfate units with an additional O-sulfate group at C6 and are linked by alternating α-1 → 4 glycosidic bond (Figure 2.26). Therefore, the content of sulfate is very high and corresponds to about 5– 6 molecules per tetra-saccharide repeating unit. The relative locations of the sulfate residues may also differ. The molecular weight of heparin ranges between 17,000 and 20,000.

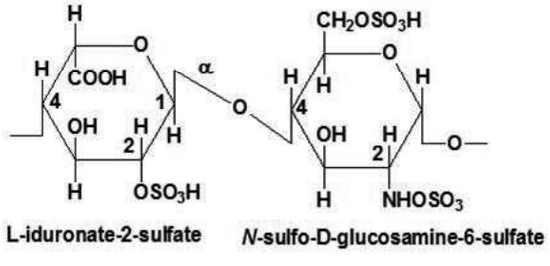


Fig. 2.26 *Structure of Heparin*

NOTES

Heparin mainly found in lung, liver, arterial walls and, certainly wherever mast cells are observe, probably for the determination of neutralizing biogenic amines e.g., histamine. It acts as an anticoagulant. It inhibits coagulation of blood by preventing the conversion of prothrombin to thrombin, leading to remove the effect of thrombin on fibrinogen.

The above described heteropolysaccharides were containing two different sugars in their repeating units or monomers. Though, there are numerous heteropolysaccharides which contain more than two carbohydrates in their repeating units. Vegetable 'gums' and agar-agar are two noticeable examples which are discussed below.

Vegetable 'gums'

Vegetable 'gums' comprise as many as four different monosaccharide units. Most common monosaccharides unit includes D-mannose, D-xylose (the second most abundant sugar in the biosphere), D-glucuronic acid, and L-rhamnose.

Agar-Agar

Agar consists of D- galactose and L-galactose, mostly with 1 $\rightarrow$ 3 glycosidic bonds and always contains certain amount of sulfuric acid (Figure 2.27). Upon hydrolysis, it yields D- and L-galactose in a ratio of 9:1. It is a gummy polysaccharide secreted by certain marine red algae of rhodophycean members such as species of Gracilaria, Gelidium, Hypnea, etc. It is soluble in hot water but remains insoluble in cold water. Agar is known to form highly viscous gels and its solidifying property at lower temperature. The melting point of agar varies between 90 to 100°F.

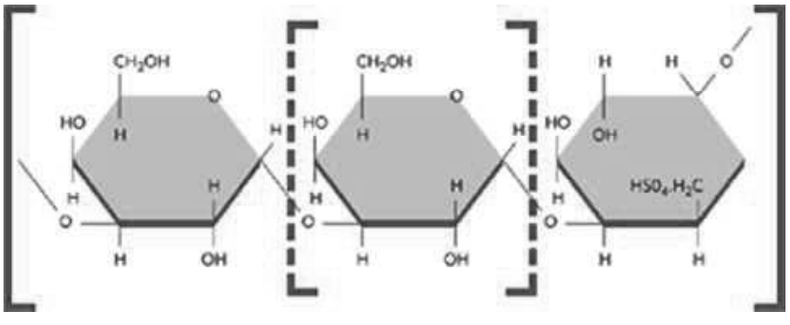


Fig. 2.27 *Structure of Agar*

Japan is the main producer of agar. The common use of agar-agar is for the preparation of culture media, especially for bacteria and fungi in biological laboratories. It is also used as laxative and sizing materials for textiles, as an emulsifier in dairy products, in the preparation of some medicines, and in cosmetics and leather industry. Agar-agar is not consumed by human beings and hence adds to the bulk of the faeces and helps in its propulsion.

2.5.4 Blood Group Substances

Blood group substances comprises the macromolecules containing highly specific, serologically active antigenic sites responsible for their immunological individuality. As their function is, they are very diverse molecules in terms of the different antigenic sites they carry and different molecules of them having the same sites. Humans

NOTES

have 60 blood group antigens which are further subdivided into fourteen independent genetic systems. The discovery of ABO blood groups by Landsteiner and associates was followed by various which led to the discovery of water-soluble substances with the same antigenic identity, in tissue fluids and secretions, particularly in mucus.

The most comprehensive are the monograph of Putkonen on tissue fluids and secretions and that of Hartmann on tissues and organs. The accumulation of all such data led to the classification of human beings into secretors and non-secretors and to the theory of the inheritance of the secretion of ABO (H) antigens. Consideration was then made on the isolation of water-soluble, mucus-borne entities, identified to bear the ABH and Lewis serological specificities. The substances of animal origin fit to the same family as those of man. Chemical and immunochemical investigations have recognized the structure of blood group substances as complex polysaccharides. As prime precursors for the isolation and purification, saliva or the contents of pseudo mucinous ovarian cysts were used in both labs.

After the experiments, human beings were classified as secretors and non-secretors and to the theory of the inheritance of the secretion of ABO (H) antigens. The Native substances shows numerous specificities so, a single molecule has specificity for AB, Lewis, and H. 'Secretors' are the one that carries the ABH and slight Lea activity (only in persons possessing the Lewis gene). In "nonsecretors," the ABH activity is almost absent from the saliva, ovarian mucin, etc., with a concomitant increase of Lea activity, while in the secretions of the gastrointestinal tract both ABH and Lea are found.

When the isolated and purifies substances are examined, they show very similar chemical and physicochemical similarity despite of the ABH and Lewis specificity. A characteristic substance is a mucopolysaccharide formed of carbohydrate and amino acids. Carbohydrate molecules in the substance can be n-galactose, L-fucose, N -acetyl-n-galactosamine, N-acetyl-glucosamine, N-acetyl-neuraminic acid in varying but small proportions.

The carbohydrate moiety is suggested to be related to the serological activity in the early stages of blood group investigation. The amino acid residue accounts for between 7 and 26 per cent of the total (20) and has to do nothing with the serological specificity but is capable of producing experimental hypersensitivity, with cross-reactivity obtainable with human materials irrespective of blood group. The quantitative composition is notable for the majority of four out of the sixteen amino acids present, viz. threonine, serine, proline, and alanine, accounting for about three fourths of the total. Their chemical diversity has the key role in the functions they play. Mucopolysaccharides explain the high degree of structural specificity and the vast diversity seen within one-and a cross-species.

The near-uniformness and orderly disappearance of the epithelial cell wall antigens during embryonal development link the function of the underlying macromolecule associated with, and limited to, the development of the epithelial anlagen e.g., in mutual cell recognition and adherence, and the presence of the antigens in stratified epithelia supports this idea. The soluble substances have shown their existence in the blood plasma irrespective of their secretion criteria, the role in blood group specificity. They play role in protecting the growing organism (foetus)

NOTES

by the diversion of maternal transplacental antibodies from their widespread tissue targets in the early embryonal life up to 35 to 40 mm stage and, in the case of incompatible pregnancies (0 mother/A, B father). If this protection is appeared late, it may explain the late high early intrauterine ovular loss in such pregnancies, variously estimated as 8 to 35 per cent. The cause of ABO erythroblastosis is also credited to soluble group specific substances which at least partially, diverts the maternal agglutinins.

Check Your Progress

5. What is the general formula of disaccharides?
6. What are polysaccharides?
7. What is cellulose?
8. How many blood group antigens do humans have?

2.6 GLYCOCONJUGATES

Glycoconjugates is the informational carbohydrate covalently linked to a protein or a lipid molecule. is They are the biologically active molecules and serve as target markers for some proteins and acts as a mediators for interactions between cells and the extracellular matrix, adhesion, cell-cell recognition, particular cell-cell interactions, cell migration during development, the immune response, wound healing and blood clotting. The common glycoconjugates includes proteoglycans, glycoprotein, and glycolipids since, their descriptive properties are elaborated below.

2.6.1 Proteoglycans

Proteoglycans are macromolecules of the extracellular matrix or cell surface in which one or more glycosaminoglycan chains are joined covalently to a secreted protein or a membrane protein. The glycosaminoglycan moiety usually forms the larger fraction (by mass) of the proteoglycan molecule, dominates the structure, and is often the major site of biological activity. In most of the cases the biological activity is the condition of multiple binding sites, affluent in opportunities for hydrogen bonding and electrostatic interactions with other proteins of the cell surface or the extracellular matrix. Proteoglycans are the most important constituents of connective tissue such as cartilage, in which their many noncovalent interactions with other proteoglycans, proteins, and glycosaminoglycans supply strength and flexibility. Mammalian cells can create at least 30 types of molecules that are members of the proteoglycan super family. These molecules may arbitrate the activities of various growth factors, act as tissue organizers, control the extracellular assemblage of collagen fibrils and manipulate the development of specialized tissues.

2.6.2 Glycoproteins

NOTES

The carbohydrate is attached at its anomeric carbon through a glycosidic link to the –hydroxyl (OH) group of a serine residues (Ser) or threonine (Thr) residue through *O*-linked or through an *N*- link glycosyl linkage to the amide nitrogen of an Asn residue (*N*-linked) (Figure 2.28). Few glycoproteins have a single oligosaccharide chain, but many have more than one oligosaccharide chain which may constitute from 1% to 70% or more of the glycoprotein by mass. The structures of a large number of *O*- and *N*-linked oligosaccharides from a variety of glycoproteins are known. The external surface of the plasma membrane has many membrane glycoproteins with arrays of covalently attached oligosaccharides of varying complexity.

Glycoproteins are affluent in information, forming highly specific sites for recognition and high-affinity binding by other proteins. They are commonly found in the extracellular matrix, on the outer face of the plasma membrane and in the blood. They are also found in specific organelles such as secretory granules, Golgi complexes, and lysosomes in the inside cells.

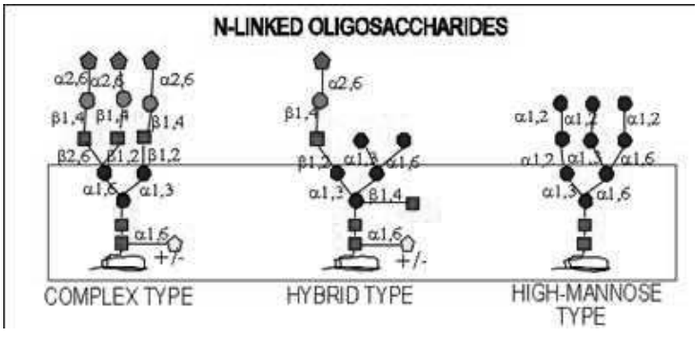


Fig. 2.28 *Glycoprotein Structure*

2.6.3 Glycolipids

The membrane lipids in which the hydrophilic head groups are oligosaccharides are called as glycolipids (Figure 2.29). It as specific sites for recognition by carbohydrate- binding proteins. Glycolipids molecules are most significant targets of the antibodies formed by the immune system of vertebrates in reaction to any bacterial infection and are therefore key determinants of serotype(strains that are differentiated on the basis of antigenic characteristics) of bacterial strains.

The outer membrane of Gram-negative bacteria such as Salmonella typhimurium and Escherichia coli contains glycoprotein leading to dominating surface characteristic of the outer membrane. The content of lipopolysaccharides in *S. typhimurium* bacteria include six fatty acids linked to two glucosamine residues, one of which is the point of attachment for a complex oligosaccharide. The lipopolysaccharides of a few bacteria are fatal to humans and other animals. For example, they are accountable for the hazardously lowered blood pressure that happens in toxic shock syndrome consequential from gram negative bacterial infections.

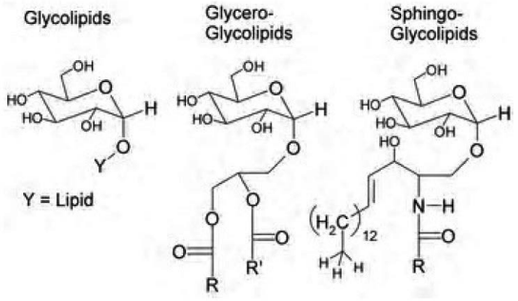


Fig. 2.29 *Showing Structure of Glycolipid*

2.7 ASCORBIC ACID

Ascorbic acid is also known as vitamin C. It is not produced or stored by the human body and needs to be supplemented from diet as it plays various important roles in humans. Naturally, it is found in citrus fruits, strawberries, broccoli, raw bell pepper, kiwifruit, brussels sprouts, leafy vegetables, potatoes, tomatoes, etc. It is also sold as a diet supplement. It is a water-soluble vitamin. It is white or slightly yellow crystal or powder and has an acidic taste. When the substance is exposed to light, it becomes dark. When Ascorbic acid is in a dry state, it is stable even in the presence of air but when mixed with a solution it oxidizes rapidly. So, it is used as a reducing agent and considered as an oxidant. The empirical formula of Ascorbic acid is $C_6H_8O_6$

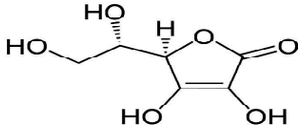


Fig. 2.30 *Ascorbic Acid*

Vitamin C is an important nutrient of certain animals including humans. The term vitamin *C* includes several vitamers that have vitamin C activity in animals. Ascorbate salts like sodium ascorbate and calcium ascorbate are present in dietary supplements. So, they upon digestion release ascorbate. Vitamin C functions act like a cofactor for many enzymatic reactions in animals (including humans) which control various essential biological functions like wound healing and collagen synthesis.

Table 2.4 *Properties of Ascorbic Acid*

Properties	
Chemical formula	$C_6H_8O_6$
Common name	Vitamin C and ascorbate
Molecular weight / ascorbic acid molar mass	176.12 g/mol
Density	1.694 g/cm^3
Melting point	190 °C
Boiling point	553 °C

NOTES

Its deficiency can lead to impaired collagen synthesis in humans contributing to the more severe symptoms of scurvy. It also acts an antioxidant (a reducing agent) by donating electrons to various enzymatic and non-enzymatic reactions. After reducing, it is converted to oxidized state either as semi dehydroascorbic acid or dehydroascorbic acid. Glutathione and NADPH-dependent enzymatic mechanisms further restore them to a reduced state.

Vitamin C is a substrate for ascorbate peroxidase in plants. This enzyme utilizes ascorbate to neutralize excess hydrogen peroxide (H_2O_2) by translating it to water (H_2O) and oxygen. The Recommended Dietary Allowance of amount of vitamin in blood serum is ≥ 50 μmol/L. There levels are considered saturated at > 65 μmol/L (1.1 mg/dL), Hypovitaminosis in the case seen when concentration of vitamin goes to μ 23 μmol/L and deficiency occurs at ≤ 11.4 μmol/L. Data from the U.S. 2003-04 NHANES survey shows that the mean and median serum concentrations of 49.0 and 54.4 μmol/L, respectively for 20 years of age or above. The percent of people reported as deficient was 7.1%. So, the daily intake of vitamin C is 40 mg/day for adults and 60mg/day for pregnant women.

Deficiency of vitamin C causes scurvy. In this condition, spores appear on the skin, bleeding is seen under the skin, it also causes spongy gums, 'corkscrew' hair growth, and poor wound healing. The skin lesions are mostly on the thighs and legs, and a person with the ailment looks pale, feels depressed, and is partially immobilized. In advanced scurvy there are open, suppurating wounds, loss of teeth, bone abnormalities and, eventually, death. It is a useful component of human body as it is not synthesised inside humans, it has various roles in medical field.

There are various functions of vitamin C like:

- **Health Benefits:** The damage caused to the body by smoking, radiations can be recovered by it due to its antioxidant properties. It helps in fixing worn-out and damaged tissues.
- **Personal Care:** The antioxidant property of vitamin C make it a useful component in cosmetics and personal care products like makeup products. It is also found in skin and hair care products.
- **Food and Beverage:** Ascorbic acid is also frequently used in preservatives, acidity regulators, colour fixative, nutrition supplements in food and beverages.
- **In Agriculture and Animal Fodder:** It is also a vital component of nutritional supplement in agriculture or animal and poultry food.

Side Effects of Ascorbic Acid

Symptoms like allergy, hives, swelling of your throat, lips, face, tongue, and difficulty in breathing shows the excess of vitamins.

- Weakness or tired feeling
- Stomach pain, joint pain, severe pain in lower or side back
- Fever and chills
- Weight loss

NOTES

- Difficult or painful urination, increased urge to urinate, and blood in the urine
- Heartburn
- Diarrhoea, nausea, and cramps in the stomach

Check Your Progress

9. What is the empirical formula of ascorbic acid?
10. Mention any two functions of vitamin C.

2.8 CARBOHYDRATE METABOLISM

Metabolism comprises the entire set of chemical reactions that occur in a living organism that allow it to reproduce, develop, maintain its structure and respond to the environment. These chemical reactions form an intricate network of pathways and cycles in which the flow of reaction products (metabolites) is determined by many regulatory mechanisms. Traditionally, metabolism is subdivided into catabolism, the breaking down of complex molecules, and anabolism, processes related to the synthesis of complex organic substances. According to the definition provided above, metabolism includes every cellular process, ranging from DNA replication to transcription and translation to enzyme function, and also involves the chemistry of small molecules in the cell. In the intermediary metabolism pathway, the structure of each enzyme plays a crucial role in determining the specific properties of each reaction. In the early seventies, Brenner introduced Caenorhabditis elegans (C. elegans) as a multicellular genetic model, mainly because this organism has very few somatic cells, which made it possible for investigators to reconstruct the cell lineage and to map the wiring of the nervous system.

Since then, it has become one of the most powerful tools for studying the genetics of metazoan development, neurobiology, aging and many other biological processes. This worm, however, is less amenable to classical biochemical study. It is sometimes difficult to obtain large quantities of synchronized individuals; eggshells and cuticles are tough barriers and it is virtually impossible to collect pure tissue in biochemically relevant quantities. Nevertheless, many metabolic studies of free-living and parasitic nematodes 3 have allowed us to gain deeper insights into the complex biochemical events that underpin the life of C. elegans.

Once in the blood stream, absorbed nutrients are distributed to the cells of the body, where they undergo by remarkable change. The sum total of these changes has been named metabolism. It includes energy requiring synthesis of new complex organic compound (anabolism) and energy degradation of absorbed nutrients to simple end products such as CO_2 and water (catabolism). When one is deprived of food, it has one overriding need—to maintain a supply of glucose to the blood, since a fall in blood glucose to below a critical level (about 2.5 mm in the human) leads to dysfunction of the central nervous system. In humans, this manifests as the symptoms of hypoglycemia, such as muscular weakness and

disharmony, mental confusion and sweating. If the blood glucose falls further, hypoglycaemic coma results.

NOTES

The immediate source of blood glucose is the store of glycogen in the liver. This can be broken down to release glucose into the bloodstream within seconds, a process known as glycogenolysis. Muscle glycogen constitutes a much larger store of glucose, does not directly contribute to blood glucose. Although glycogen is the immediate source of liver output of glucose, the liver can rapidly switch to producing glucose by de novo synthesis from ammo acids that are released from muscle and metabolized in liver a process known as gluconeogenesis. At the same time ketogenesis begins; this is ' a process in the liver that converts fatty acids, released from adipose tissue, to ketone bodies, which fan be utilized in some tissues to prevent excessive use of glucose. The release of fatty acids from the storage lipid, triacylglycerol, is known as lipolysis.

Adrenaline and glucagon activate the enzymes that release glucose from liver glycogen. Glucagon also activates the process of gluconeogenesis. Glucocorticoids activate the release of amino acids from the muscle, and adrenaline, Adreno Cortico Tropin Hormone (ACTH) and growth hormone activate lipolysis.

Carbohydrates supply 50 per cent of the energy requirement of the body. The metabolism of carbohydrate (Figure 2.31) may be sub divided as follows:

- **Glycolysis**: The oxidation of glucose or glycogen to pyruvate and lactate by Embden – Mayerhof pathway (EMP)
- **TCA Cycle**: The tricarboxylic Acid or Citric Acid or Krebs cycle is the final common pathway of oxidation of glucose, fatty acid and amino acid through which acetyl CoA is completely oxidized to CO_2 and water.
- **Glycogenesis**: The synthesis of glycogen from glucose.

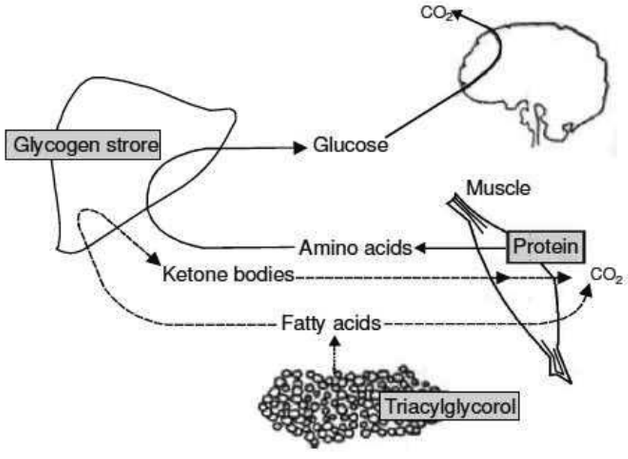


Fig. 2.31 *Metabolism of Carbohydrates*

- **Glycogenolysis**: The conversion of glycogen to glucose
- **Gluconeogenesis (Cori cycle)**: The formation of glucose from non-carbohydrates such as glycerol, pyruvic acid

NOTES

The major pathway of carbohydrate metabolism is shown in Figure 2.32.

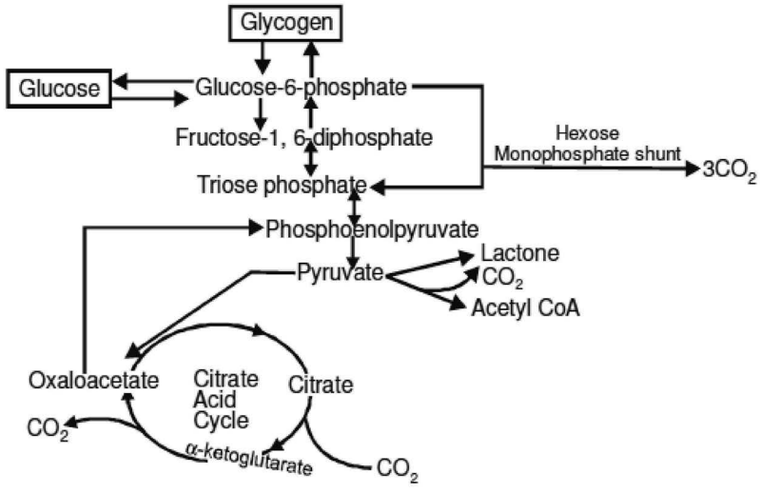


Fig. 2.32 *Major Pathway of Carbohydrates Metabolism*

2.8.1 Glycogenesis

The synthesis of glycogen from glucose can occur in most of the tissues of body. Liver and muscles are the most important sites of glycogenesis. Liver can also synthesize glycogen from monosaccharide other than glucose. In the absence of urgent demand of oxidative energy or conversion to any other compound, excess of glucose in thus converted to glycogen and stored in tissues. Glycogen is the only immediately available reserve glucose in fasting condition (Figure 2.33).

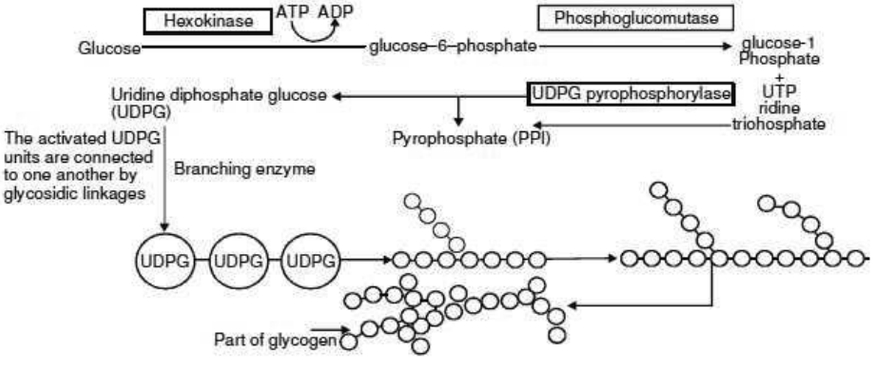


Fig. 2.33 *Glycogenesis*

NOTES

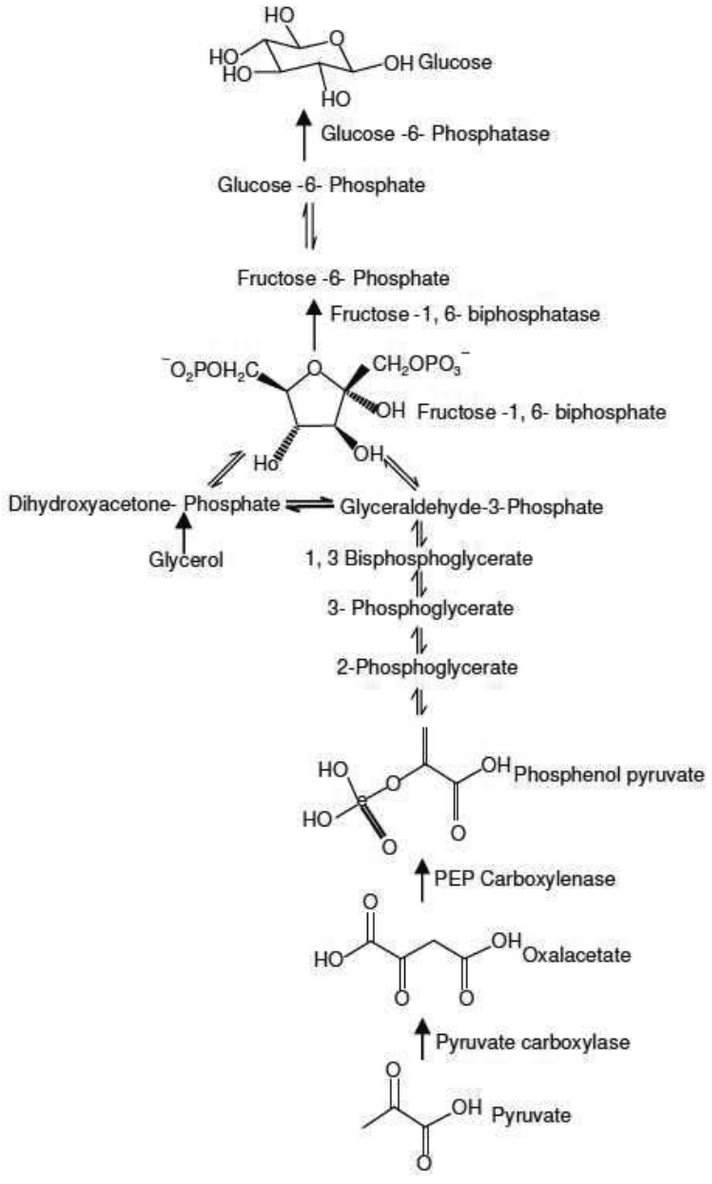


Fig. 2.34 *Schematic Representation of Glycogenesis*

2.8.2 Glycogenolysis

In this process, breakdown of glycogen to glucose takes place. Both muscles and liver glycogen undergo glycogenolysis (Figure 2.35). The enzyme responsible for this, phosphorylase, removes terminal glucose units one by one by phosphorolysis, a reaction analogous to hydrolysis, except that phosphate is the attacking group rather than water. The product is glucose 1- phosphate. Degradation of glycogen to glucose involves the following four enzyme reactions:

1. Phosphorylase, which hydrolyses a 1-4 bonds
2. A debranching step, which removes a 1-6 bonds

3. Phosphoglucomutase, which converts glucose 1-phosphate to glucose-6-phosphate
4. Glucose-6-phosphatase, which converts glucose 6-phosphate to glucose

NOTES

Phosphorylase action proceeds along the chain until it approaches a branch point. About three residues from the branch point its action ceases and glycogen-debranching enzyme hydrolyses the 0tI~6 bond. This enzyme has two activities associated with a single polypeptide chain: a transferase activity that transfers all except one of the remaining residues of the branch from the chain being degraded to the end of a longer chain, and a 1,6-glucosidase activity that hydrolyses the a 1-6 link, a reaction that releases free glucose, as shown in Glycogen degradation to glucose then requires conversion of the glucose 1- phosphate released by phosphorylase to glucose 6-phosphate, and conversion of this to glucose by glucosc-6-phosphatase. The effect of the combined action of all these enzymes is illustrated in Glucose-6-phosphatase is not present in muscle, so muscle glycogen cannot directly act as a source of blood glucose. Instead, it provides a source of energy for muscle contraction, the glucose 6-phosphate being acted on by enzymes of the glycolytic pathway.

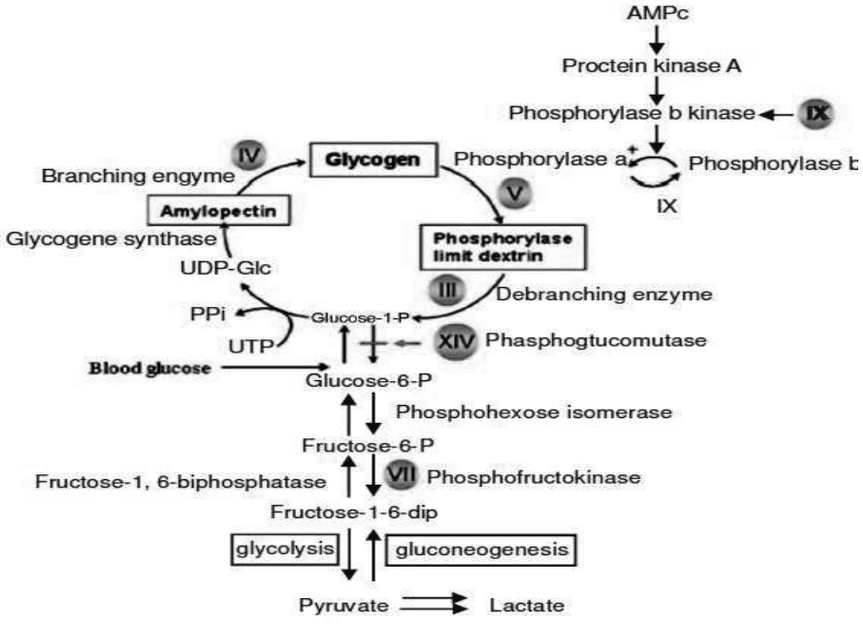


Fig. 2.35 *Schematic Representation of Glycogenolysis*

2.8.3 Gluconeogenesis—The Gluconeogenic Pathway (The Cori Cycle)

During fasting, the supply of glucose to the blood is dependent on de novo synthesis in liver, utilizing the carbon skeletons of amino acids derived from muscle, this process is known as gluconeogenesis. Protein breakdown in muscle to release the amino acids is activated by glucocorticoid hormones. The muscle to some extent assists the liver by converting many of the amino acids into alanine. When this arrives in the liver, the amino group is removed by transamination and converted

to urea, and the resulting pyruvate enters the gluconeogenic pathway. Kidney and small intestine also possess the capacity for gluconeogenesis, but participate to a lesser extent than liver. Figure 2.36 shows the outline of reactions of gluconeogenesis.

NOTES

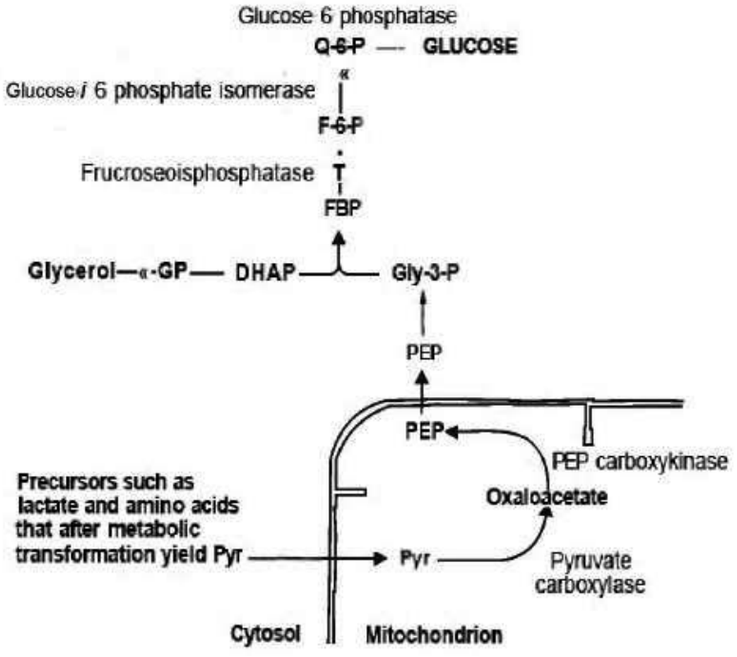


Fig. 2.36 *Outlines of Reactions of Gluconeogenesis*

The lack of glucose-6-phosphatase in muscle prevents the formation of free glucose in the tissue. However, muscle metabolism can contribute to blood glucose indirectly, in that lactate formed in muscle can be converted to glucose in the liver, as illustrated in the sequence of reactions shown in Figure 2.37. It has been termed as the Cori cycle, from the name of the husband and wife team, Carl Cori and Gerty Cori who first realized its significance.

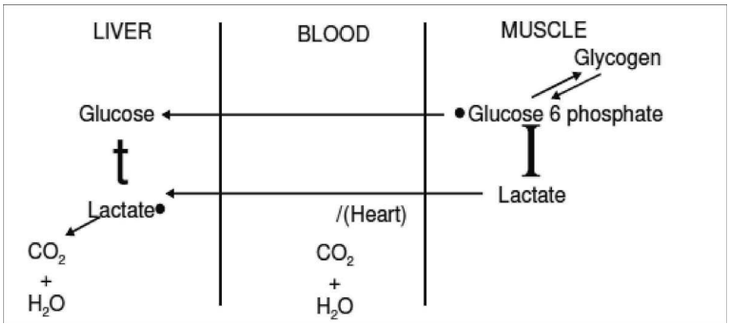


Fig. 2.37 *Cori Cycle*

Ketogenesis

Gluconeogenesis is usually accompanied by ketogenesis. Almost all of the fatty acids of body tissues have an even number of carbon atoms, and yield acetyl CoA on degradation. They thus cannot contribute to the net synthesis of glucose. In general, they are not used as a fuel by the brain. The ketone bodies synthesized

NOTES

from them during fasting can be used by many tissues as a fuel, thus sparing the utilization of glucose. During a prolonged fast, even the brain adapts to the use of ketone bodies. An outline of the interactions between tissues is shown in Ketone bodies are synthesized in the liver, which not itself utilizes them. The structures of the three compounds known as ketone bodies acetoacetate, 3-hydroxybutyrate and acetone - together with details of the pathway by which they are formed are given in all of the enzymes of ketogenesis are mitochondrial.

The scheme shows that three molecules of acetyl CoA are needed for the formation of hydroxymethylglutaryl CoA (HMG CoA), one of these being released again on its conversion to acetoacetate. 3-Hydroxybutyrate dehydrogenase is a mitochondrial enzyme that maintains equilibrium between 3-hydroxybutyrate and acetoacetate. Sometimes acetoacetate spontaneously decomposes to acetone. Thus occurrence of variations in the liver mitochondrial redox state affects the accuracy of the measurement as an indication of the total amount of ketone bodies.

Ketone Body Utilization

Our body produced in the liver utilized by peripheral tissues such a, muscle. During a prolonged fast, the brain adapts to the use of ketone bodies. The reactions involved. Most of these reactions relate to pathways that have already been discussed. The reaction that has not previously been described is that citric acid cycle intermediate. This reaction forms acetoacetyl CoA and succinate.

2.8.4 Glycolysis

D-glucose is the major fuel of most organisms and occupies a central position in metabolism. By storing glucose as a high molecular weight polymer, a cell can stock pile large quantities of hexose units while maintaining a relatively low cytosolic osmolarity. When the cells energy demand suddenly increase, glucose can be released quickly from these intracellular storage polymers. The sequence of reactions by which glucose is degraded anaerobically is called as glycolysis or glycolytic sequence. This refers to the production of two molecules of lactic acid from one mole of glucose.

$$C_6H_{12}O_6 \xrightarrow{\text{Glycolysis}} 2CH_3CHOHCOOH$$

Monosaccharides other than glucose can be broken down by glycolysis provided they can be converted into an intermediate in that sequence. Energy is released in the form of ATP as the monosaccharide is degraded and several important metabolites are produced for use elsewhere in the intermediary metabolism.

All organisms, with the exception of blue green algae, possess the ability to degrade glucose by this glycolytic process as far as pyruvic acid. Those cells and tissues that actually convert pyruvic acid to lactic acid as a major end product are much more limited. The examples are of the skeletal (white) muscle of animals, lactic acid bacteria and some plant tissues, i.e., potato tubers. Skeletal muscle with its poor oxygen supply and relatively few mitochondria but high concentration of glycolytic enzymes is ideally designed for carrying out glycolysis; heart muscle well supplied with oxygen and mitochondria will convert only small quantities of pyruvic acid to lactic acid.

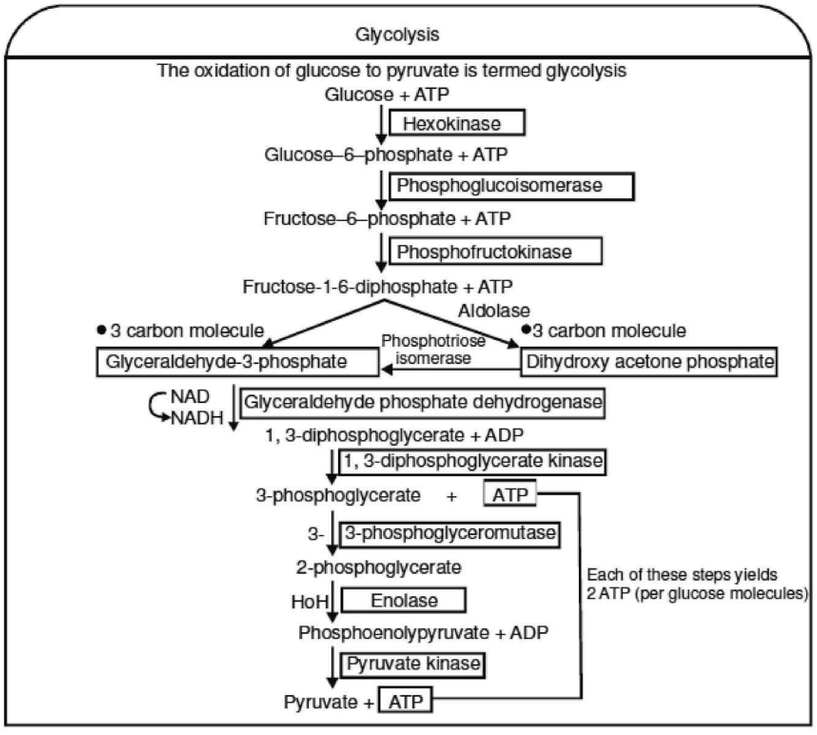


Fig. 2.38 *Reaction Sequence of Glycolysis (EMP) Pathway*

Reactions of the Glycolytic Sequence

Ten reactions are involved in the conversion of glucose to lactic acid; these are divided into two groups. The first four reactions are concerned with converting glucose into a compound D-glyceraldehyde 3-phosphate, whose oxidation subsequently releases energy to the cell. In contrast, the four reactions of the preparative phase require an expenditure of energy as the glucose molecule is phosphorylated prior to the formation of the glyceraldehyde 3 phosphates.

The initial step in the utilization of glucose in glycolysis is its phosphorylation by ATP to yield glucose 6 phosphate in the presence of hexokinase. In liver, instead of hexokinase another enzyme glucokinase is present. The activity of glucokinase is affected by nutritional status. This step is irreversible.

The next reaction in glycolysis is the issmerization of glucose 6 phosphate catalyzed by phosphoglucoisomerase. Fructose 6 phosphate is phosphorylated with ATP by phosphofructokinase to form fructose 1, 6 diphosphate. The enzyme requires Mg^{2+} and is an irreversible reaction. The next reaction in the glycotic sequence involves the cleavage of fructose 1, 6 diphosphate to form the two triose phosphate sugars, i.e., dihydroxy acetone phosphate and D-glyceraldehyde 3 phosphate. The finding of this enzyme in high concentration in a particular tissue is indicative of a functioning glycolysis pathway.

The production of D-glyceraldehyde 3 phosphate in the aldolase reaction technically completes the preparative phase of glycolysis. The second or energy yielding phase involves the phase involves the oxidation of glyceraldehyde 3 phosphate. Only half of the glucose molecule has been converted to D-glyceraldehyde 3 phosphates in the above reactions.

NOTES

If cells were unable to convert dihydroxyacetone phosphate to glyceraldehyde 3 phosphate, half of the glucose molecule would accumulate in the cells as the ketose phosphate or be disposed off by other reactions. This next reaction is the first in energy yielding or second phase of glycolysis, and is also the first reaction in the glycolytic sequence to involve oxidation reduction. It is also the first reaction in which a high energy phosphate compound has been formed. The following reaction accomplishes the transfer of the phosphate from the acyl phosphate formed in the preceding reaction to ADP to form ATP. In this reaction, the acyl phosphate group is utilized to drive the phosphorylation of ADP and make it ATP.

The next reaction in the degradation of glucose involves the dehydration of 2 phosphoglyceric acid to produce phosphoenol pyruvic acid, a compound with a high energy phosphate group. Pyruvate kinases catalyse the transfer of phosphate from phosphoenol pyruvic acid to ADP produce ATP and pyruvic acid.

Production of ATP

A net formation of high energy phosphate in the form of ATP is glycolysis. Table 2.5 gives the reaction sequence, which provides ATP.

Table 2.5 *Reaction Sequence*

Reaction	ATP Formed
Glyceraldehyde 3 phosphate 1, 3 Diphosphoglycerate	6
1, 3 diphosphoglycerate 3 phosphoglycerate	2
phosphoenol pyruvate "Enol pyruvate	2
Total	10

However, in glycolysis there is a net expenditure of 2 ATP molecules.

Table 2.6 *Net ATP Calculation*

Reaction	ATP Consumed
Glucose→Glucose 6 phosphate	1
Fructose 6 phosphate→Fructose 1, 6 diphosphate	1
Net ATP synthesized	10 – 2=8

Inhibitors of Glycolysis

There are different inhibitors which inhibit the enzymes involved in the glycolytic sequence or interfere with the mechanism of reactions. For example, Indoacetate is the inhibitor of glyceraldehyde 3 phosphate dehydrogenase involved in the conversion of glyceraldehyde 3 phosphate to 1, 3 diphosphoglycerate. Arsenate inhibits the synthesis of ATP by accomplishing incoupling of oxidation and phosphorylation in the conversion of 1, 3 diphosphoglycerate to 3 phosphoglycerate. Fluoride inhibit enolase involved in the conversion of 3-phosphoglycerate to 2 phosphoglycerate.

NOTES

Effect of Hormones in Glycolysis

Insulin and glucagon hormones are two important hormones which greatly influences theoverall carbohydrate metabolism. Insulin stimulates hexokinase and glucokinase, which catalyze the conversion of glucose to glucose 6 phosphate. It also stimulates phosphofructokinase which catalyse the conversion of fructose 6 phosphate to fructose 1, 6 diphosphate. Glucagon stimulates liver glucose 6 phosphatase which is involved in the conversion of glucose 6 phosphate to glucose and also fructose 1, 6 diphosphatase involved in the conversion of fructose 1, 6 diphosphate to fructose 6 phosphate.

Regulation of Glycolysis

Glycolysis is regulated in all cells so that energy is released from carbohydrates only as it is needed by those cells. This was shown that some tissues that possess not only the capacity to convert glucose to lactate by glycolysis, but also can oxidize pyruvic acid completely to CO_2 and H_2O via the Krebs cycle. Such tissues utilize glucose much more rapidly in the absence of O_2 than when they do when O_2 is present. The functional significance of this inhibition of glucose consumption by oxygen is known as Pasteur Effect. This is appreciated when much more energy is made available as ATP when glucose is oxidized aerobically to CO_2 and H_2O than, when it is anaerobically converted only to lactic acid or alcohol and CO_2. Since more ATP is formed under aerobic conditions, less glucose is needed to be consumed to do the same amount of work in the cell.

Glycolysis in muscle is regulated at the reactions catalysed by the following enzymes:

- Phosphorylase (for glycogen) and hexokinase (for glucose)
- Phosphofructokinase and pyruvate kinase

In a reciprocal manner, gluconeogenesis in which carbon flows back into carbohydrate from lactate or pyruvate, is regulated at the steps catalyzed by pyruvic caboxylase, fructose 1, 6 diphosphatase glucose 6 phosphatase and glycogen synthetase. The regulation occurs at those reactions in glycolysis and gluconeogenesis that are unidirectional. Thus, glycolysis is not only the main pathway for glucose metabolism leading to the formation of acetyl CoA and oxidation in the citric acid cycle, but it also provides the principal pathway for the metabolism of fructose and galactose derivedfrom the diet. It produces ATP in absence of oxygen because this allows skeletal muscle to perform efficiently when aerobic oxidation becomes insufficient and it allows tissues with significant glycolytic ability.

Energetics of Glycolysis

Calculating the energy balance sheet for a single molecule of glucose under-going glycolysis there is an overall gain of 4 ATP molecules and 2 NADred molecules. Since two ATP molecules are used during the glocolysis process, there is a net gain of 2 ATP molecules. Pyruvic acid can dissociate to produce pyruvate and a hydrogen ion (proton).

These two forms, including pyruvic acid and pyruvate, exist in dynamic equilibrium and both the terms are used interchangeably. Two molecules of pyruvate formed still contain a large amount of energy stored within the molecule, and

NOTES

glycolysis is virtually carried out by all the living cells including prokaryotes, eukaryotes, plants, animals and man.

One mole of glucose contains 686 kcal of energy. That is, if a mole of glucose is burned completely, 686 kcal will be released as heat. Two moles of pyruvic acid have a total energy content of about 547 kcal. Therefore, the energy released during glycolysis is about 139 kcal per mole of glucose. If this reaction took place all at once, then all the energy released would be dissipated. However, the breakdown of glucose is accomplished by a series of separate reactions, each of which is catalyzed by a different enzyme. The chemical modification that occurs at each step in the series is very small. The molecule is broken apart, a little at a time. In this way, only a small amount of energy is released each time a chemical bond is broken. If the energy of the molecule was liberated at all once, then it would have produced heat. This heat would be of little use to the cell and could be rather damaging. By controlling the release of energy, the cell loses only a portion of it as heat and is able to conserve or recover much of it in the form of useful chemical energy.

The energy difference between ADP and ATP is approximately 7 kcal per mole. Therefore, in the oxidation of glucose to pyruvic acid, about 14 kcal are captured in the form of ATP. Two molecules of ATP are generated from ADP and Pi in the course of glycolysis or fermentation—either the formation of lactic acid or ethyl alcohol. During this process, about 7 per cent of the total available energy of the glucose molecule and about 52,000 cal are released and 93 per cent still conserved in the glycolytic or fermentative end products. Of the 52,000 cal released, only 14,000 cal are trapped in the form of 2 ATP molecules. Therefore, energetically glycolysis is not at efficient process. The fact that the glycolytic sequence does not require oxygen suggests that this series of reactions evolved early, before free oxygen was available and the primitive heterotrophs used glycolysis to extract energy from the organic compounds. Summing up, an anaerobic cell or organism is primarily in need of larger quantities of glucose or other substrate, per unit time, in comparison to an aerobic cell. Consequently an aerobic cell has a better 'equipped' machinery to supply maximum amount of energy than its anaerobic counterpart.

Oxidation of Pyruvate to Acetyl CoA

The pyruvic acid is not an intermediate in the tricarboxylic acid cycle. The keto acid is first converted to acetyl CoA by the multienzyme complex known as pyruvic dehydrogenase complex. This conversion which is known as a -oxidative decarboxylation is carried on in the mitochondria following the formation of pyruvic acid in the cytosol during glycolysis. The reaction involves six cofactors coenzyme A, NAD, lipoic aicd, FAD, Mg^{2+} and thiamin pyrophosphate. This is an irreversible step. The pyruvate dehydrogenase complex is constituted of three enzymes:

1. **Pyruvate Dehydrogenase:** It brings about decarboxylation of pyruvic acid.
2. **Dihydrolipoyl Transacetylase:** S acetyl lipoate reacts with coenzyme A to form acetyl CoA and reduced lipoate in the presence of this enzyme.
3. **Dihydrolipoyl Dehydrogenases:** The reduced lipoate is re-oxidized by FAD in presence of this enzyme.

NOTES

Formation of Acetyl CoA

Pyruvic acid represents the terminal product of the glycolytic process. If sufficient oxygen is available, then pyruvic acid undergoes oxidative decarboxylation to form a two-carbon compound sailed acetyl CoA. The reaction takes place in a sequence and requires five cofactors, namely: (i) Thiamine Pyrophosphate (TPP); (ii) Coenzyme A (CoA); (iii) Mg+ ions, (iv) NAD+; and (v) Lipoic Acid. The reaction is exothermic and a molecule of NADPH is produced from NADOX.

The formation of acetyl CoA may be summarized as follows:

Formation of a TPP complex between TPP and pyruvate: Formation of acetyl lipoic acid complex as a result of interaction between TPP complex (from Step 1) and oxidized lipoic acid. Release of acetyl group from lipoic acid and combination with CoA forms acetyl CoA. Besides, a reduced form of lipoic acid is also produced.

Formation of Oxidized Lipoic Acid Using NAD+ as an Electron Acceptor

Coenzyme A is a large molecule, a portion of which is a nucleotide (composed of adenine, ribose and a pyrophosphate bridge) and the other portion is a vitamin pantothenic acid— a B complex vitamin .

The overall reaction for the formation of acetyl CoA may thus be written as:

$$CH_3COCOOH + NAD_{ox} + CoA - SH \rightarrow CH_3CO - S - CoA + CO_2 + NAD_{red}$$

The electrons removed from pyruvic acid are accepted by NADOI and eventually carried on to the respiratory chain. The standard free energy of hydrolysis of acetyl CoA at pH 7.0 is—7.5 kcal. The acetyl group carried by CoA — SH is, therefore, in a highly activated form and may be enzymatically transferred to various acetyl group acceptors. The acetyl group of acetyl CoA is thus a 'fuel' of the Krebs cycle and can be made available by an enzymatic transfer reaction. Besides glucose, fat and amino acids can also be converted to acetyl CoA. A fat molecule is first hydrolysed to glycerol and three fatty acids. Then, successive two-carbon groups are removed beginning at the carboxyl end. A molecule such, as Palmitic Acid [$CH_3 (CH_2)_{14}$ COOH] which contains sixteen carbon atoms yields eight molecules of acetyl CoA.

Release of Energy in Glycolysis and Formation of Pyruvic Acid

Glycolysis literally means 'dissolution of sugar'. All the intermediary products formed during the glycolysis are esters of phosphoric acid except the initial reactant, glucose and the final product, pyruvic acid. Glucose and pyruvic acid are freely permeable through the cell membranes. However, the phosphate esters of the glycolytic cycle are unable to penetrate through cell membranes. All the eleven enzymes which catalyze the glycolytic sequence exist free in solution in the soluble portion of the cytoplasm.

Glycolysis is a primary pathway in higher organisms as well as in several bacteria. However, some bacteria have evolved quite different pathways. Glycolysis takes place in the ground substance of the cytoplasm and involves three important events. The energy-rich glucose molecule is broken down into two simple and

NOTES

relatively energy-poor molecules of pyruvic acid. Formation of ATP from ADP with the addition of inorganic phosphate (Pi). This phosphorylation that takes place during glycolysis is called substrate phosphorylation. Formation of two molecules of NADr«j (NADH + H+) from each molecule of glucose during glycolysis. Thus, glycolysis ends up with the formation of two molecules of pyruvic acid, two molecules of ATP and two molecules. The various steps in the process of glycolysis and formation of pyruvic acid can be given as follows: Conversion of glucose to fructose 1, 6-diphosphate. This conversion requires three important enzymes including hexokinase for converting glucose to glucose 6-phosphate, phosphoglucoiso-merase for isomeric conversion of glucose 6-phosphate to fructose. 6-phosphate and phosphofructokinase for converting fructose 6-phosphate to fructose 1, 6-diphosphate.

Formation of fructose 1, 6-diphosphate is followed by its splitting into two molecules of three carbon compounds, 3-phosphogly-ceraldehyde and dihydroxyacetone phosphate. The reaction is catalysed by aldolase. Both the products formed are intercovertible with the help of an enzyme phosphotriose isomerase. However, because the glyceraldehyde phosphate is used up in subsequent reactions, all of the dihydroxyacetone phosphate is eventually converted to glyceraldehyde phosphate. With the termination of this step, the preparatory reactions that require an input of ATP energy are complete. The reaction is oxidative and NADPH becomes reduced. This is the first reaction from which the cell gains energy and is catalysed by the enzyme phosphoglyceraldehyde dehydrogenase.

Then, the high-energy phosphate is released from the diphosphoglyceric acid and is used to recharge a molecule of ADP. This causes the release of two molecules of ATP per molecule of glucose. The reaction is highly exothermic and pulls all other previous reactions forward and 3-phosphoglyceric acid is formed.

The remaining phosphate group is transferred from the third position to the second position resulting in the formation of 2-phosphoglyceric acid. The reaction is catalysed by an enzyme phosphoglyceromutase. A molecule of water is removed from 2- phosphoglyceric acid res-ulting in the formation of phosphoenolpyruvate possessing a high energy phosphate bond.

The high energy phosphate is transferred to a molecule of ADP forming another molecule of ATP. Here also, two molecules of ATP are formed per molecule of glucose. The reaction is catalysed by an enzyme called pyruvic kinase and the product formed is pyruvic acid which marks the termination of glycolysis. The sequence of reactions leading from glucose to pyruvic acid is often who postulated and experimentally proved various steps in this pathway during 1920–1930.

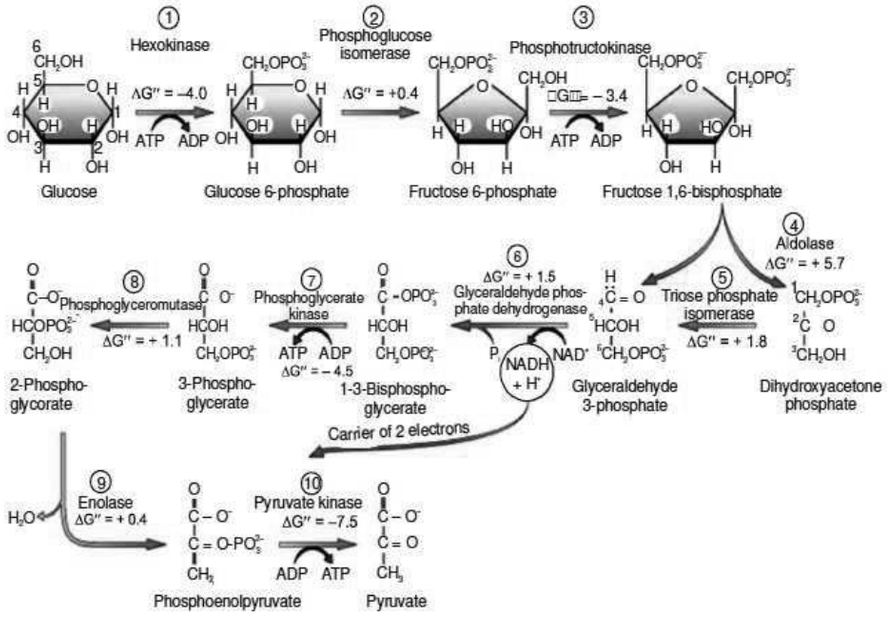


Fig. 2.39 *Schematic Representation of Glycolysis*

2.8.5 Tricarboxylic Acid (TCA) Cycle or Kreb's Cycle

The tricarboxylic acid or citric acid cycle is the process in which acetate (in the form of acetyl CoA) is oxidized completely to CO_2 and water. Since acetyl CoA is readily produced from pyruvate the cycle is also the process in which the oxidation of glucose to CO_2 and H_2O is completed. The electrons are removed from the substrates as they are oxidized and transferred eventually to molecular oxygen; thus the process is aerobic one, the reactions of cycle, which accounted for the oxidation of pyruvic acid to CO_2 and H_2O were contributed by distinguished. English biochemist Sir Hans Kreb, and so the cycle is also called as Kreb cycle.

The reactions of tricarboxylic acid cycle are:

- The synthesis of citric acid from acetyl CoA is the first reaction in the Krebs cycle. This reaction is catalyzed by citrate synthetase, and it is found in the matrix of mitochondrion.
- The reaction of interest that is catalysed by aconitase is the interconversion of citric acid to isocitric acid.
- Isocitric acid dehydrogenase catalyze the oxidative p-decarboxylation of isocitric acid to 3-ketoglutaric acid and CO_2 in the presence of a divalent cation Mg^{2+} or Mn^{2+}), and a mcotinamide nucleotide as the oxidant. The evidence, however indicates that oxalosuccinate, if formed is firmly bound to the surface of the enzyme and is not released as a free intermediate in either the oxidative decarboxylation of isocitrate or the reverse reaction, the reductive carboxylation of a-ketoglutarate.
- The next step of tricarboxylic acid cycle involves the formation of succinyl CoA by the oxidative a - decarboxylation of a ketoglutaric acid. This reaction is catalysed by the a ketoglutaric dehydrogenase complex which requires

NOTES

TPP, Mg^{2+}, NAD+, FAD, lipoic acid and coenzyme A as cofactors. Succinic dehydrogenase catalyses the removal of two hydrogen atoms from succinic acid to fumaric acid. The immediate acceptor (oxidizing agent) of the electrons is a flavin coenzyme (FAD) which in contrast to other flavin enzymes is bound to succinic dehydrogenases through a covalent bond. Succinic dehydrogenases are firmly associated with the inner mitochondrial membrane and are rendered soluble with difficulty.

- The next reaction is the addition of H_2O to fumaric acid to form L – malic acid.
- The tricarboxylic acid cycle is completed with the oxidation of L – malic acid to oxaloacetic acid which is accomplished by the enzyme malic dehydrogenase. The reaction is the fourth oxidation reduction reaction to be encountered in the cycle. Figure 2.40 gives you a diagrammatic representation of the tricarboxylic acid (TCA) cycle.

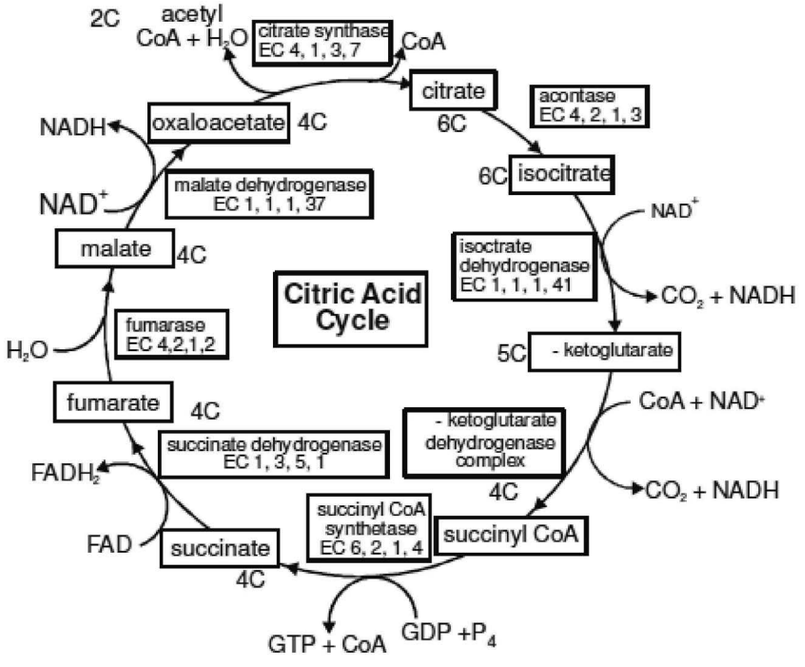


Fig. 2.40 *Reactions of Tricarboxylic Acid Cycle*

Formation of citric acid and release of CoA

Before the beginning of the Krebs cycle reactions, there is a condensation reaction between acetyl CoA and oxaloacetic acid. The condensation product is a six-carbon compound called citric acid which is a tricarboxylic acid. During this condensation reaction CoA is released.

Regeneration of Oxaloacetic Acid and the Krebs Cycle

Tricarboxylic acid cycle commonly known as the Krebs cycle was elucidated in 1937 by an English biochemist, Sir Hans Krebs. It was called a cycle because of the cyclic manner in which oxaloacetic acid is regenerated. Later on, in 1948, Kennedy and Lehninger discovered that the complete oxidation of pyruvate and all other intermediates of tricar-boxylic acid take place in the mitochondria. This

was another important achievement because it determined the site of tricarboxylic acid cycle.

Hans Krebs was awarded the Nobel Prize in 1953 for formulating the tricarboxylic acid cycle. Various steps in the Kreb cycle may be explained as follows:

NOTES

- There is transconversion of citric acid, cis-aconitic acid and isocitric acid. The reaction involves the addition and removal of water molecule and is catalysed by the enzyme aconitase.
- Conversion of isocitric acid to oxalosuccinic acid. The reaction is catalysed by an enzyme isocitric acid dehydrogenase and requires NADP+. The reaction is oxidative and it removes two electrons and two protons from isocitric acid. NADP+ acts as an electron acceptor and is reduced to NADP 4 H+.
- Decarboxylation of oxalosuccinic acid to form a-ketoglutaric acid. The reaction is catalysed by a carboxylase. a-ketoglutaric acid is one of the most important intermediate of the Kreb cycle and is the major source of most of the amino acids.
- a-Ketoglutaric acid undergoes oxidation and this process requires thiamine pyrophosphate for initial decarboxylation. Besides, it also requires oxidized lipoic acid. The reaction ultimately releases succinyl moiety which combines with CoA forming succinyl CoA and is catalysed by enzymes x-ketoglutaric dehydrogenase.
- Formation of succinic acid from succinic CoA. The reaction is oxidative and there is concomitant release of CoA. The excess energy released is immediately stored in the form of guanosine triphosphate (GTP).
- Oxidation of succinic acid to fumaric acid. The reaction is unique because it does not require NAD+ like other reactions of the Krebs cycle. Instead, it requires flavin adenine dinucleotide (FAD) as an electron acceptor and is catalysed by succinic dehydrogenase. The enzyme succinic dehydrogenase is very characteristic of the Krebs cycle and is apparently absent in cyanobacteria which have an incomplete Kreb cycle.
- Fumaric acid formed from the succinic acid is hydrated to yield malic acid and the reaction is catalysed by the enzyme fumarase.
- Conversion of malic acid to oxaloacetic acid, NAD+ is used to accept protons forming NADH+ H+. With the regeneration of oxaloacetic acid, the cycle is complete. The reaction is catalysed by the enzyme malic dehydrogenase. Keeping an overall estimate of the cycle, one molecule of the two-carbon compound acetic acid combines with one molecule of the four carbon compound oxaloacetic acid. Two atoms of carbon appear as two molecules of CO_2 and one molecule of the four-carbon compound oxaloacetic acid is regenerated. These reactions thus describe the fate of the carbon skeleton of acetic acid. The regeneration of oxaloacetic acid at the end of each turn of the cycle is of considerable importance be-cause one molecule of oxaloacetic acid can bring about the oxidation of an infinite number of acetic acid molecules to CO, and water.

NOTES

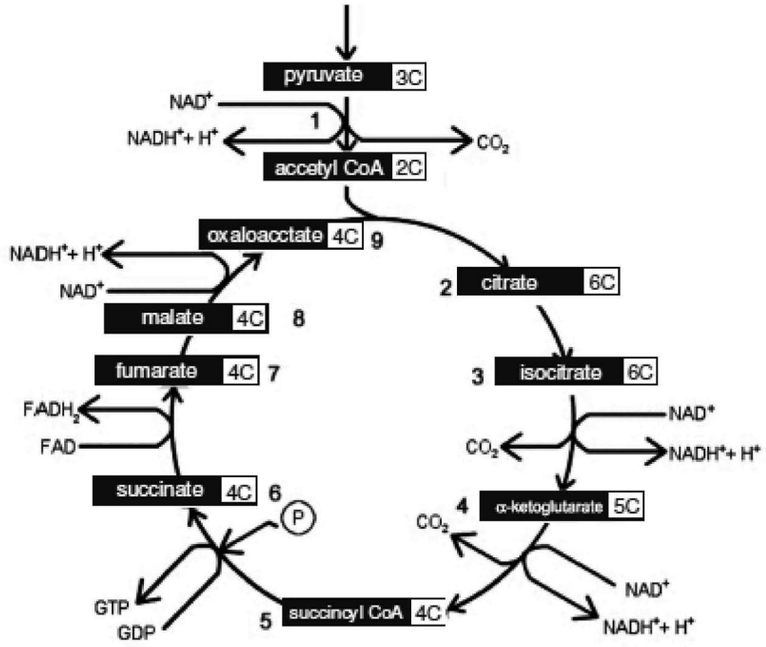


Fig. 2.41 *Schematic representation of Krebs Cycle*

Amphibolic Pathway

The primary function of TCA cycle is the oxidation of acetate to CO_8 with concomitant conservation of energy as ATP. In this category, acetyl CoA is the only substrate which enters the Krebs cycle. However, in some other cells, there is a substantial flux of four, five-, and six-carbon intermediates into and out of TCA cycle. These anabolic reactions are extreme importance to meet the cellular demand for various biomolecules. Therefore, TCA cycle can have both catabolic and anabolic reactions and such pathways are referred to as amphibolic. As a matter of fact, several biosynthetic pathways draw their raw material from intermediates in the Krebs cycle.

Amphibolic Nature of TCA Cycle

The Krebs cycle is the primary site of the oxidation of carbon atoms derived from food. It is sometimes called the 'final common pathway' because all carbon atoms from food are oxidized here.

Net Stoichiometry:

$$\text{Acetyl CoA} + 3\text{NAD+} + \text{FAD} + \text{GDP} + \text{Pi} + \text{H2O}\,,$$
$$2\text{CO2} + 3\,\text{NADH} + \text{FADH2} + \text{CoA–SH} + \text{GTP} + 3\text{H+}$$

Anaplerotic Reactions of TCA Cycle

Due to the amphibolic nature of the cycle, sometimes the supply of intermediates can become limiting. Remember than when we add 2 carbons as Acetyl CoA, 2 carbons come out as CO_2, and thus we can't add to the carbon pool of the Krebs cycle with Acetyl CoA. A corollary to this is that we can't make glucose from Acetyl-groups. There are three reactions that can add carbon to the Krebs cycle:

(i) Pyruvate carboxylase (liver and kidney):

$$\underset{(3C)}{\text{Pyruvate}} + CO2 + ATP = \underset{(4C)}{OAA} + ADP + Pi$$

(ii) This enzyme is allosterically activated by Acetyl CoA. Thus, when the supply of Acetyl CoA is high, more OAA is made so that it can enter the cycle (this will be important later). As with most carboxylases, this enzyme requires biotin as a cofactor.

NOTES

Phosphoenolpyruvate carboxykinase (PEPCK) (heart and skeletal muscle):

$$PEP + CO2 + GDP = OAA + GTP$$

(iii) Malic enzyme (widely distributed-cytosolic in mammals):

$$\text{Pyruvate} + NADPH + H+ = \text{Malate} + NADP+ + CO2$$

Electron Carriers

In the Krebs cycle, there are four dehydrogenation steps. In three steps, NADO, serves as the electron acceptor for the specific dehydrogenases and, consequently, three molecules of NAD red are formed in each turn of the cycle. In the fourth step, FAD acts as an electron acceptor. FAD (a flavoprotein) contains flavin, nucleotide instead of a pyridine nucleotide (NAD+) and has riboflavin or vitamin B2 as a building block. These four electron carriers, including three NADred molecules and one FADP. Molecule, are now in a position to donate their electrons to another series of enzymes that constitute the respiratory chain. It is through this electron-transport or respiratory chain that these electron carriers are reoxidized and the energy released during oxidation is conserved as ATP.

The electron transport system is composed of a series of cytochromes as electron-transferring enzyme molecules containing heme which consists of porphyrin and iron. However, before electrons are accepted by cytochromes, another common acceptor called coenzyme a, accepts the electron and is then transported to series of cytochromes. Cytochromes were first reported by McMunn in 1886 and rediscovered by Keilin in 1927. Keilin demonstrated that cytochromes were present in almost all the living organisms and designated them as a, b, c. These types are distinguished from one another on the basis of the absorption spectrum. Cytochromes are conjugated proteins having an iron porphyrin as a prosthetic group. Cytochrome is surrounded by a protein made up of about 100 amino acids. Cytochromes differ in their protein chains and also in the energy levels at which they hold electrons.

Characteristically, they function as part of an electron transport chain in which a cytochrome will accept an electron from a compound that holds it at a higher level and will pass it to a cytochrome that holds it at a slightly lower level, which then passes it 'downhill' to yet another cytochrome. It is the iron that actually combines with the electrons, each iron atom these studies have considerably enhanced the understanding of the electron-transport system of the respiratory chain. Of immense use has been the sequence of electron carriers, which became possible because of such investigations.

NOTES

Production of ATP

The following steps in Table 2.7 summarize the net production of ATP in the TCA cycle.

Table 2.7 Total ATP Production in TCA Cycle

Reaction	No. of ATP Formed
Isocitrate → Oxalosuccinate	3
ketoglutarate → Succinyl CoA	3
Succinyl CoA → Succinate	1
Succinate → Fumarate	2
Malate → Oxaloacetate	3
Total ATP	12

Total number of ATP produced in the complete oxidation of one molecule of glucose. One molecule of glucose forms two molecules of pyruvic acid by glycolysis.

Table 2.8 Net Gain of ATP

Number of ATP formed in glycolysis	8
Number of ATP formed in the oxidation of pyruvate to acetyl CoA	3*2=6
Number of ATP formed in the citric acid cycle	12*2 = 24
Total	38 ATP

Thus, the overall reaction of aerobic respiration is

$$\underset{\text{Glucose}}{C_bH_iA_oO_o} + 6O_2 + 36 \text{ or } ADP + 36 \text{ or } 38 \rightarrow$$

$$\mathbf{6CO_2 + 6H_2O + 36 \text{ or } 38 \text{ ATPs}}$$

Inhibitors of TCA Cycle

The inhibitors either inhibit the reactions or inhibit the enzymes involved in the TCA cycle. For example, Fluoroacetate inhibits the enzyme aconitase and prevents the conversion of citrate to isocitrate. Arsenite - inhibits a ketoglutarate dehydrogenase and causes a ketoglutarate to accumulate. Malonate or oxaloacetate inhibits succinate dehydrogenase competitively resulting in succinate accumulation.

Regulation of TCA Cycle

A continuous supply of oxidized NAD is required to permit the Kreb cycle to operate. The enzymes of the electron transport chain carryout this vital activity and the concomitant process of oxidative phosphorylation. Where these process are inhibited, the Kreb cycle cannot function. Pyruvic dehydrogenase, which provides a supply of acetyl CoA for oxidation via the Krebs cycle is inhibited when the level of NADH or acetyl CoA builds up. Additional control may also be exerted by cyclic AMP produced when the concentration of ATP increases. Citrate synthase is under fine control both ATP and NADH can inhibit this initial reaction of the Kreb cycle. The inhibition of ATP and NADH on isocitric dehydrogenase

have also been noted. Thus the energy charge of the cell can readily affect the rate at which the tricarboxylic acid cycle operates.

2.8.6 The Pentose Phosphate Pathway (Hexose Monophosphate Shunt)

NOTES

Another pathway of carbohydrate metabolism that is of importance in the absorptive state (although it also has other functions) is the pentose phosphate pathway, or pentose shunt. It was given the name pentose shunt at the time it was being elucidated, because it was seen as an alternative to glycolysis as a pathway of glucose degradation, but it is now seen as a pathway having special functions of its own. One of these functions is to provide pentoses for nucleotide and nucleic acid synthesis. Another function is to catalyse reduction of NADP+ to NADPH to support biosynthetic processes, such as fatty acid biosynthesis, that require NADPH.

Glucose-6-phosphate dehydrogenase catalyzes oxidation of the aldehyde (hemiacetal), at C1 of glucose-6-phosphate, to a carboxylic acid in ester linkage (lactone). NADP+ serves as electron acceptor. 6-Phosphogluconolactonase catalyzes hydrolysis of the ester linkage (lactone) resulting in ring opening. The product is 6-phosphogluconate. Although ring opening occurs in the absence of a catalyst, 6-Phosphogluconolactonase speeds up the reaction, decreasing the lifetime of the highly reactive and thus potentially toxic, 6-phosphogluconolactone. Figure 2.42 gives the structures of Glucose-6-phosphate, 6-phosphogluconolactone and 6-phosphogluconate.

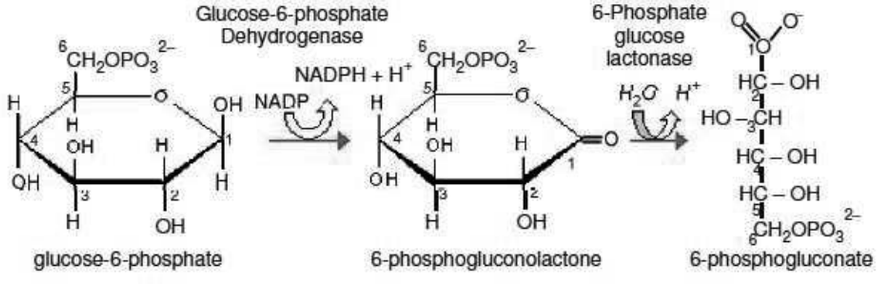


Fig. 2.42 *Structures of Glucose-6-Phosphate, 6-Phosphogluconolactone and 6-Osphogluconate*

NAD+ and NADP+ differ only in the presence of an extra phosphate on the adenosine ribose of NADP+. This difference has little to do with redox activity, but is recognized by substrate-binding sites of enzymes. It is a mechanism for separation of catabolic and synthetic pathways.

NADPH, a product of the Pentose Phosphate Pathway, performs the functions of a reductant in various synthetic (anabolic) pathways, including fatty acid synthesis. NAD+ serves as electron acceptor in catabolic pathways in which metabolites are oxidized. The resultant NADH is re-oxidized by the respiratory chain, producing ATP.

NOTES

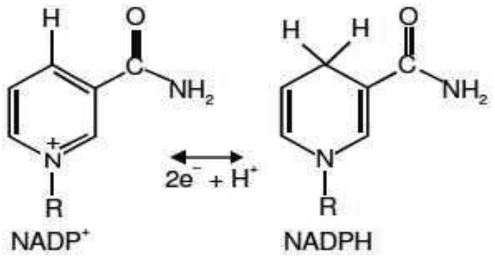


Fig. 2.43 *Structures of NAD+ and NADP+*

The pentose phosphate pathway occurs in the cytosol, and the reactions are outlined. NADPH is a product of two dehydrogenases, glucose-6-phosphate dehydrogenase and 6-phosphogluconate dehydrogenase, which convert glucose 6- phosphate to the pentose, ribulose 5-phosphate. Ribulose 5-phosphate undergoes isomerization to both xylulose 5-phosphate and ribose 5-phosphate, and these two compounds serve as substrates for transketolasc to yield sedoheptulose 7-phosphate and glyceraldehyde 3-phosphate.Transaldolase then converts these to fructose 6-phosphate and erythrose 4-phosphate. Transketolase also catalyses a reaction between xvlulose 5- phosphate and erythrose 4-jphosprute to form glyceraldehyde 3-phosphate and fructose 6-phosphate.

Glucose-6-phosphate Dehydrogenase is the committed step of the Pentose Phosphate Pathway. This enzyme is regulated by availability of the substrate NADP+. As NADPH is utilized in reductive synthetic pathways, the increasing concentration of NADP+ stimulates the Pentose Phosphate Pathway, to replenish NADPH. The remainder of the Pentose Phosphate Pathway accomplishes conversion of the 5-C ribulose-5-phosphate to the 5-C product ribose-5-phosphate, or to the 3-C glyceraldehyde-3-phosphate and the 6-C fructose-6-phosphate. Additional enzymes include Ribulose-5-phosphate Epimerase, Ribulose-5- phosphate Isomerase, Transketolase, and Transaldolase. Epimerase interconverts the stereoisomers ribulose-5-phosphate and xylulose-5-phosphate. Isomerase converts the ketose ribulose-5-phosphate to the aldose ribose-5-phosphate. Both reactions involve deprotonation to form an endiolate intermediate, followed by specific reprotonation to yield the product. Both reactions are reversible.

Transketolase and transaldolase catalyze transfer of 2-C and 3-C molecular fragments respectively, in each case from a ketose donor to an aldose acceptor. D. E. Nicholson has suggested that the names of these enzymes should be changed, since Transketolase actually transfers an aldol moiety (glycoaldehyde) and Transaldolase actually transfers a ketol moiety (dihydroxyacetone). Thiamine pyrophosphate binds at the active sites of enzymes in a "V" conformation. The amino group of the aminopyrimidine moiety is close to the dissociable proton, and serves as the proton acceptor. This proton transfer is promoted by a glutamate residue adjacent to the pyrimidine ring.

The thiazolium carbanion (ylid) that results from proton dissociation reacts with the carbonyl C of xylulose-5-P to form an addition compound. The positively charged N in the thiazole ring acts as an electron sink, promoting C-C bond

NOTES

cleavage. The 3-C aldose glyceraldehyde-3-phosphate is released. A 2-C fragment remains on TPP. Completion of the reaction is by reversal of these steps. The 2-C fragment condenses with one of the aldoses erythrose-4-phosphate (4-C) or ribose-5-phosphate (5-C) to form a 6-C or 7-C ketose-phosphate product. Transfer of the 2-C fragment to the 5-C aldose ribose-5-phosphate yields sedoheptulose-7-phosphate (see above). Transfer instead to the 4-C aldose erythrose- 4-phosphate yields fructose-6-phosphate. Explore at right the structure of Transketolase with bound substrate erythrose-4-phosphate. Transaldolase catalyzes transfer of a 3-C dihydroxyacetone moiety, from sedoheptulose-7-phosphate to glyceraldehyde-3- phosphate. The e-amino group of an active site lysine residue reacts with the carbonyl C of sedoheptulose-7-phosphate to form a protonated Schiff base intermediate. Aldol cleavage results in release of erythrose-4-phosphate. The Schiff base stabilizes the carbanion on C3.

Completion of the reaction occurs by reversal, as the carbanion attacks instead the aldehyde carbon of the 3-carbon aldose glyceraldehyde-3-phosphate to yield the 6- carbon fructose-6-phosphate. Depending on relative needs of a cell for ribose-5-phosphate, NADPH, and ATP, the Pentose Phosphate Pathway can operate in various modes, to maximize different products. There are three major scenarios as shown in Figure 2.44.

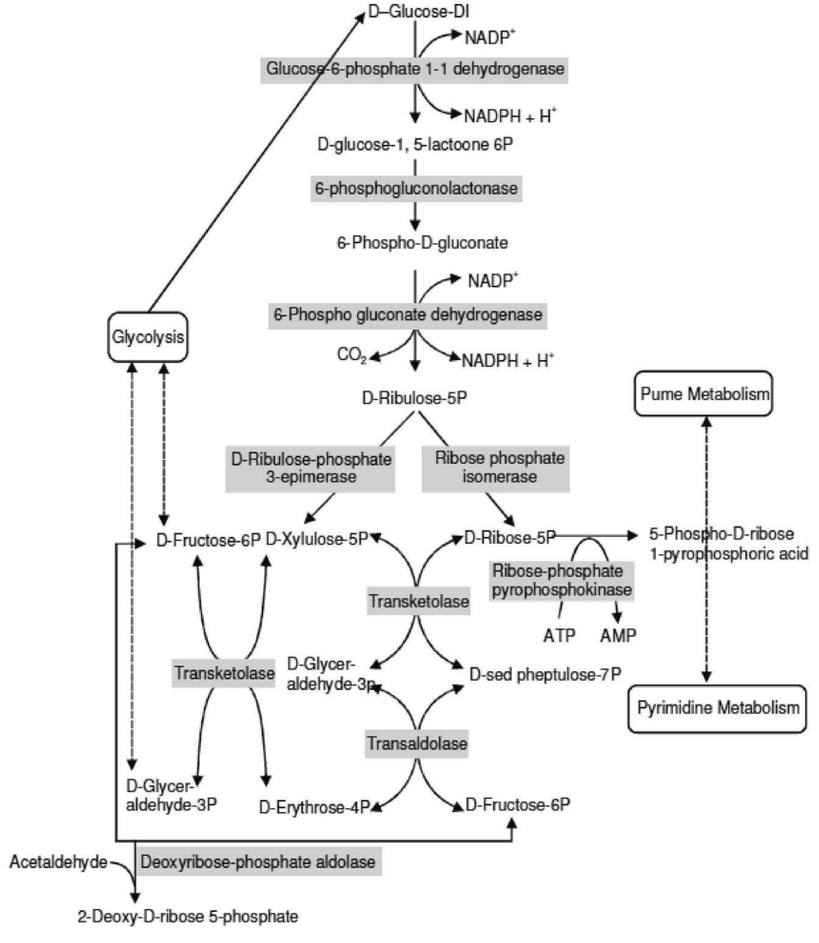


Fig. 2.44 *Different Scenarios in Pentose Phosphate Pathway Operation*

NOTES

Primary Functions of HMP Shunt

The primary functions of the HMP Shunt are as follows:

- To generate reducing equivalents, in the form of NADPH, for reductive biosynthesis reactions within cells.
- To provide the cell with Ribose-5-Phosphate (R5P) for the synthesis of the nucleotides and nucleic acids.
- Although not a significant function of the PPP, it can operate to metabolize dietary pentose sugars derived from the digestion of nucleic acids as well as to rearrange the carbon skeletons of dietary carbohydrates into glycolytic/ Gluconeogenic intermediates.
- Located exclusively in the cytoplasm, the pathway is one of the three main ways the body creates molecules with reducing power, accounting for approximately 60 per cent of NADPH production in humans.

Check Your Progress

11. What is the immediate source of blood glucose in liver?
12. What is gluconeogenesis?
13. What is glycogenesis?
14. Which are the most important sites of glycogenesis?
15. How many reactions are involved in the conversion of glucose to lactic acid?
16. What is the energy difference between ADP and ATP?

2.9 ANSWERS TO 'CHECK YOUR PROGRESS'

1. Carbohydrates is an organic compound which consists of the elements carbon (C), hydrogen (H) and oxygen (O) with a ratio of hydrogen twice that of carbon and oxygen, i.e., 1:2:1.
2. Monosaccharide or simple sugars are those carbohydrates that cannot be hydrolyzed into simpler carbohydrate containing a single polyhydroxy aldehyde or ketone unit.
3. On the basis of oxidizing activity, monosaccharides are categorized in two classes. Monosaccharides (such as glucose) which are capable of reducing ferric or cupric ion are known as reducing sugars.
4. The general formula of disaccharides is $C_n(H_2O)_{n-1}$.
5. Polysaccharides are condensation yield of more than 20 or more so monosaccharide unit and few may have hundreds or thousands of units. They are also known as Glycans.
6. Cellulose is water-insoluble, fibrous, tough substance found in mainly in the cell walls of plants and predominantly in stalks, stems, trunks, and all the woody portions of the plant body.

7. Humans have 60 blood group antigens which are further subdivided into fourteen independent genetic systems.
8. The simplest unit of glycosylation present on a molecular level is O-linked N-acetylglucosamine (O-GlcNAc) resulted from modification of hundreds, and most likely thousands, of nuclear and cytosolic proteins.
9. The empirical formula of ascorbic acid is $C_6H_8O_6$.
10. There are various functions of vitamin C like:
    a. **Health Benefits:** The damage caused to the body by smoking, radiations can be recovered by it due to its antioxidant properties. It helps in fixing worn-out and damaged tissues.
    b. **Personal Care:** The antioxidant property of vitamin C make it a useful component in cosmetics and personal care products like makeup products. It is also found in skin and hair care products.
11. The immediate source of blood glucose is the store of glycogen in the liver.
12. Glycogen is the immediate source of liver output of glucose, the liver can rapidly switch to producing glucose by de novo synthesis from ammo acids that are released from muscle and metabolized in liver a process known as gluconeogenesis.
13. Glycogenesis is the synthesis of glycogen from glucose.
14. Liver and muscles are the most important sites of glycogenesis.
15. Ten reactions are involved in the conversion of glucose to lactic acid.
16. The energy difference between ADP and ATP is approximately 7 kcal per mole.

2.10 SUMMARY

- Carbohydrates have important structural and metabolic roles. In plants, glucose is synthesized from carbon dioxide and water by photosynthesis and stored as starch or used to synthesize cellulose for plant framework.
- The different types of carbohydrates are monosaccharides, oligosaccharides, polysaccharides and glycoconjugates.
- Glycocalyx is the major component of mammalian cell surface. It is composed of complex layer of carbohydrates which determines the interaction of cells with its surroundings.
- Specific surface glycans, or the united cellular complement of these molecules, reflect the internal workings and status of a cell and thus serve as available biomarkers for, amongst other conditions, cancer and stem cell status.
- Monosaccharides or simple sugars are those carbohydrates that cannot be hydrolysed into simpler carbohydrates containing a single polyhydroxy aldehyde or ketone unit.
- Oligosaccharides are compound sugars that yield two to ten molecules of the identical or different monosaccharides on hydrolysis (removal of water molecules).

NOTES

- Polysaccharides are condensation yield of more than twenty monosaccharide units and few may have hundreds or thousands of units.
- The main storage polysaccharide of animal cells is glycogen. Glycogen is a polymer of subunits of glucose with glycosidic linkage at the linear site and glycosidic linkage at the branching site.
- Peptidoglycan is the most inflexible layer of the bacterial cell envelope. It is partly made up of a heteropolysaccharide made from two corresponding monosaccharide units that are heteropolymer of corresponding 1-4-linked N-acetyl-glucosamine (NAG) and N-acetylmuramic acid (NAM) residues.
- Blood group substances comprises of the macromolecules containing highly specific, serologically active antigenic sites responsible for their immunological individuality.
- The carbohydrate moiety is suggested to be related to the serological activity in the early stages of blood group investigation.
- Glycoconjugate is the biologically active molecule which constitutes the informational carbohydrate covalently linked to a protein or a lipid. Glycoconjugates include proteoglycans, glycoprotein, and glycolipids.
- Carbohydrates possess active groups which are responsible for their chemical behaviour. These groups are glycosidic (OH), alcoholic (OH) and aldehyde (CHO) or ketonic (CO).
- Ascorbic acid is also known as Vitamin C. It is not produced or stored by the human body and needs to be supplemented from diet as it plays various important roles in humans.
- Vitamin C is an important nutrient of certain animals including humans. The term vitamin C includes several vitamers that have vitamin C activity in animals.
- Carbohydrate metabolism begins with digestion in the small intestine where monosaccharides are absorbed into the blood stream. Blood sugar concentrations are controlled by three hormones: insulin, glucagon and epinephrine.
- If the concentration of glucose in the blood is too high, insulin is secreted by the pancreas. Insulin stimulates the transfer of glucose into the cells, especially in the liver and muscles, although other organs are also able to metabolize glucose.
- In the liver and muscles, most of the glucose is changed into glycogen by the process of glycogenesis (anabolism). Glycogen is stored in the liver and muscles until needed at some later time when glucose levels are low.
- If blood glucose levels are low, then epinephrine and glucagon hormones are secreted to stimulate the conversion of glycogen to glucose. This process is called glycogenolysis (catabolism).
- If glucose is needed immediately upon entering the cells to supply energy, it begins the metabolic process called glycolysis (catabolism). The end products of glycolysis are pyruvic acid and ATP.

- Since glycolysis releases relatively little ATP, further reactions continue to convert pyruvic acid to acetyl CoA and then citric acid in the citric acid cycle.
- The majority of the ATP is made from oxidations in the citric acid cycle in connection with the electron transport chain.
- During strenuous muscular activity, pyruvic acid is converted into lactic acid rather than acetyl CoA. During the resting period, the lactic acid is converted back to pyruvic acid.
- The pyruvic acid in turn is converted back to glucose by the process called gluconeogenesis (anabolism).
- If the glucose is not needed at that moment, it is converted into glycogen by glycogenesis.

2.11 KEY TERMS

- **Carbohydrates:** These are polyhydroxy compounds that contain a carbonyl (C=O) group of the general formula, (CH2O)n.
- **Monosaccharide:** Also known as simple sugars, these are those carbohydrates that cannot be hydrolysed into simpler carbohydrates containing a single polyhydroxy aldehyde or ketone unit.
- **Disaccharide:** It is the sugar formed when two monosaccharides are joined by glycosidic linkage.
- **Polysaccharides**: They are long chain polymeric carbohydrates made of monosaccharide units that are bound together by glycosidic linkages.
- **Glycogenesis:** It is the process of glycogen synthesis. In this process, addition of glucose molecules to chains of glycogen takes place for storage.
- **Glycogenolysis: It is the process that in which** glycogen molecules break down into glucose.
- **Starch:** It is a homopolymer of glucose forming an -glycosidic chain called glucosan or glucan.
- **Cellulose:** It is a water-insoluble, fibrous, tough substance found mainly in the cell walls of plants and primarily in stalks, stems, trunks and every wooded part of the plant body.

2.12 SELF ASSESSMENT QUESTIONS AND EXERCISES

Short-Answer Questions

1. What are the functions of carbohydrates?
2. What are oligosaccharides? What are its types?
3. Briefly explain the storage and structural polysaccharides.

4. Write a short note on glycoproteins and glycolipids.
5. What is the composition of blood group substances?
6. Give a brief account of ascorbic acid.

NOTES

Long-Answer Questions

1. Explain the role of sugars in biological recognition.
2. Discuss the structure and types of monosaccharides.
3. Evaluate the concept of mucopolysaccharides.
4. Discuss the process of glycolysis in detail.
5. Describe the different steps of reaction in the TCA cycle.
6. Discuss how the pentose-phosphate pathway works.

2.13 FURTHER READING

Blackstock, James C. 2014. *Guide to Biochemistry*. Oxford: Butterworth-Heinemann.

Park, K. H. 2008. *Carbohydrate-Active Enzymes: Structure, Function and Applications*. Amsterdam: Elsevier.

Bhardwaj, Uma. 2012. *Biochemistry for Nurses*. Noida: Pearson Education India.

Fromm, Herbert J. and Mark Hargrove. 2012. *Essentials of Biochemistry*. Berlin: Springer.

Fearon, William Robert. 2014. *An Introduction to Biochemistry*. Amsterdam: Elsevier.

UNIT 3 LIPIDS

NOTES

Structure

3.0 INTRODUCTION

Lipids perform a variety of functions in the human body and there has been a lot of research conducted on them in the recent years because of the increasing importance assigned to them by the media and the medical world. Lipids refer to the class of organic molecules, which include fats, waxes and steroids. They, together with proteins and carbohydrates, are a vital component of the living cells. They act as reservoirs of fuel and regulators of body heat in cold climates and help in the sustenance of human beings. In this unit, we will discuss the meaning, structure and nomenclature of various lipids, along with their classification into many broad types and sub-types, and their physical and chemical properties. It will also focus on the lipid aggregates and metabolism.

3.1 OBJECTIVES

After going through this unit, you will be able to:

- Describe the meaning, structure and nomenclature of various lipids
- Explain the classification of lipids and their physical and chemical properties
- Discuss the lipid aggregates and metabolism

NOTES

3.2 LIPIDS: STRUCTURE AND FUNCTIONS

Lipids are a heterogeneous group of compounds, including fats, oils, sterols, waxes and related compounds which are associated more by their physical and by their chemical properties. These are oily or greasy organic substances, moderately insoluble in water and significantly soluble in organic solvents like ether, chloroform and benzene. They are, thus, hydrophobic in nature. They are also known as lipins or lipoids. The concluding term is, however, occasionally used to refer to 'fat-like' substances which may not essentially be related to the fatty acids. The term 'lipid' was first used by the German biochemist Bloor in 1943 for a chief class of tissue components and food. The fats and oils used generally as stored forms of energy in living organisms are derivatives of fatty acids. The general structure of lipids of is shown in Figure 3.1.

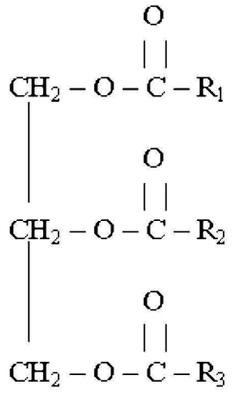


Fig. 3.1 *General Structure of Lipid where R Denotes the Alkyl Group*

3.2.1 Fatty Acids and Essential Fatty Acids (Saturated and Unsaturated Prostaglandins)

In many biological organisms, fats and oils are the major stored forms of energy. Phospholipids and sterols constitute the chief structural elements of biological membranes. The other lipids which play many critical roles are enzyme cofactors, electron carriers, light absorbing pigments, hydrophobic anchors for proteins, which help in membrane proteins folding, emulsifying agents in the digestive tract, hormones and intracellular messengers. Lipids are important constituents of the diet because of their high energy value and also because they are precursors of the fat-soluble vitamins. In the body, the fats serve as an ingenious source of energy which is stored in the adipose tissues. They also serve as an insulating material in the subcutaneous tissues and around certain organs. Fats combine with proteins to form lipoproteins which are important constituents of the cell membranes and mitochondria of the cell. Saturated and unsaturated fatty acids are commonly found in both plant and animal world. Plant and animal fats differ in variety of ways but some of the basic differences are listed in Table 3.1.

Table 3.1 *Significant Differences between Animal and Plant Fats*

Animal fats	Plant fats
1. Relatively rich in saturated fatty acids esp., C16 and C18 acids.	1. Relatively rich in unsaturated fatty acids, esp., polyunsaturated acids.
2. Solid at ordinary room temperature	2. Liquid at ordinary room temperature
3. These have usually low iodine number.	3. These have usually high iodine number.
4. These have usually high Reichert-Meissl Number.	4. These have usually low Reichert-Miessl number.
5. These are stored mainly in liver and bone marrow.	5. These are stored mainly in seeds and fruits.
6. Oxidative rancidity is observed more frequently.	6. Oxidative rancidity is observed less frequently.
7. Examples– Butterfat, Beef fat, Tallow	7. Examples– Olive oil, Castor oil, Soybean oil, Corn oil

NOTES

Proteins, polysaccharides, DNA and RNA are macromolecules but lipids are not usually considered as macromolecules even though they share some of their features. Most of the lipids are synthesized as linear polymers of a smaller molecule, such as in case of acetyl-CoA. It is notable that most of the mass of both mammalian and bacterial cells are comprised of water and protein. The cellular oxidation reaction of fatty acids to carbon dioxide (CO_2) and water (H_2O) is highly exergonic, which shows resemblances to the internal combustion engines in which controlled and rapid burning of fossil fuels takes place.

Fatty acids are long-chain organic acids which generally have carbon atoms from 4 to 30 and they have a single carboxyl group with a long, nonpolar hydrocarbon 'tail', which facilitates most of the lipids to their hydrophobicity and oily or greasy nature. Fatty acids do not occur in free or uncombined state in cells or tissues but are present in covalently bound state in different classes of lipids. Fatty acids which are found in natural fats are called as monocarboxylic acids and contain an even number of C atoms as these are synthesized from two carbon units which are usually straight-chain derivatives. The chemical names and sources of common fatty acids are listed in Table 3.2.The chain may be saturated which means containing single bonds only or unsaturated which means containing one or more double bonds. Some fatty acids may have hydroxyl group(s) in the chain designated as hydroxy or oxygenated fatty acids and still others may possess ring structure referred to as cyclic fatty acids. Fatty acids are stored as an energy reserve in the form of fat linkage through an ester joined to glycerol to form triglycerides.

NOTES

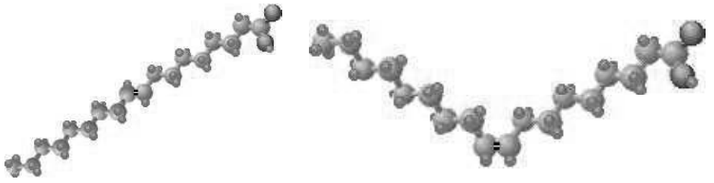
Table 3.2 *List of Common Fatty Acids*

Chemical Names and Descriptions of some Common Fatty Acids				
Common Name	**Scientific Name**	**Carbon Atoms**	**Double Bonds**	**Sources**
Butyric acid	Butanoic acid	4	0	Butter fat
Caproic acid	Hexanoic acid	6	0	-do-
Caprylic acid	Octanoic acid	8	0	Coconut oil
Capric acid	Decanoic acid	10	0	-do-
Lauric acid	Dodecanoic acid	12	0	-do-
Myristic acid	Tetradecanoic acid	14	0	Palm kernel oil
Palmitic acid	Hexadecanoic acid	16	0	Palm oil
Palmitoleic acid	9-hexadecenoic acid	16	1	Animal fats
Stearic acid	Octadecanoic acid	18	0	-do-
Oleic acid	9-octadecenoic acid	18	1	Olive oil
Ricinoleic acid	12-hydroxy-9-octadecenoic acid	18	1	Castor oil
Vaccenic acid	11-octadecenoic acid	18	1	Butterfat
Linoleic acid	9,12-octadecadienoic acid	18	2	Grape seed oil
Alpha-linolenic acid (ala)	9,12,15-octadecatrienoic acid	18	3	Flaxseed (linseed) oil
Gamma-linolenic acid (gla)	6,9,12-octadecatrienoic acid	18	3	Borage oil
Arachidic acid	Eicosanoic acid	20	0	Peanut oil, fish oil
Gadoleic acid	9-eicosenoic acid	20	1	Fish oil
Arachidonic acid (aa)	5,8,11,14-eicosatetraenoic acid	20	4	Liver fats
Epa	5,8,11,14,17-eicosapentaenoic acid	20	5	Fish oil
Behenic acid	Docosanoic acid	22	0	Rapeseed oil
Erucic acid	13-docosenoic acid	22	1	-do-
Dha	4,7,10,13,16,19-docosahexaenoic acid	22	6	Fish oil
Lignoceric acid	Tetracosanoic acid	24	0	Small amounts in most fats

The structured nomenclature of the fatty acids is based on the Genevan system. According to this system, the fatty acid is named after the hydrocarbon by adding the suffix 'oic' acid as written in place of the final letter e in the name of the hydrocarbon with the same number of carbon atoms. The names of saturated

NOTES

fatty acids end with the suffix 'anoic' acid and those of unsaturated fatty acids end with the suffix 'enoic' acid. The arrangement of carbon atoms in the fatty acid chain is denoted either by numbering in which the carboxyl carbon is numbered as C1, the carbon adjacent to C1 is numbered as C2 and so on or by the use of Greek letters in which case C2 is denoted as α-carbon, C3 is denoted as â-carbon and so on. Thus, a broadly used convention to specify the number and position of the double bond(s) in case of unsaturated fatty acids is to write the number of carbon atoms, the number of double bond(s) and the position of the double bonds(s) below the name of the acid. For example, oleic acid having 18 carbon atoms and a double bond between carbon atoms 9 and 10 is written as 18:1; 9. Similarly, linoleic acid (18 carbon atoms and 2 double bonds at C 9 and C 12) is written as 18:2; 9, 12. An alternative method to write the name of an unsaturated fatty acid is to first write the position of double bond(s) in numerals and then the total number of carbon atoms in Roman followed by the suffix 'enoic' acid. Thus, oleic acid may be written as 9-octadecenoic acid and linoleic acid as 9, 12-octadecadienoic acid. The structure of cis and trans-form of 9-octadecadienoic acid is shown in Figure 3.2.



Trans-*9-octadecenoic acid (Elaidic acid)* **Cis**-*9-octadecenoic acid (oleic acid)*

Fig. 3.2 *Structure of Cis and Trans- Form of 9-Octadecenoic Acid*

A description of the various categories of fatty acids involved in lipid formation is as follows

Saturated Fatty Acids

The fatty acids whose carbon chain cannot absorb additional hydrogen atoms are termed as saturated fatty acids. The general formula for saturated fatty acids is $C_nH_{2n+1}COOH$. Some even-numbered straight chain saturated fatty acids are found in both plant and animal worlds. In addition, lipids from all sources contain small amounts of saturated fatty acids with an odd number of carbon atoms (C 5 through C 17). Generally, these odd-carbon acids account for less than 1 per cent of the total fatty acids. In animal fats, palmitic and stearic acids (C16 and C18) are the most abundant saturated fatty acids, next in order are shorter chain fatty acids (C14 and C12) and longer chain fatty acids (C20, C22 and C24). Fatty acids of 10 carbon atoms or less are present in limited amounts in animal lipids except milk fat which contains considerable amount of lower molecular weight fatty acids. The high proportion of these acids may, thus, be shown in the descending order as follows:

C16, C18 > C14, C12, C20, C22, C24 > C10 and less

Unsaturated Fatty Acids

The fatty acid whose carbon chain can absorb additional hydrogen atoms are termed as unsaturated fatty acids. Any class of aliphatic monocarboxylic acids

NOTES

that form part of a lipid molecule and can be derived from fat by hydrolysis are termed as unsaturated fatty acids. Unsaturated fatty acids are simple molecules built from around a series of carbon atoms linked together in a chain. The structure of saturated and unsaturated fatty acids is shown in Figure 3.3.

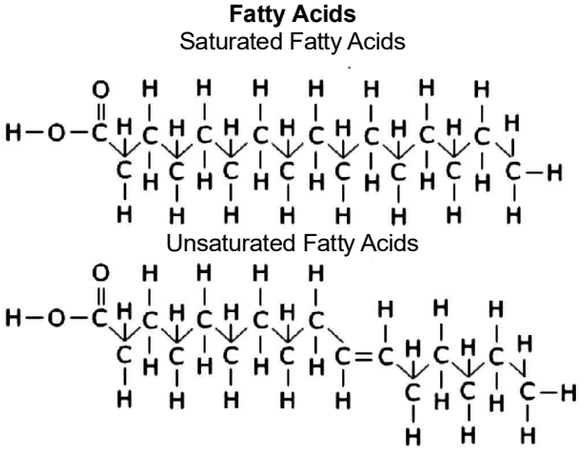


Fig. 3.3 *Structure of Saturated and Unsaturated Fatty Acid*

Unsaturated fatty acids may be categorized on the basis of the degree of unsaturation as described in the following lines:

- **Monoethenoid Acids —** These contain one double bond and conform to the general formula, CnH_{2n}–1COOH. The common example is oleic acid.
- **Diethenoid Acids —** These contain two double bonds and have general formula, CnH_{2n}–3COOH.The common example is linoleic acid.
- **Triethenoid Acids —** These contain three double bonds and have general formula ,CnH_{2n}–5COOH. The common example is linolenic acid.
- **Tetraethenoid Acids —** These contain four double bonds and have general formula, CnH_{2n}–7COOH. The common example is arachidonic acid

Monoethenoid acids are usually known as Monounsaturated Fatty Acids (MUFAs) and the residual ones are known as Polyunsaturated Fatty Acids (PUFAs). In most of the unsaturated fatty acids, there is a double bond (designated Δ9) between carbon atoms 9 and 10. This is predominantly accurate for the unsaturated fatty acids which are commonly found in the plant world. If there are additional bonds, they usually occur between Δ9 and the methyl-terminal end of the chain. In case of mammals, polyunsaturated fatty acids can have up to 22 carbon atoms and 6 double bonds but in plants these acids do not exceed 18 carbon atoms and 4 double bonds. Human body can convert stearic acid to oleic acid by inserting a double bond but is unable of inserting additional double bonds because of which oleic acid cannot be converted to linoleic, linolenic or arachidonic acid. Anyone of these acids is needed for normal functioning of the cell and especially in skin tissues. Since they cannot be synthesized by the cells, they must be obtained from the diet. These three acids are collectively called as Essential

Fatty Acids (EFA) on account of the important physiological role. The term 'Essential Fatty Acids' (EFA) was introduced by Burr and Burr in 1930.

NOTES

Vegetable oils contain two types of Polyunsaturated Fatty Acids (PUFAs) – linoleic acid (lin with 2 double bonds) and á-linolenic acid (len with 3 double bonds). In chemical terminology, Lin and longer chain fatty acids derived from it are referred to as n-6 fatty acids since the first double bond in their molecule occurs on carbon no. 3. Len and fatty acids derived from it are referred to as n-3 fatty acids. Since the first double bond in their molecule occurs on carbon no.3, Lin is more abundant in nature and is present in all vegetable oils, whereas len is present only in some vegetable oils such as mustard oil, rapeseed and soyabean. Fish oils are however, good sources of both lin and len fatty acids and hence are considered nutritionally rich. Hydrogenated fat or 'vanaspati' is manufactured from vegetable oils by the process of hydrogenation. It is a saturated fat and, though derived from vegetable oils, behaves like other saturated fats of animal origin. All foods contain small quantities of bound fat or invisible fat. Such bound fat is found in green leafy vegetables and pulses in plenty and is a rich source of n-3 fatty acids.

About half of the daily human necessity of fat (near about 40 gm) can be derived through such bound fat. Thus, the minimum daily requirement of visible fat or oil etc is only 20 gm. For long time, these two types of PUFAs, n-6 and n-3, were thought to have similar effects but latest study shows that n-6 and n-3 PUFAs have diverse types of effects on blood lipids and blood clotting and that they should be present in the diet in a certain proportion. Too much of n-6 with very little n-3 is not advantageous. For healthy cooking practice, it is therefore valuable and beneficial to use vegetable oils containing both lin and len, instead of using anyone type of oil, including even the much-exposed unsaturated oils like safflower or rice bran which are very effective in reducing blood cholesterol.

Physical Properties of Fatty Acids

The physical properties of fatty acids are as follows:

1. **State:** Fats containing saturated fatty acids are solid at ordinary room temperature. The animal fats belong to this category. Most plant fats, on the contrary, possess unsaturated fatty acids and are, henceforth, liquid at room temperature.
2. **Colour, Odour and Taste:** When pure, the fats are colourless, virtually odourless and possess an extremely bland taste. They are capable of absorbing a variety of odours and hence flavour during storage. In some cases, however, this absorbing property of fats is of advantage. For example, the perfumes of some flowers can be isolated by placing their petals in contact with the fat for a certain period, then extracting the fat with alcohol and concentrating the essence.
3. **Solubility:** The fats are, however, only sparingly soluble in water. These are, therefore, described as hydrophobic in contrast to the water-soluble or hydrophilic substances like many carbohydrates and proteins. However, these are freely soluble in organic solvents like chloroform, ether, acetone and benzene. These solvents, as they dissolve fats in them, are also known

NOTES

as 'fat solvents'. The solubility of the fatty acids in organic solvents, in fact, decreases with the increase of chain length. The introduction of hydroxyl groups, however, increases solubility.

4. **Melting Point:** The melting point of fats depends on the chain length of the constituent fatty acid and the degree of unsaturation. Fats containing saturated fatty acids from C 4 to C 8 are liquid at room temperature but those containing C 10 or higher saturated fatty acids are solid and their melting points increase with increasing chain length. With the introduction of double bond in the fat molecule, the melting point lowers considerably. It may be stated, in general, that greater the degree of unsaturation (or higher the number of double bonds) of the constituent fatty acid, the lower is the melting point of the fat. In fact, short chain length and unsaturation enhance the fluidity of fatty acids and of their derivatives.
5. **Specific Gravity:** The specific gravity of the fats is less than 1 (about 0.86) and, therefore, they float on water surface. Solid fats are lighter than the liquid fats. Oils spread on water to form thin monomolecular layers. In general, either unsaturation of the fatty acid chains increase or increase in chain length of the fatty acid residues tend to increase the specific gravity.
6. **Geometric Isomerism:** As stated earlier, the presence of double bond (s) in the unsaturated fatty acid part of the fat molecule produces geometric (or cis-trans) isomerism.
7. **Insulation:** The fats possess high insulating power, i.e., they are bad conductor of heat. A layer of fat below the skin provides a sort of blanket for warm-blooded animals (or homoiotherms). This is especially important for whales and seals which have to maintain a high temperature in cold waters. The fishes are cold-blooded animals (or poikilotherms) and, therefore, do not require maintenance of high temperature and so have very little subcutaneous fat.
8. **Emulsification:** It is the process by which a lipid mass is converted into a number of small lipid droplets. The fats may be emulsified by shaking either with water or with emulsifying agents like soaps, gums, proteins etc. An emulsifying agent helps in the production of a finely divided suspension of a fat in an aqueous medium. The hydrocarbon portions of the two (the emulsifier and the fat) tend to aggregate. This leaves the water-soluble group of the emulsifier projecting into the aqueous phase. A fat droplet will associate with a number of molecules of the emulsifier, thus producing a new water soluble surface. Water molecules, henceforth, tend to be held in a layer or 'cloud' around each droplet, thus disallowing the aggregation of the fat droplets.The process of emulsification is of great metabolic significance. In fact, the fats have to be emulsified before they can be absorbed by the intestinal wall. The process is accomplished by the bile juice secreted from liver.
9. **Surface Tension:** The force with which the surface molecules are held together is called the surface tension. When liquid fat is poured on water, it spreads uniformly over the surface of water in the form of a unimolecular layer and thus reduces the surface tension of water.

Chemical Reactions of Fatty Acids

The chemical reactions of the fats reflect the reactivities of the ester linkage and the degree of unsaturation of the hydrocarbon chain. The various chemical reaction of fatty acids have been discussed in this section.

NOTES

Reactions Involving -COOH Group

1. **Hydrolysis:** The fats are hydrolyzed by the enzymes lipases to yield fatty acids and glycerol. The lipases catalyze this reaction at a slightly alkaline pH (7.5–8.5) in a stepwise manner. The fats first split to produce diglycerides, part of these are then split to monoglycerides. Finally, part of the monoglycerides split to yield fatty acid and glycerol. In the intestine, the absorption of mono-, diand triglycerides is so rapid that very little free glycerol is formed.

   Simple lipid + H_2O $\xrightarrow{\text{hydrolysis}}$ fatty acid + alcohol

2. **Saponification:** The hydrolysis of fats by alkali is called saponification. This reaction results in the formation of glycerol and salts of fatty acids which are called soaps (Figure 3.4). The soaps are of two types: hard and soft. Hard soaps such as the common bar soaps are the sodium salts of the higher fatty acids. Soft soaps are the potassium salts of higher fatty acids and are marketed as semisolids or pastes. The fatty acid salts of calcium, magnesium, zinc and lead are, however, insoluble in water. Calcium soaps are used industrially as lubricating greases. Zinc soaps are employed in the manufacture of talcum powder and other cosmetics. Lead and magnesium soaps are used in paints industry to hasten the process of drying Soaps are important cleansing agents. Their cleansing property is due to their emulsifying action (i.e., capacity to render more prolonged the mixing of oil and water). This is accomplished by means of negative charge the soap anion confers on oil droplets. The electrostatic repulsion then prevents the coalescence of soap and oil droplets into an oil phase.

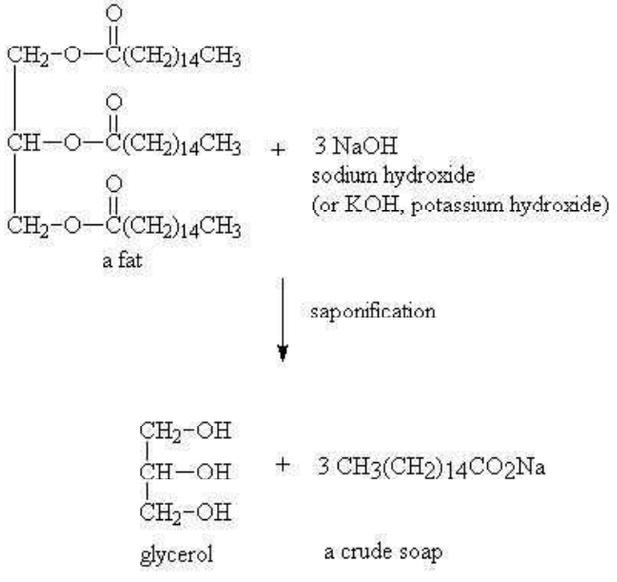


Fig. 3.4 *Process of Saponification*

NOTES

3. **Hydrolytic Rancidity:** When butter or other fats are stored, they often become rancid and hence unpalatable. Rancidity is caused by the growth of microorganisms which secrete enzymes like lipases (Figure 3.5). These split the fats into glycerol and free fatty acids. The fatty acids impart unpleasant odour and flavour to the fat. However, butter may be prevented from becoming rancid by refrigeration or by exclusion of water.

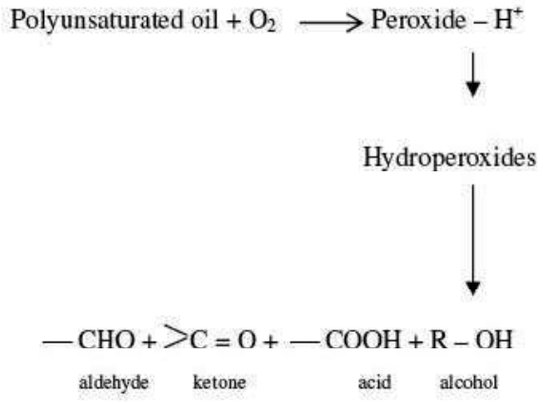


Fig. 3.5 *Process of Hydrolytic Rancidity*

Reactions Involving Double Bond

Reactions involving double bond in include:

1. Hydrogenation

Unsaturated fatty acids, either free or combined in lipids, react with gaseous hydrogen to yield the saturated fatty acids. The reaction is catalyzed by platinum, palladium or nickel. The addition of hydrogen takes place at the C—C double bond (s). Thus, 1 mole of oleic, linoleic or linolenic acid reacts with 1, 2 or 3 moles of hydrogen respectively to form stearic acid. This reaction is of great commercial importance since it permits transformation of inexpensive and unsaturated liquid vegetable fats into solid fats. The latter are used in the manufacture of candles, vegetable shortenings and of oleomargarine.

2. Halogenation

Unsaturated fatty acids and their esters can take up halogens like Br2 and I2 at their double bond (s) at room temperature in acetic acid or methanol solution. This reaction is the basis of the 'iodine number determination'.

3. Oxidation

Unsaturated fatty acids are susceptible to oxidation at their double bonds. Oxidation may be carried with ozone or $KMnO_4$.

(a) **With Ozone:** An unstable ozonide is formed which later cleaves by water to give rise to 2 aldehydic groups

(b) **With $KMnO_4$:** Under mild conditions, the glycols are formed at the sites of double bonds. The oxidation reactions have been extensively used in establishing the position of double bond(s) in the fatty acid chain. This gives important clues regarding lipid structure.

4. Oxidative Rancidity

Oils containing highly unsaturated fatty acids are spontaneously oxidized by atmospheric oxygen at ordinary temperatures. The oxidation takes place slowly and results in the formation of short chain fatty acids (C4 to C10) and aldehydes which give a rancid taste and odour to the fats. This type of rancidity or rancidification is called 'oxidative rancidity' and is due to a reaction called 'autoxidation'. Autoxidation proceeds by a free radical mechanism in which the a- methylene group is primarily attacked. A hydrogen atom is removed from an a-methylene group. This initiates a chain of reactions leading to oxidation (Holman, 1954). Oxidative rancidity is observed more frequently in animal fats than in vegetable fats. This is due to the presence, in the vegetable oils, of natural 'antioxidants' such as tocopherols (= vitamin E), phenols, naphthols etc., which check autoxidation. Vitamin E is, therefore, some-times added to foods to prevent rancidity. Animal shortenings such as lard are nowadays protected against oxidative rancidity by the addition of synthetic antioxidants such as nordihydroguiaretic acid (NDGA), tertiary butyl hydroxy anisole (BHA) etc. Linseed oil, a plant oil used as a base for paints, is highly rich in unsaturated fatty acids. It undergoes autoxidation when exposed to air, followed by polymerization to a hard, resinous coating as it 'dries' or oxidizes. The action of antioxidants is opposed by a group of compounds present in the fats and oils. These accelerate the oxidation of the parent compound and are called pro-oxidants. Majority of these substances are formed during the processing and refining of fats. Among the noteworthy pro-oxidants are the copper, iron and nickel salts of organic acids like lactic, etc.

Reactions Involving –OH Groups

Reactions involving –OH groups include dehydration or aerolein test.

Fats, when heated in the presence of a dehydrating agent, NaHSO4 or KHSO4 produce an unsaturated aldehyde called acrolein from the glycerol moiety. Acrolein (Figure 3.6) is easily recognized by its pungent odour and thus forms the basis of the test for the presence of glycerol in fat molecule.

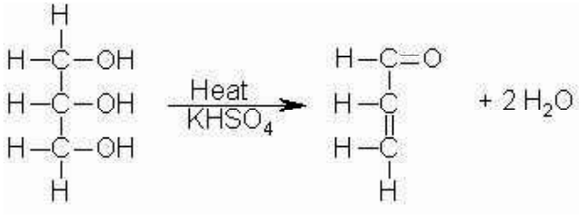


Fig. 3.6 *Acrolein Test*

Quantitative Tests

The reactions described above give valuable information about the chemical nature of fatty acids and the number of hydroxyl groups present in the fat molecule. Such chemical determinations involve various analytical tests. These are called chemical constants and include the following:

1. **Acid Value:** It is the number of milligrams of KOH required to neutralize the free fatty acids present in 1 gm of fat. The acid number, thus, tells us of

NOTES

the quantity of free fatty acid present in a fat. Obviously, a fat which has been both processed and stored properly has a very low acid number.

2. **Saponification Number:** It is the number of milligrams of KOH required to saponify 1 gm of fat. The saponification number, thus, provides information of the average chain length of the fatty acids in the fat. It varies inversely with the chain length of the fatty acids. The shorter the average chain length of the fatty acids, the higher is the saponification number.
3. **Iodine Value (or Koettstorfer Number):** It is the number of grams of iodine absorbed by 100 g of fat. The iodine number is, thus, a measure of the degree of unsaturation of the fatty acids in the fat. Oils like soybean, corn and cottonseed have higher iodine numbers (133, 127 and 109, respectively) than the solid fats such as, beef fat or tallow (42) because the former possess more unsaturated fatty acids in the fat molecule. However, the iodine number gives no indication as to the number of double bonds present in the fatty acid molecule.
4. **Polenske Number:** It is the number of millilitres of 0.1N KOH required to neutralize the insoluble fatty acids (i.e., those which are not volatile with steam distillation) obtained from 5 gm of fat.
5. **Reichert-Meissl Number:** It is the number of millilitres of 0.1N KOH required to neutralize the soluble, volatile fatty acids derived from 5 g of fat. The Reichert-Meissl number, thus, measures the quantity of short chain fatty acids (up to C 10 inclusive) in the fat molecule. The Reichert-Meissl numbers of coconut and palm oils range between 5 and 8. Butterfat is exceptional in having a high Reichert-Meissl number, ranging from 17 to 35. This high value makes possible the detection of any foreign fats which are, sometimes, adulterated in the manufacture of butter.
6. **Acetyl Number:** It is the number of milligrams of KOH required to neutralize the acetic acid obtained by saponification of 1 gm of fat after it has been acetylated (The treatment of fat or fatty acid mixture with acetic anhydride results in acetylation of all alcoholic OH groups). The acetyl number is, thus, a measure of the number of OH groups in the fat. For example, the castor oil has a high acetyl number (146) because of high content of a hydroxy acid, ricinoleic acid, in it.

3.2.2 Function of Lipids

The lipids perform a wide variety of functions which are as follows:

- **Food Material:** Lipids provide food, highly rich in calorific value. One gram lipid produces 9.3 kilocalories of heat.
- **Food Reserve:** Lipids are insoluble in aqueous solutions and hence can be stored readily in the body as a food reserve.
- **Structural Component:** Lipids are an important constituent of the cell membrane.
- **Heat Insulation:** The fats are characterized by their high insulating capacity. Great quantities of fat are deposited in the subcutaneous layers in aquatic mammals such as whale and in animals living in cold climates.

- **Fatty Acid Absorption:** Phospholipids play an important role in the absorption and transportation of fatty acids.
- **Hormone Synthesis:**The sex hormones, adrenocorticoids, cholic acids and also vitamin D are all synthesized from cholesterol, a steroidal lipid.
- **Vitamin Carriers:** Lipids act as carriers of natural fat-soluble vitamins such as vitamin A, D and E.
- **Antibiotic Agent:** Squalamine, a steroid from the blood of sharks, has been shown to be an antibiotic and antifungal agent of intense activity. This seems to explain why sharks rarely contract infections and almost never get cancer.

Check Your Progress

1. Why are lipids considered important constituents of the diet?
2. What are fatty acids?
3. Which suffix is used in the names of saturated fatty acids?
4. What is saponification?
5. Name the two best sources of carotene.
6. Which factors regulate the melting point of fats?

3.3 CLASSIFICATION OF LIPIDS

Lipids can be classified into three broad categories, which are as follows:

1. Simple Lipids
2. Conjugated Lipids
3. Derived Lipids

A detailed description of all these three types along with their sub-types is as follows:

1. Simple Lipids

- Neutral fats - Triglycerides
- Waxes

2. Conjugated Lipids

- Phospholipids which contain a phosphoric acid molecule and a fat molecule.
- Glycolipid which contains a carbohydrate and a fat molecule.

  Examples -cerebrosides, globosides, gangliosides
- Sulfolipids which contain a sulfate radical.
- Lipoprotein

3. Derived Lipids

- Fatty acids
- Glycerol
- Cholesterol and other steroids (Vit. D)
- Vitamins A, E, K

3.3.1 Simple Lipids: Structure and Function of Triacylglycerols

In the follwing sections, we will discuss the simple lipids.

NOTES

Triacylglycerols

Triacylglycerols, also referred to as triglycerides, fats or neutral fats are the simplest lipids constructed from fatty acids. Triacylglycerols possess three fatty acids each in ester linkage with a single glycerol molecule. Simple triacylglycerols are those containing the similar type of fatty acid in all three positions. They are named after the fatty acid they contain. Simple triacylglycerols, for example tristearin, tripalmitin, and triolein are of 16:0, 18:0, and 18:1, respectively. The triacylglycerols which are the most naturally occurring are mixed as they contain two or more different fatty acids. In order to name these compounds unambiguously, the name and position of each fatty acid must be specified. Triacylglycerols are nonpolar, hydrophobic molecules essentially insoluble in water because the polar hydroxyls of glycerol and the polar carboxylates of the fatty acids are bound in ester linkages. As compared to water, lipids have lower specific gravities which explains why mixtures of oil and water (for example, oil and vinegar salad dressing) have two phases and oil with the lower specific gravity floats on the aqueous phase. Figure 3.7 shows the structure of Triacylglycerols.

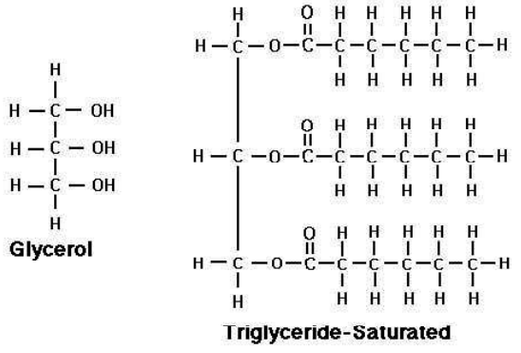


Fig. 3.7 *Structure of Triglyceride*

In most eukaryotic cells, triacylglycerols are found as a separate phase of microscopic, oily droplets in the aqueous cytosol, serving as depots of metabolic fuel. Specialized cells called as adipocytes, or fat cells, are found in vertebrates and can store large amounts of triacylglycerols as fat droplets that nearly fill the cell. Triacylglycerols are also preserved as oil in the seeds of many types of plants and provide energy and biosynthetic precursors during seed germination. Adipocytes and germinating seeds contain lipases which are enzymes that catalyze the hydrolysis of stored triacylglycerols, releasing fatty acids for export to sites where they are required as fuel. There are two significant advantages of using triacylglycerol as stored fuels, rather than polysaccharides such as glycogen and starch. The first advantage is that carbon atoms of fatty acids are more reduced than those of sugars and oxidation of triacylglycerols yields more than twice as much energy, gram for gram, as the oxidation of carbohydrates. The second advantage is that because triacylglycerols are hydrophobic and therefore unhydrated, the organism

that carries fat as fuel does not have to carry the extra weight of water of hydration that is associated with stored polysaccharide (2 g per gram of polysaccharide). Humans have fat tissue (composed primarily of adipocytes) under the skin, in the abdominal cavity, and in the mammary glands. Moderately obese people with 15 to 20 kg of triacylglycerols deposited in their adipocytes could meet their energy needs for months by drawing on their fat stores. In contrast, the human body is able to store less than a day's energy supply in the form of glycogen.

Carbohydrates like glucose and glycogen have certain advantages as quick sources of metabolic energy and one of them is their ready solubility in water. In some animals, triacylglycerols stored under the skin serve as energy stores and insulators against low temperatures. Seals, walruses, penguins, and other warm-blooded polar animals are sufficiently loaded with triacylglycerols. In hibernating animals such as bears, the huge fat reserves accumulated before hibernation serve the dual purposes of insulation and energy storage. The low density of triacylglycerols depicts another crucial function of these compounds. In sperm whales, triacylglycerols and waxes which are stored, allow the animals to match the buoyancy of their bodies to that of their surroundings during deep dives in cold water.Triacylglycerols such as tristearin which contain only saturated fatty acids are white, greasy solids at room temperature.

When lipid-rich foods are exposed for a long time to the oxygen in air, they may get damaged and become rancid. The unpleasant taste and smell associated with rancidity result from the oxidative cleavage of the double bonds in unsaturated fatty acids, which produces aldehydes and carboxylic acids of shorter chain length and therefore higher volatility.

Waxes

Biological waxes are esters of long-chain (C14 to C36) saturated and unsaturated fatty acids with long-chain (C16 to C30) alcohols. Their melting points (60 to 100 °C) are generally higher than those of triacylglycerols. In plankton, the free-floating microorganisms at the bottom of the food chain for marine animals, waxes are the chief storage form of metabolic fuel. Waxes also serve a diversity of other functions related to their water-repellent properties and their firm consistency. Certain skin glands of vertebrates secrete waxes to protect hair and skin and keep it pliable, lubricated, and waterproof. Birds, particularly waterfowl, secrete waxes from their preen glands to keep their feathers water-repellent.

The shiny leaves of holly, rhododendrons, poison ivy, and many tropical plants are coated with a thick layer of waxes, which prevents excessive evaporation of water and protects against parasites. Biological waxes find a variety of applications in the pharmaceutical, cosmetic, and other industries. Lanolin (from lamb's wool), beeswax, carnauba wax (from a Brazilian palm tree), and wax extracted from spermaceti oil from whales are widely used in the manufacture of lotions, ointments, and polishes.

3.3.2 Conjugated Lipids: Structure and Functions of Glycerophspholipids Sphingo Lipids, and Lipoproteins

NOTES

Phospholipids or Glycerophospholipids

Phospholipids are also known as phosphatids and are the most abundant membrane lipids. They serve primarily as structural components of membranes and are never stored in large quantities. As their name implies, phospholipids contain phosphorus in the form of phosphoric acid groups. They differ from triglycerides in possessing usually one hydrophilic polar 'head' group and usually two hydrophobic nonpolar 'tails' (Figure 3.8). For this reason, they are often called polar lipids. Thus, phospholipids are amphipathic, whereas the storage $-H_2O$ lipids (triglycerides and waxes) are not. In phospholipids, two of the OH groups in glycerol are linked to fatty acids while the third OH group is linked to phosphoric acid. The phosphate is further linked to one of a variety of small polar head groups (alcohols). Folch and Sperry (1955) have classified phospholipids into phosphoglycerides, phosphoinositides and phosphosphingosides.

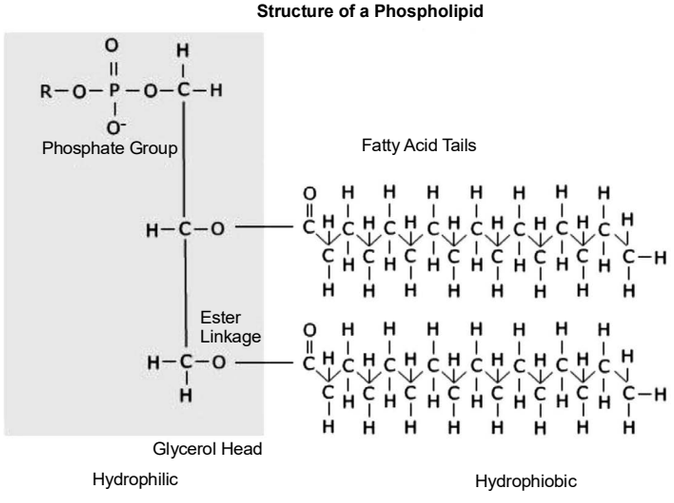


Fig. 3.8 *Structure of Phospholipid*

Phosphoglycerides

These are the major phospholipids found in membranes and contain two fatty acid molecules or 'tails' esterified to the first and second hydroxyl groups of glycerol. The third hydroxyl group of glycerol forms an ester linkage with phosphoric acid. In addition, phosphoglycerides contain a second alcohol, which is also esterified to the phosporic acid. This is referred to as 'head alcohol group' as it is present at one end ('head') of the long phosphoglyceride molecule. The various phosphoglycerides differ in their head alcohol groups. However, all of them contain two nonpolar tails, each consisting of a long chain (usually C16 or C18) fatty acid (see Figure 3.9). Usually one of the fatty acids is saturated and the other unsaturated; the latter is always esterified to the middle or β-hydroxy group of glycerol. A noteworthy feature of the phosphoglycerides is that they contain an asymmetric carbon atom at position 2 in the glycerol part of their molecule. It has the L-

configuration since it is related to L-glyceraldehyde. All phosphoglycerides have a negative charge on phosphoric group at pH 7. In addition, the head alcohol group may also have one or more electric charges at pH 7.

NOTES

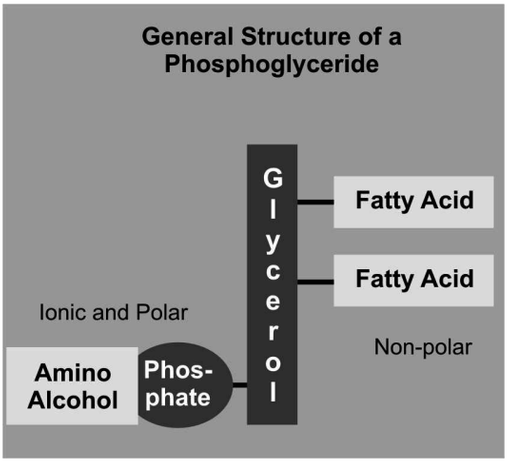


Fig. 3.9 *Structure of Phosphoglyceride*

Phosphoinositides (Phosphatidyl Inositols)

Phosphoinositides have been found to occur in phospholipids of brain tissue and of soybeans and are of considerable importance because of their role in transport processes in cells. These are phospholipids where a cyclic hexahydroxy alcohol called inositol replaces base (Figure 3.10). The inositol is present as the stereoisomer, myo-inositol. On hydrolysis, the phosphoinositides yield 1 mole of glycerol, two moles of fatty acid, 1 mole of inositol and 1, 2, or 3 moles of phosphoric acid. Accordingly, mono-, di- or triphosphoinositides are found.

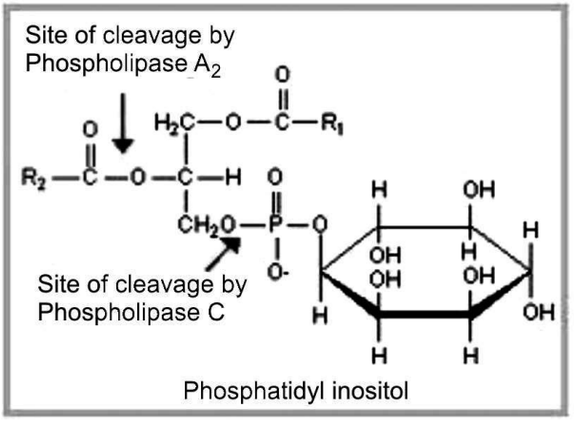


Fig. 3.10 *Structure of Phosphatidyl Inositol*

Phosphosphingosides (Sphingomyelins)

These compounds are commonly found in nerve tissue esp., in the myelin sheath of the nerve (hence their name, sphingomyelins) and apparently lack in plants and the microorganisms. In a syndrome called Niemann Pick disease, the sphingomyelins are stored in the brain in large quantities. These differ from other phospholipids in

NOTES

their lack of glycerol and the presence of another nitrogenous base sphingosine or a closely related dihydrosphingosine, besides choline, in place of glycerol. Sphingomyelins are electrically charged molecules and contain phosphocholine as their polar head groups. On hydrolysis, the phosphosphingosides yield equimolar amounts of fatty acid, phosphoric acid, choline and sphingosine or dihydrosphingosine but no glycerol. Thus, in these compounds the atomic ratio N/P is 2, in contrast to phosphoglycerides where this ratio equals unity. Phosphoinositides, however, do not contain a nitrogen base. It may be observed from the formula of sphingomyelins (Figure 3.11) that sphingosine carries the phosphoric acid on its primary alcohol group and the fatty acid by amide linkage on its primary amino group.

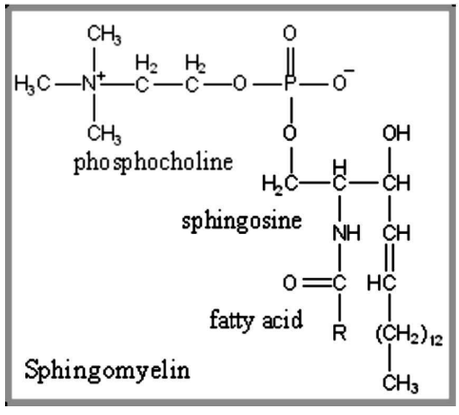


Fig. 3.11 *Structure of Sphingomyelin*

Lecithins (Phosphatidyl Cholines)

Lecithins are widely distributed in nature. Various oil seeds like soybean and the yeasts are important sources from plant world. In animals, the glandular and nervous tissues are rich in these lipids. The lecithins are required for the normal transport and utilization of as other lipids esp., in the liver of animals. In their absence, accumulation of lipids occurs in the liver to as much as 30 per cent against a normal value of 3–4 per cent, giving rise to a condition called 'fatty liver'. This fatty infiltration may lead to fibrotic changes, characteristic of the liver disease cirrhosis.

In addition to glycerol and 2 moles of fatty acids, the lecithins (Figure 3.9) also contain phosphoric acid and a nitrogen base choline at either the end or middle carbon atom of glycerol unit. Accordingly, two forms of lecithins, α and β are recognized. On complete hydrolysis, lecithin yields choline, phosphoric acid, glycerol and 2 moles of fatty acids. But partial hydrolysis of lecithins by lecithinases (active principles found in snake venoms) causes removal of only one fatty acid to yield substances called lysolecithins. These, therefore, contain only one acyl radical. When subjected into the blood stream by sting as a result of snake bite or by needle, the lysolecithins cause rapid rupture (hemolysis) of the red blood corpuscles.

NOTES

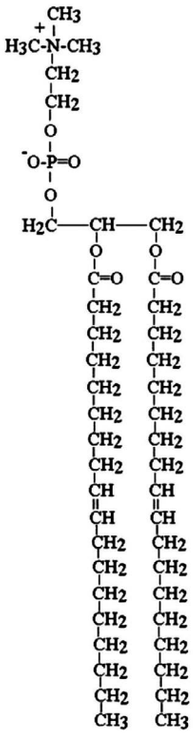


Fig. 3.12 Structure of Lecithin

Cephalins

The cephalins are closely associated with lecithins in animal tissues. These have also been identified from soybean oil. These are similar in structure to the lecithins except that the choline is replaced by either ethanolamine or serine. Serine is the biochemical precursor of ethanolamine. Accordingly, two types of cephalins are recognized, phosphatidyl ethanolamine and phosphatidyl serine. Like lecithins, the cephalins (Figure 3.13) also exist in 2 forms, α and β,depending upon the relative positions of the two substituent fatty acids. Since the primary amino group of ethanolamine is a weaker base than the quaternary ammonium group of choline, the cephalins are more acidic than lecithins. Moreover, the cephalins are comparatively less souble in alcohol than lecithins. Venoms containing lecithinases also split off fatty acids from cephalins, leaving hemolytic lysocephalins.

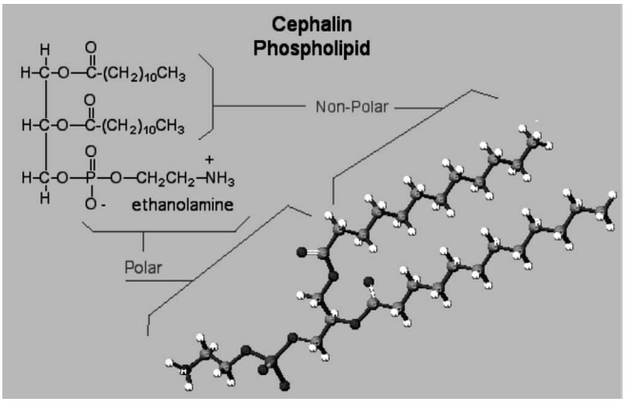


Fig. 3.13 Structure of Cephalin

NOTES

Plasmalogens (phosphoglyceracetals)

Plasmalogens constitute about 10 per cent of the phospholipids of the brain and muscle. These are apparently not found in significant quantities in plant tissues. Structurally, these resemble lecithins and cephalins but have one of the fatty acids replaced by unsaturated ether. Since the nitrogen base can be choline, ethanolamine or serine, three types of Plasmalogens (Figure 3.14) are accordingly distinguished: phosphatidal choline, phosphatidal ethanolamine and phosphatidal serine.

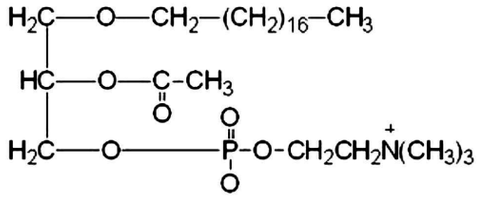


Fig. 3.14 *Structure of Plasmalogens*

Cerebrosides (Glycosphingosides)

The cerebrosides, as the name suggests, are important constituent of brain where they amount to about 8 per cent of the solid matter. These may also occur in tissues other than brain. Since the head group characteristically consists of one or more sugar units, cerebrosides are often called glycosphingosides. Like phospholipids, glycolipids are composed of a hydrophobic region, containing two long hydrocarbon tails, and a polar region, which now contains one or more sugar residues and no phosphate. Both phospholipids and glycolipids form self-sealing lipid bilayers that are the basis for all cellular membranes. In Gaucher disease, the cerebrosides appear in relatively large amount in the liver and the spleen. They are also present in large amount in the brain in Niemann–Pick disease. These are present in much higher concentration in medullated than in nonmedullated nerve fibres. There is evidence that they also occur in some plant organs. The structure of cerebrosides (Figure 3.15) is somewhat similar to that of phosphosphingosides. They contain a high molecular weight fatty acid, sphingosine and either galactose or glucose instead of choline but no phosphoric acid.

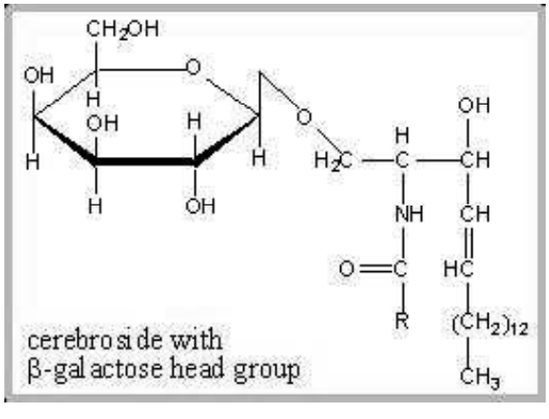


Fig. 3.15 *Structure of cerebrosides*

They have no electric charge since their polar head groups are neutral. In general properties, they resemble sphingomyelins. The name ceramide is commonly

used to designate the sphingosine— fatty acid (or N-acylsphingosine) portion of the cerebrosides. Here again the sphingosine carries the galactose by glycosidic linkage on its primary alcohol group and the fatty acid by an amide linkage on its primary amino group. Individual cerebrosides are differentiated on the basis of their fatty acid component (Figure 3.16).

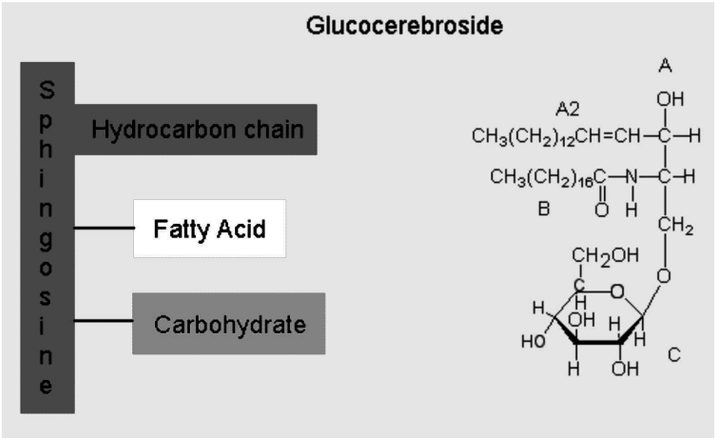


Fig. 3.16 *Structure of Glucocerebrosides*

The various classes, so differentiated, are as follows:

(a) **Kerasin**—contains saturated C-24 lignoceric acid.

(b) **Phrenosin** (cerebron)—contains a 2-hydroxy derivative of lignoceric acid called cerebronic acid.

(c) **Nervon**—contains an unsaturated homologue of lignoceric acid called nervonic acid

(d) **Oxynervon**—contains a 2-hydroxy derivative of nervonic acid called oxynervonic acid.

Gangliosides

Klenk in 1955 isolated a new type of glycolipid from brain tissue and named it as ganglioside**.** These are found in significant concentrations in ganglion cells of nervous tissue (hence so named) and also in most parenchymatous tissues like spleen and erythrocytes. They make up about 6 per cent of the membrane lipids in the gray matter of the brain. They are also found in lesser amounts in the membranes of most nonneural tissues. Gangliosides are thought to act as receptors for toxic agents like the pathogens, Vibrio cholerae influenza virus and tetanus toxin. They are also implicated to play a role in cell-cell interaction. The structure of gangliosides (Figure 3.17) is complex and related to that of cerebrosides in that they contain a ceramide (N-acylsphingosine) linked to a carbohydrate (galactose or glucose). In addition to these, the gangliosides also generally contain 2 additional moles of carbohydrates, 1 mole of N-acetylgalactosamine and from 1 to 5 moles of N-acetylneuraminic acid (NANA). In Tay- Sachs disease, the gangliosides are stored in relatively large amounts in the brain and spleen. More than 15 different gangliosides have been characterized and their structures determined. These are commonly abbreviated by the letter G, with a subscript M, D or T to designate that they contain one, two or three sialic acid and (N-acetylneuraminate or N-

NOTES

glycolyl neuraminate) residues respectively, and a number or letter to distinguish different members of a group from one another. In essence, they consist of an oligosaccharide chain attached to ceramide (N-acylsphingosine) by a sugar residue which is usually glucose. They lack phosphoric acid.

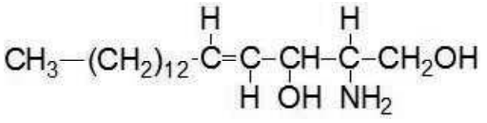


Fig. 3.17 Structure of Ganglioside

Sulfolipids

A glycolipid that contains sulfur is widely distributed in plants. It is localized in the chloroplasts but is also found in the chromatophores of photosynthetic bacteria. As the sulfur in this compound is present as a sulfonic group in a hexose, this may be included under a class of compounds called sulfolipids.

Sulphatides and Sphingolipids

A sulfate ester analogue of phrenosin, abundant in white matter of brain, is another sulfur-containing glycolipid (Figure 3.18). In it, the sulfate is present in ester linkage at C 3 of the galactose portion of the molecule. Members of this group of cerebroside sulfuric esters have been designated as sulfatides. The classification of compound lipids as presented here is now becoming obsolete. It is, however, more logical to classify them according to the main alcohol component. According to this system, the compound lipids fall under 2 categories: one in which the main alcohol component is glycerol (glycerophosphatides) and the other in which sphingosine is the main alcohol component (sphingolipids). The further classification of these two categories depends on the nature of the other alcohol components.

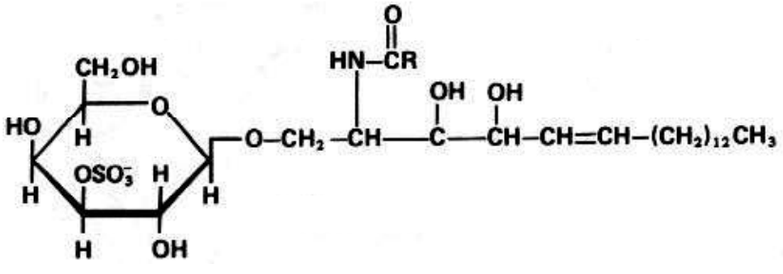


Fig. 3.18 Structure of Sulphatides

Lipoprotein: Composition, Function, and Role in Atherosclerosis

Lipoprotein are globular, micelle-like particles that consist of a nonpolar core of triacylglycerol and cholesteryl esters, surrounded by an amphipathic monolayer of protein, phospholipid, and cholesterol that is about 20 Å thick. These lipids and proteins associate noncovalently.

Classification of Lipoprotein

Lipoproteins with high lipid content will have low density, larger size and so float on centrifugation. Those with high protein content sediment easily, have compact

size and have a high density. They are separated by Ultracentrifugation. On the basis of their functional & physical properties, primarily on their densities, lipoproteins are categorized into 5 broad categories are as follows:

NOTES

1) Chylomicrons

Chylomicrons derived from intestinal absorption of triacylglycerol and other lipids. Their density is generally less than 0.95 while the mean diameter lies between 100- 500 nm. Chylomicrons, in connection with the movement of dietary triacylglycerols from the intestine to other tissues, are the largest of the lipoproteins and the least dense, containing a high proportion of triacylglycerols. Chylomicrons are synthesized in the endoplasmic reticulum of epithelial cells that line the small intestine, then move through the lymphatic system and enter the bloodstream via the left subclavian vein.

2) Very Low Density Lipoproteins (VLDL)

Very-low-density lipoprotein (VLDL) is derived from the liver for the export of triacylglycerol. Their density lies between 0.95- 1.006 and the mean diameter lies between 30-80 nm.

3) Intermediate Density Lipoproteins (IDL)

Intermediate density lipoproteins (IDL) are derived from the catabolism of VLDL, with a density ranging intermediate between Very low density and Low density lipoproteins i.e. ranging between 1.006- 1.019 and the mean diameter ranges between 25- 50nm.

4) Low-Density Lipoprotein (LDL)

Low-density lipoproteins (LDL), representing a last stage in the catabolism of VLDL. Their density lies between 1.019-1.063 and mean diameter lies between 18-28 nm. LDLs is responsible for transportation of cholesterol to extra-hepatic tissues.

5) High-Density Lipoproteins (HDL)

High-density lipoprotein (HDL) is involved in cholesterol transport and also in VLDL and chylomicron metabolism. Their density ranges between 1.063-1.121 and the mean diameter varies between 5-15 nm. It originates in the liver and small intestine as small, protein-rich particles that contain no cholesteryl esters and relatively little cholesterol.

3.3.3 Derived Lipids: Structure and Function of Cholesterol and Bile Acid

The group derived lipids is a 'catch all' group in Bloor's classification. It includes the hydrolysis products of simple and compound lipids and also various other compounds such as steroids, terpenes, fatty acids, alcohols, fatty aldehydes, ketones etc.

Steroids and Bile Acids

The steroids are one of the most studied classes of biological compounds and are often found in association with fat. Since they contain no fatty acids, they are

NOTES

nonsaponifiable, i.e., cannot be hydrolyzed by heating with alkali to yield soaps of their fatty acid components. Fats, on the other hand, are saponifiable and form soaps when hydrolyzed with alkali. The steroids may be separated from the fat after the latter is saponified since they occur in 'nonsaponifiable residue'. All steroids may be considered as derivatives of a fused and fully saturated ring system called cyclopentanoperhydrophenanthrene or sterane. This system consists of 3 cyclohexane rings (A, B and C) fused in nonlinear or phenanthrene manner and a terminal cyclopentane ring (D). The sterane nucleus along with the conventional numbering of the carbon atoms is shown in Figure 3.19. It may, however, be emphasized that in steroids the hexagonal rings are not the benzene rings as in them the valences of C atoms are fully satisfied by hydrogen bonds, unless shown otherwise.

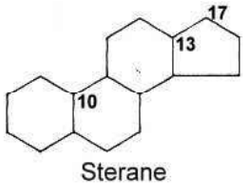


Fig. 3.19 *Structure of Sterane*

The double bonds, if present, are shown as such. The 'angular' methyl groups occur typically at positions 10 and 13 and constitute carbon atoms 19 and 18 respectively. The steroids may have one or more hydroxyl groups, one OH group being present usually on C 3. A side chain at C 17 is usual. This side chain serves as a convenient basis for classification of steroids. For example, the side chain contains 8, 9 or 10 C atoms in sterols, 5 C atoms in bile acids, 2 C atoms in adrenal cortical steroids and in progesterone and none in naturally occurring estrogens and androgens. Stereochemical considerations of steroids. Without even considering the substituent groups, there are 6 possible asymmetric centres present in the steroid nucleus. Consequently, the steroids have many potential stereoisomers. Each of 3 cyclohexane rings, on the basis of the tetrahedral theory, is capable of existing in 3-dimensional conformation either of a 'chair' or of a 'boat'.

Cholesterol

Cholesterol is undoubtedly the most publicized lipid in nature, because of the strong correlation between high levels of cholesterol in the blood and the incidence of diseases of the cardiovascular system in humans. It is not only an important component of some cell membranes and of plasma lipoproteins but also the precursor of many other biologically important steroids, such as bile acids and various steroid hormones. It is the principal sterol of higher animals and is especially abundant in nerve tissues and in gallstones. In occurs either free or as fatty esters in all animal cells. It was first isolated in 1784, from human gallstones which consist almost entirely of cholesterol and hence so named (cholesterol literally means 'solid alcohol from bile'). Its main sources are fish liver oils and the brain and spinal cord of cattle. White matter contains as much as 14 per cent, gray matter 5 per cent, spinal cord 12 per cent and liver about 1 per cent cholesterol.

Cholesterol is however, not found in plant fats. Its parent hydrocarbon is cholestane, C27H48. The structure of cholesterol was determined by the German

NOTES

chemist, Adolph Windaus, who received 1928 Nobel Prize in Chemistry. Cholesterol (Figure 3.20) has a molecular formula, C27H45OH. In addition to an OH group at C3, there is a double bond at C5. The hydroxyl group constitutes its polar head; the rest of the molecule is hydrophobic. It is a white crystalline solid and is optically active. The crystals are rhombic plates with one of the angles broken. It has a melting point of 149°C. Konrad S. Bloch, a Germany-born American biochemist, elucidated the biosynthetic pathway of cholesterol synthesis, one of the most complex known, for which he received the coveted Nobel Prize in Chemistry along with Feodor Lynen, a German and John Cornforth, a Briton, in 1964. Bloch combined work and pleasure by traveling to Bimini for collection of shark livers. No vegetable oil contains any cholesterol; hence, it is misleading for oil manufacturers to advertise any oil as 'low cholesterol' oil. Saturated fats of animals origin (such as butter, ghee and lard) do contain cholesterol but vanaspati, which is derived from vegetable oils, does not contain cholesterol.

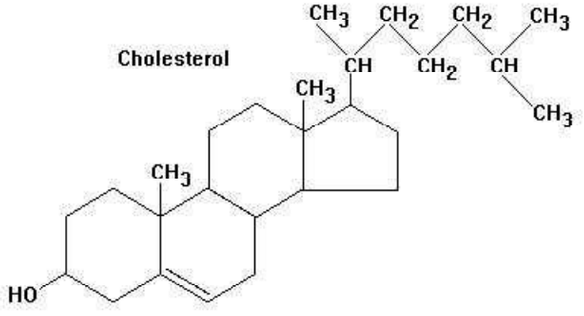


Fig. 3.20 *Structure of Cholesterol*

Only a little portion of the body cholesterol is derived from diet. The bulk of it is synthesized in the body. Unsaturated fatty acids lower blood cholesterol levels by controlling cholesterol's synthesis as well as its elimination from the body. Thus, vegetable oils when consumed within the recommended amounts help to maintain proper levels of lipids in the blood than do the saturated fats like butter, ghee, lard and vanaspati. These saturated fats elevate blood lipids (cholesterol and triglycerides) and reduce the ratio of good to bad cholesterol even when consumed within the recommended limits. High intake of dietary cholesterol can probably raise blood cholesterol levels and hence should be avoided, esp., by the elderly people and those having high blood cholesterol.

The level of cholesterol in the blood is measured in milligrams per decilitre (mg/dl), which is equivalent to parts per 1, 00,000). The levels range from less than 50 in infants to an average of 215 in adults to 1,200 or more in individuals suffering from a rare inherited disease called familial cholesterolemia. For those persons in the normal range, about two-thirds of their cholesterol is transported as LDLs. Most of the rest is carried by HDLs. An often-posed question relates to the consumption of eggs and meat, and sometimes even milk, because of the relatively high content of cholesterol in these foods. Milk is no doubt a very good and safe food. Persons aspiring to control their cholesterol intake (and calories) must use defatted or skimmed milk. Egg and meat proteins are of high grade and

NOTES

desirable especially for the growing children. Those desirous of reducing their cholesterol intake should eat only the white of egg and lean on white meat like chicken and fish. Nutritionally, fish is a very good food since it has high-quality protein (fish fat is rich in n-3 PUFA) and its flesh has very little cholesterol. Organ meats like brain and liver are rich in cholesterol and should hence be avoided by those wanting to restrict cholesterol.

The notion that cholesterol is a poisonous substance in the body is misleading. It is the parent hydrocarbon from which many important hormones (corticosteroids and sex steroids, for example) are synthesized. The problem creates when blood cholesterol level rise and the level of LDL cholesterol, which carries cholesterol to the tissues, goes below the desired level. HDL-cholesterol or good cholesterol removes cholesterol from the tissues and helps eliminating it from the body. High blood cholesterol results in the deposition of cholesterol in the arteries and narrows the passage, a condition called atherosclerosis. Also, the arterial surface roughens due to cholesterol deposition, which leads to frequent blood clotting or thrombosis. Both are major health hazards. Studies, however, point out that even very low levels of blood cholesterol are too not desirable. The thumb rule, therefore, is moderation. Both too much and too little of fat may be harmful. In blood about two-thirds of the cholesterol is esterified mainly to unsaturated fatty acids, the remaining portion occurring as the free alcohol. Reduction of the double bond gives rise to 2 products, coprostanol and cholestanol.

Coprostanol (Coprosterol)

It occurs in feces and is produced in the intestine as a result of bacterial action on the double bond of cholesterol (Figure 3.21). The A/B junction is -cis in coprostanol in contrast to -trans in cholesterol.

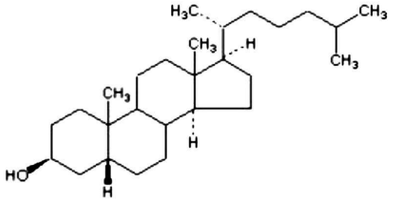


Fig. 3.21 *Structure of Coprostanol*

Cholestanol

It occurs as a minor constituent of the sterols of blood and other tissues (Figure 3.22). Here, the A/B junction is in -trans form.

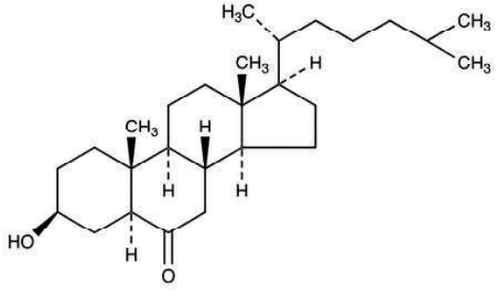


Fig. 3.22 *Structure of Cholestanol*

NOTES

Ergosterol

It is present in ergot (hence its nomenclature), yeast and the mould Neurospora. Its parent hydrocarbon is ergostane, $C_{28}H_{50}$. Ergosterol (Figure 3.23) has a molecular formula, $C_{28}H_{43}OH$ with one OH group at C3 and 3 double bonds at C5, C7 and C22. It is also optically active. Rupture of the ring B by UV radiation produces vitamin D2 which is chemically known as ergocalciferol. A similar compound cholecalciferol (or vitamin D3) is, however, obtained from 7- dehydrocholesterol on irradiation with UV light.

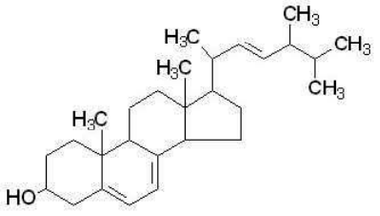


Fig. 3.23 *Structure of Ergosterol*

Lanosterol (Cryptosterol)

It is a major constituent of wool fat and is also present in minor quantities in liver and yeast. Lanosterol is a C 30 compound with twin methyl groups at C 4 and a third angular methyl group on C 14. There are 2 double bonds at C8 and C24 (see Figure 3.24). It is an intermediate in the biosynthesis of cholesterol. A number of sterols have also been obtained from various plant and animal sources, for example **stigmasterol** (from soyabean and wheat germ oils), **spinasterol** (spinach and cabbage), **sitosterol** (many higher plants), **ostreasterol** (oysters) and **chondrillasterol** (marine sponges). The bile acids are the important end products of cholesterol metabolism in higher plants. About 20 natural bile acids have been characterized. All these are derived from a C 24 parent steroid, cholanic acid and resemble coprostanol in having rings A and B in cis form. The most abundant bile acids in human bile are cholic acid (25–60 per cent of the total bile acids), chenodeoxycholic acid (30–50 per cent) and deoxycholic acid (5–25 per cent). Various bile acids differ from each other in the number and position of OH groups which are all in a configuration. In the bile acids, the number of OH group(s) may be 1, 2, or 3 and the position of OH group(s) may be any of the following: 3, 6, 7, 11, 12 and 23. The side chain is usually made up of 5 carbon atoms and bears the carboxyl group. They are water-soluble and are powerful detergents. Several important poisons are based on the steroid structure.

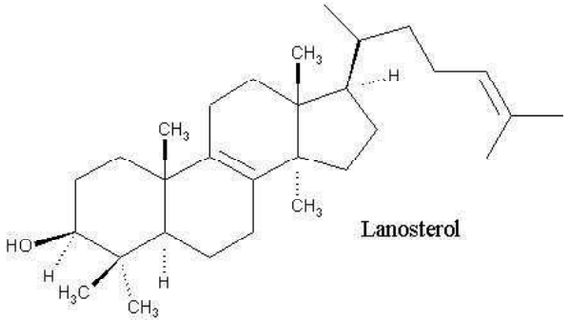
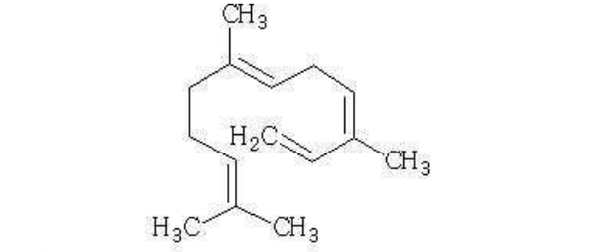


Fig. 3.24 *Structure of Lanosterol*

NOTES

Terpenes

Among the nonsaponifiable lipids found in plants are many hydrocarbons known as terpenes (from turpentine). In general, these hydrocarbons and their oxygenated derivatives have lesser than 40 carbon atoms. The simplest terpenes are called monoterpenes and conform to the formula $C_{10}H_{16}$ (equivalent to 2 isoprene units), those with the formula $C_{15}H_{24}$ are called as sesquiterpenes, with $C_{20}H_{32}$ as diterpenes and with $C_{30}H_{48}$ as triterpenes (Figure 3.25). Terpenes with 40 carbon atoms or tetraterpenes) include compounds called carotenoids. In fact, O. Wallach (1910 Nobel Laureate in Chemistry) was the first to point out, in 1887, that nearly all the terpenoids are made of varying number of repetitive units (C_5H_8), called isoprene units. His finding later came to be known isoprene rule. Structurally, isoprene is a 5-carbon diene. The carbon skeletons of open-chain monoterpenoids and sesquiterpenoids are even not only the presence of isoprene units but their special type of arrangement is found to be present in nearly all the terpenoids. In 1925, Ingold, formulated this observation under another rule, the special isoprene rule, according to which the isoprene units in terpenoids are usually joined in head-to-tail linkages or 1,4 linkages (the branched end of the isoprene unit was considered as the head). Exceptions, however, do occur for these two isoprene rules. For example, cryptone, a natural terpenoid, contains only 9 carbon atoms instead of 10. Also, in lavandulol, the 2 isoprene units are not joined in head-to-tail manner, and in carotenoids, the two halves (each with 4 isoprene units) are linked with each other by tail-to-tail at the centre.



(3*Z*,6*E*)-3,7,11-trimethyldodeca-1,3,6,10-tetraene
Farnasene

Fig. 3.25 *Structure of Terpenes*

Monoterpenes and Sesquiterpenes

The fragrances of many plants arise from volatile C10 and C15 terpenes. These and their oxygenated derivatives occur as components of the essential oils, some of which are used as perfumes. Important monoterpenes are myrcene (from oil of bay), geraniol (from rose oil), limonene (from lemon oil) and menthol (from peppermint oil). The structure of monoterpenes have been shown in figure 3.26.

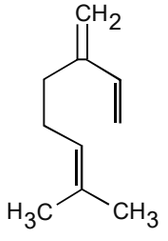


Fig. 3.26 *Structure of Monoterpenes*

Diterpenes

These are usually found as substituents of the resins and balsams. The two resin acids, abietic and sapietic are the best known tricyclic diterpenes. Vitamins A1 and A2 and their aldehydes called retinenes are important monocyclic derivatives of diterpenes.

Phytol

It is an acyclic diterpene and is obtained from the hydrolysis of chlorophyll. It was isolated from nettles by Paul Karrer et al in 1943. The phytol molecule has 2 chiral (= asymmetric) centres, 7 and 11 (Figure 3.27). Natural phytol is very weakly dextrorotatory.

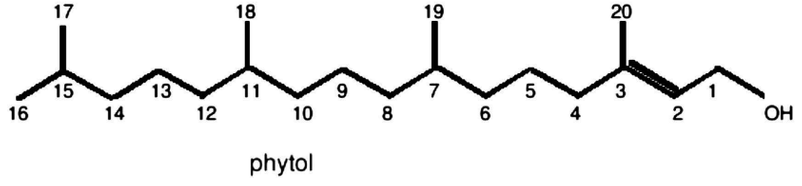


Fig. 3.27 *Structure of Phytol*

Triterpenes

These are not widespread in nature but are significant in that some of them are intermediates in the biosynthesis of cholesterol. Two such compounds are: a tetracyclic alcohol, lanosterol and an acyclic hydrocarbon, squalene. Squlaene (C30 H50; b.p. 240°–242° C) was first isolated from the liver of sharks (genus Squalus), hence so named. The olive oil and several other vegetable oils are other sources. It has also been detected in the leaves. Each squalene molecule has 6 double bonds. The double-bond system present is of nonconjugated type. The conjugated double bonds are, however, absent from the molecule.

Polyterpene

However what should be mentioned here is rubber, a polyterpene present in the latex of many tropical plants. A molecule of rubber is composed of about 500 to 5,000 isoprene units joined in a long straight chain. When the bark of the rubber tree is cut, latex slowly exudes from the cut. Addition of acetic acid coagulates the rubber, which is then separated from the liquor and either processed into blocks or rolled into sheets, and finally dried in a current of warm air, or smoked.

Carotenoids

The carotenoids are tetraterpenes. These are widely distributed in both the plant and animal kingdoms but are exclusively of plant origin. These occur in unsaponifiable residue of plant and animal lipids. They are isoprene derivatives with a high degree of unsaturation. Because of the presence of much tomato (lycopene) and that of carrot (α- and β-carotene) are red while many oxygen-containing carotenoids are yellow (xanthophylls). Since the olefinic (= double) bonds permit cis-trans isomerism, numerous forms are possible. Most carotenoids, however, exist in all-trans form. The presence of long hydrocarbon chain in

carotenoids makes them lipid-soluble; they are hence also called lipochromes or chromolipids. The structure of carotene and Xanthin is shown in Figure 3.28.

NOTES

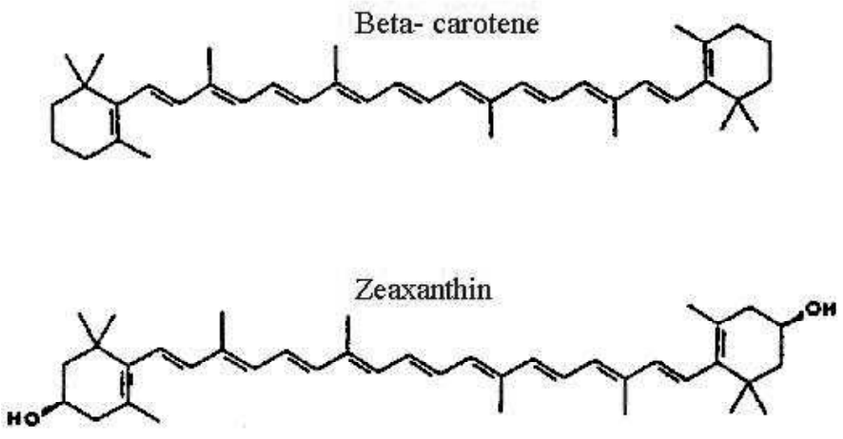


Fig. 3.28 *Structure of Carotene and Xanthin*

Lycopene

Lycopene ($C_{40}H_{56}$) is the main pigment of tomato, paprika and many other fruits. It is a highly unsaturated, unbranched, long chain hydrocarbon (a polyene) and is composed of two identical units ($C_{20}H_{28}$), joined by a double bond between C15 and C152 . Each of these units may, in turn, be considered to have been derived from 4 isoprene units (C_5H_8). A molecule of lycopene (Figure 3.29) in all contains as many as 13 double bonds, of which 11 are conjugated double bonds.

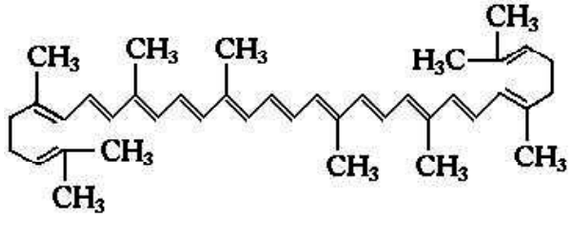


Fig. 3.29 *Structure of Lycopene*

Carotene

Another group of naturally occurring carotenoids, with the same molecular formula as that of lycopene, is carotene. Carotene was first isolated by Wackenroder (1831) from carrots (this was the origin of the name carotin, which was later on modified to carotene). The three types of carotenes are:

1. α-carotene—violet crystals; m.p. 187°C; optically active (dextrorotatory).
2. β-Carotone—red crystals; m.p. 183°C; optically inactive
3. γ-Carotene—dark red crystals; m.p. 152"154°C; optically inactive

The carotenes are obtained commercially by chromatography. The two best sources are carrots and alfalfa. Their relative proportion, however, varies with the source; for example carrots contain 15 per cent α-, 85 per cent β- and 0.1 per cent γ-form. These are unstable to air, heat, acids and alkalies. It is to be noted that the ends of the lycopene chain can easily close up to form rings, with the disappearance of the double bonds.

Xanthophylls (Figure 3.30) are characterized by the presence of hydroxyl groups in the ionone rings of carotenes, in the para position to the long chain. The xanthophyll of the leaf (lutein, $C_{40}H_{56}O_2$) is derived from α-carotene while that of corn (zeaxanthine) from β-carotene. The xanthophylls of crustaceans (astaxanthine) is, however, more rich in oxygen. Astaxanthine is responsible for the appetizing redness of boiled lobsters.

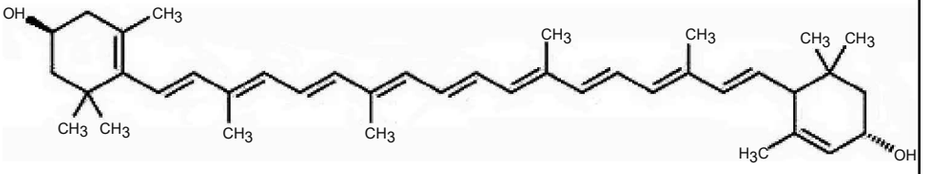


Fig. 3.30 *Structure of xanthophylls*

Check Your Progress

7. What are triacylglycerols?
8. Give two examples of biological waxes.
9. Name two important sources of lecithins.

3.4 PROPERTIES OF LIPID AGGREGATES

Lipid molecules are composed primarily of hydrocarbon groups which make them hydrophobic (water insoluble). Some of the lipid possess hydrophilic groups too which make them water soluble. Such molecules which contain both hydrophobic (a nonpolar tail) and hydrophilic (a polar head) fractions are known as amphipathic molecules (amphi-both, pathos-passion). This nature of lipid allows them to form microscopic aggregates in aqueous environment because when amphipathic molecules are mixed with water, the polar groups orient (heads) themselves towards the aqueous phase while the non-polar orients (tails) towards the opposite direction.

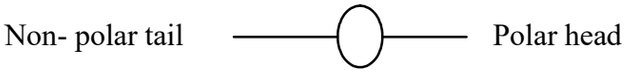


Fig: 3.31 *Amphipathic Lipid*

Examples of amphipathic lipids are phospholipids, sphingolipids, bile salts and cholesterol (to some extent). This leads to the formation of a phase or a system which is separated from the surrounding aqueous phase.

NOTES

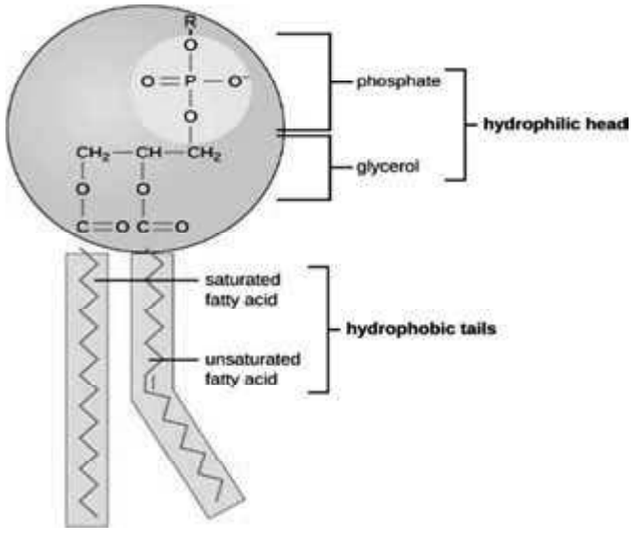


Fig: 3.32 *Structure of Phospholipid Showing Hydrophilic Head and Hydrophobic Tail*

Lipids may form the stable aggregated like micelles and thermo thermodynamically unstable aggregates like bilayer sheets. Lipids are usually single and double chain amphiphiles. Single chain amphiphiles form a monolayer and micelle. Oppositely, double chain amphiphiles form a bilayer and liposome.

3.4.1 Micelles, Bilayers, and Liposomes and their Possible Biological Functions

Now, we will discuss the structure and functions of Micelles, Bilayers, and Liposomes in the following sections.

Micelles

Micelles are the aggregates of amphipathic molecules formed when the molecules are mixed with an aqueous solution. As an amphipathic molecule has two groups, so they form micelle in which polar or ionic heads form an outer shell in contact with water, while non-polar tails are positioned in the interior. Aggregation number is the quantity of amphipathic molecules forming an aggregate and is used to describe the size of the micelle. When the concentration of the amphiphilic molecule reaches a given concentration called Critical Micelle Concentration (CMC), it led to the formation of micelle aggregates. Before the concentration is reached, micelles are absent.

The property of micelles to micellize, depend on two factors: Repulsion of water by the non-polar tails and the repulsion between the polar or charged heads, a destabilizing effect on the aggregation process. Non-polar tails point in interior as a consequence of avoiding water and due to hydrophobic interaction among these non-polar tails, the favourable arrangement of the micelle occurs in the interior. It is the dominant effect in the formation process of these aggregates. The repulsion among the charged heads on the surface of the micelle is reduced by the presence of oppositely charged ions (counter-ions).

Micelle formation is not only seen in polar solvents or aqueous solutions, but it is also seen in non-polar solvents. This aggregates of micelles in non-polar

solvents are known as inverse micelles. In this case, hydrocarbon tails are exposed to the solvent, while the polar heads point towards the interior of the aggregate to escape the contacts with the solvent and thus, they are able to hold relatively large amounts of polar solute or solvent in their interior.

NOTES

This arrangement is useful for transportation of polar solutes through a non-polar solvent as the polar solutes can be dissolved in this pocket.

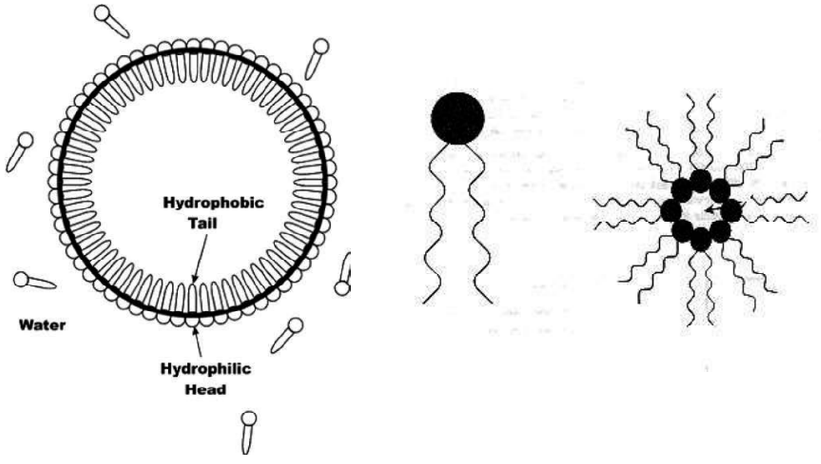


Fig. 3.33 *(A) Micelle Formation (B) Reverse Micelle: Arrow Depicting Water 'Sequestered' in the Interior*

Micelles have various functions in industrial and biological fields for their capability of dissolving and moving non-polar substances through an aqueous medium.

- They are broadly used in electrophoresis and chromatography as a media of separation.
- They are efficiently used as nano-carries in for various applications namely gene delivery, diagnostic imaging, and drug.
- They act like emulsifiers when surfactants are above the critical micelle concentration allowing a compound to dissolve which are usually insoluble. As, they can incorporate insoluble species in their core. For example, detergents has the ability to clean lipophilic material like oil stains, wax, etc. from the dirty surfaces as they are less soluble and cannot be removed by washing from water alone. Detergents lowers the surface tension of water too, which make it easier to remove the stain from the surface.
- Micelles plays an important role in removing complex lipids and fat-soluble vitamins from the human body by absorbing them. Complicated lipids (e.g., lecithin) and lipid-soluble vitamins (A, D, E, and K) in our body can be absorbed within the micelle in the small intestine by the micelles of fatty acids of bile salts which are formed in the liver and secreted by the gall bladder.
- Micelle has been proved very benefial for their use in chemical reactions. The interior of micelles can be used to perform the chemical reactions as they create the conditions which are suitable for the reaction to the

NOTES

performed and feaseable to collect the reaction products like hydrophobic molecules. Because of this property, they can make a multi-step reaction very easy to perform which definetely incraeses the yield of the reaction. They can also to used to create required conditions (e.g., extreme pH) of the reaction and reduces the need for extra solvents and thus, side products. So, Micellular chemistry is also known as green chemistry. But, with the following advantages, they can also inhibit the chemical reactions for example if the reacting molecule form micelles, which now shield a molecule component, it makes it deprived of oxidation.

- Micelles also helps in the process of targeted drug delivery by trapping the gold nanoparticles within them.

Lipid Bilayer

Lipid bilayer is a membranous structure consisting of two layers of amphipathic lipid molecules. As discussed in the previous section, when mixed with the polar solvents, the arrangement of an amphipathic molecule is such that the hydrophilic head point to the outside or are in contact of the polar phase whereas the hydrophobic tail points to the interior repelled by the water. By the hydrophobic interaction between these tails, they are attracted towards each other and aggregate together. So, when exposed to water, the lipid molecules transform into a two-layer structure with the hydrophobic tails of both the layers pointing at the centre of the sheet. Giving the structure of two leaflets, the layers contain no water between them and so excludes all those molecules also which are dissolved in water like sugars and salts. As the polar groups are exposed outside, they create hydrogen bonds with other polar molecules.

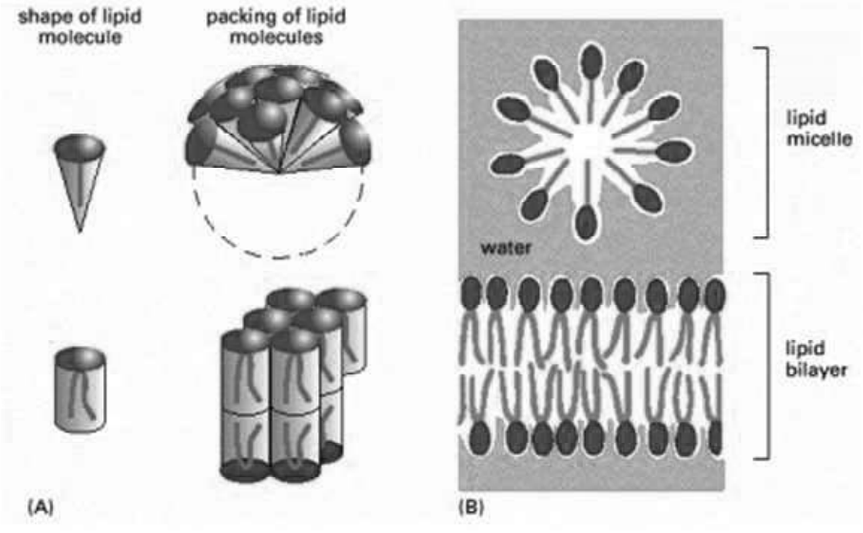


Fig: 3.34 *Arrangement of Lipid Molecules in an Aqueous Environment*

(***Source:*** Alberts, B. 2015. *Molecular Biology of the Cell*. United Kingdom: Garland Science, Taylor and Francis Group.).

Figure 3.34 shows:

- Wedge shaped molecules giving the structure of a micelle on aggregation and cylinder-shaped molecules forming lipid bilayer structure on aggregation.
- Cross section views of micelle and lipid bilayer.

NOTES

The major lipid molecules found in lipid bilayer are phospholipids, glycolipids and sterols. Lipid bilayer performs various biological functions, like:

- Lipid bilayer structure is the foundation of all the membranes found in the biological system, together with the several membrane proteins. Inner and outer monolayers have different lipid compositions, which make the two faces of bilayer functionally different.
- They show prominent role in separation. It forms the barrier of the aqueous compartment from their neighbouring surroundings. The existence of a cell without a membrane cannot be even imagined. It differentiates the cells, provide the knowledge of 'self' and 'non-self'. The amphipathic property of lipid bilayer does not allow hydrophilic molecules to cross the hydrophobic bilayer core.
- As lipid bilayer structures is composed of integral membrane proteins also, they provide the site for signal transduction. Example of such protein is CD59 which identifies cells and thus prevents their destruction by the immune cells. The integrated proteins which are embedded such that they have there one face at the outside of the cells and other in the interior, form the channel to pass the contents from surroundings to into the cells. Such a protein is G protein-coupled receptor (GPCR) which are important as they can sense the surroundings and protects the cells. This function of them makes them the suitable target for drug designing for many drugs.

Liposomes

The structure of liposomes was deduced for the first time by British haematologist, Alec D Bangham in 1961**.** Liposome consists of words 'Lipos' and 'Soma', meaning fat and body, respectively. Liposomes are the closed, sphere-shaped vesicles which consist of bilayer arrangement of phospholipids. As phospholipids molecules contain one hydrophobic tail and one hydrophilic head region, they form the tightly arranged structure of a bilayer in which the two tails being attracted hydrophobically face towards each other and hydrophilic head face outwards towards the polar medium when mixed with it. The cavity which is created inside can carry aqueous solution as the polar heads interact with them. This property is used for the transportation of aqueous, polar solute or solvent by enclosing it in the hollow cavity inside liposome.

NOTES

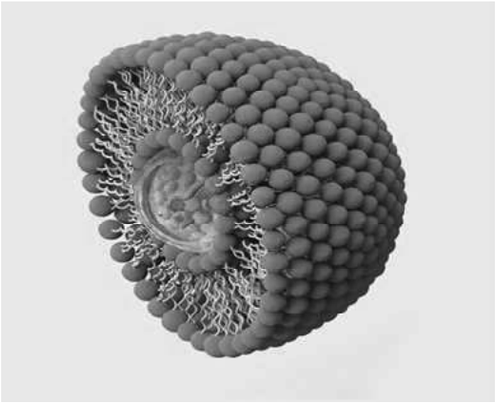


Fig 3.35 *Structure of a Bilayer Liposome Carrying the Nutrient inside the Cavity*

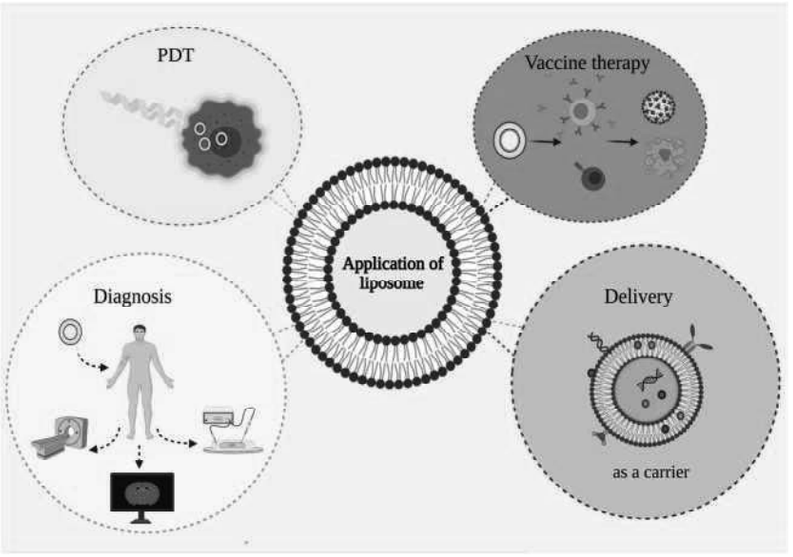
A micelle is the monolayer arrangement of phospholipids in the form of sphere. But liposome is the bilayer arrangement of phospholipids in the form of sphere. Liposomes can be made with various different combinations of structure, size, shape, and surface properties as they show a very wide diversity in terms of size and shape. The liposome has a size varying from very small (0.025 ìm) to large (2.5 ìm) vesicles. Also, it can have one or bilayer membranes. Size and number of bilayers affects the functioning and use of liposomes. So, liposomes can also be classified on the basis of their size and number of bilayers into: a. multilamellar vesicles (MLV) and b. oligolamellar vesicles (OLV) and c. unilamellar vesicles.

Unilamellar liposomes, have the single phospholipid bilayer sphere vesicle enclosing an aqueous solution. Whereas, multilamellar liposomes vesicles resemble the onion structure with many phospholipid bilayers. Unilamellar vesicles can be further classified into two categories: a. giant unilamellar liposomes/vesicles (GUV) groups, b. large unilamellar vesicles (LUV) and c. medium unilamellar vesicles (MUL) and d. small unilamellar vesicles (SUV).

Today, they have shown their important roles in various fields. Like they are a very important tool in various scientific disciplines, including mathematics and theoretical physics, biophysics, chemistry, colloid science, biochemistry, and biology and very useful in reproduction and reagent. They are extensively used as a model for studying cell membranes and organelles. This has helped in making the modern cell theory by studying the internal solution compared to the external solution by integrating various proteins into the lipid bilayer and testing them.

Due to their simple structure, they are easy to watch and are used for predicting and identifying the different methods cell used for the transportation of molecules through them.

As liposomes are just a small cell and can carry aqueous solution protected in their cavity, it is being researched for using them a delivery agent or a small machine to deliver required nutrient to the plants as they are biodegradable over time but up to that can still carry the aqueous intake protectively.



(Pooria Nakhaei et al., 2021)

Fig. 3.36 *Diverse Applications of Liposome*

The most important use of liposome is in the delivery method. The principle behind this function is that the special proteins, which can be recognised by the receptor proteins on the target cell can be embedded on them. Then, as the protein identifies the target, it delivers its contents to the target cell by binding to it. So, this can be used to deliver drugs to the targeted cells, can identify and kill cancerous cells, and in gene therapy. Even DNA can be delivered to the specific tissues by using targeted liposomes. In which the DNA after integration and transportation inside the cell, can be read by the cell machinery and can be used to produce the protein and reverse the deficiency a cell would be having. If successfully researched further, this function can be used to a very high extent for treating various genetic and other disease which occurs due to the loss of a function gene due to some mutation or which occurs due to some dysfunction in the protein synthesis by the cell. They have created an advanced technology for delivering active molecules to the site of action. From being used as a conventional to the progression of them as a 'second-generation liposomes', allowing formulation of long-circulating liposomes by modulation of the lipid composition, size, and charge of the vesicle.

3.4.2 Biological Membranes

The biological membrane covers the cells, shield and separate it from the surroundings. Biological membrane is a semi-permeable membrane which allows the selective movement of molecules only from outside into the cell and from cell to outside. It consists of a lipid bilayer or phospholipid bilayer structure. With lipids, biological membranes also form membrane proteins and sugars, which plays a key role in maintain integrity of their structure, maintain organization and selective movements of molecules to and fro the membrane. Sugars are found on one side of the bilayer only, and are attached by covalent bonds to some lipids and proteins.

NOTES

Major lipids found in biological membranes are namely phospholipids, glycolipids and sterols.

The sugars molecules are used as markers because of structural diversity of their chains. These molecules are bonded to the lipids and proteins. For example, the blood group of an individual is determined by the antigens of sugar chains present on the surface of red blood cells. Also, this are the ones recognized by antibodies to cause an immune response which necessities the matching blood groups for blood transfusions. Other carbohydrate markers are present in disease (e.g., specific carbohydrates on the surface of cancer cells), and can be used by doctors and researchers to diagnose and treat various conditions.

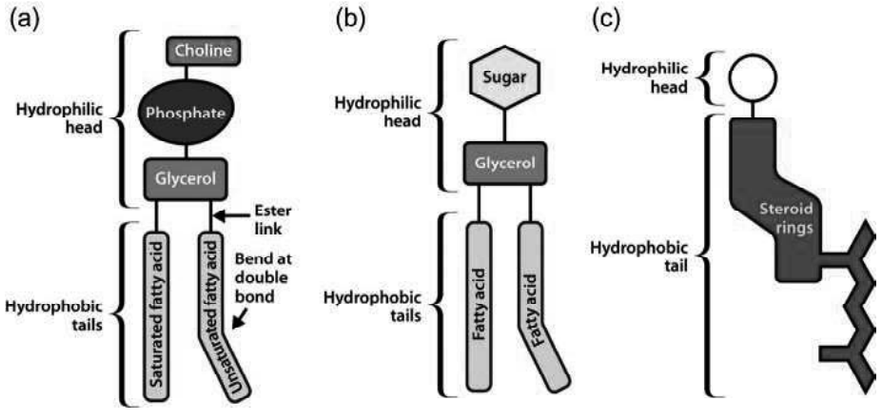


Fig. 3.37 *Types of Membrane Lipids:*
A. Phosphatidylcholine, a Glycerophopholipid
B. Glycolipid
C. Sterol

The Biological membranes are formed by addition to a pre-existing membrane. Like, in prokaryotes it takes place at the inner leaflet of the plasma membrane which faces the cytoplasm and in eukaryotes, on ER membrane synthesis occurs on the cytoplasmic leaflet of the ER membrane (termed the 'inside' of the cell). From there, Lipids leaves and travel through the secretory pathway for distribution to various subcellular compartments or the plasma membrane.

Both the membranes differ in their lipid composition. Mammalian cells contain on their outer leaflet of the plasma membrane predominantly Phosphatidyl Choline (PC) and sphingomyelin, whereas Phosphatidyl Serine (PS) and Phosphatidyl Ethanolamine (PE) are found on the inner leaflet. Lipid composition varies within organelles and within eukaryotic cells. For example, cholesterol synthesis occurs on ER, but their membrane has a very low quantity of cholesterol unlike the large quantity transported to other cellular membranes from here. Some membrane-bound enzymes need specific lipid head groups for their functioning. The head groups of some lipids create docking sites for specific cytosolic proteins.

Membrane proteins act like a nanomachines facilitating membranes to allow the transfer of molecules from cell comportments and cells and sending and receiving the signals for signal transduction. If membrane proteins would not be present, the membrane would be a impassable barrier in which no cells would be able to communicate with one another, signal molecules, transport nutrients into the cells,

take out the waste products. In order to survive, both unicellular and multicellular organisms need membrane proteins. This determines the molecule which will be permeable and molecules a cell can recognize. Membrane proteins control the active or passive transport of materials across the cell membrane or other subcellular membranes surrounding organelles. For a cell or an organism to survive, it is crucial that the right substances enter cells (e.g., nutrients) and the right substances are transported out of them (e.g., toxins).

Biological membranes have various functions like:

- Selective permeability as the movement across it is truly based on the size, charge, and other chemical properties of the atoms and molecules attempting to cross it. The mechanical or elastic properties membrane also favours the movement selectively as they can change their shape like a fluid and facilitate the movement.
- Different types of membranes also create intracellular organelles: endosome; smooth and rough endoplasmic reticulum; sarcoplasmic reticulum; Golgi apparatus; lysosome; mitochondrion (inner and outer membranes); nucleus (inner and outer membranes); peroxisome; vacuole; cytoplasmic granules; cell vesicles (phagosome, autophagosome, clathrin-coated vesicles, COPI-coated and COPII-coated vesicles) and secretory vesicles (including synaptosome, acrosomes, melanosomes, and chromaffin granules). Different types of biological membranes carry different protein and lipid composition, which make their physical and biological properties differ.
- The membrane is fluid in nature which has several functions like, it enables membrane proteins for their rapid diffusion from the plane of the bilayer which make their interaction easier with one another and is also beneficial in cell signalling. Also, it allows membrane lipids and proteins to diffuse from their synthesis sites to the sites where they are needed. It allows different membranes to fuse with one another and mix their molecules, and make it sure to divide the equal content when the cell divides. Without the fluid nature of membranes, the reproduction, growth of cell would not be possible.
- They participate directly in the process of cell signalling, for example, phosphatidylserine triggered phagocytosis. Normally position of phosphatidylserine is on the interior side of cell membrane. Scramblase, a protein during programmed cell death equilibrates this distribution, displaying phosphatidylserine on the extracellular bilayer face. The presence of phosphatidylserine then triggers phagocytosis to remove the dead or dying cell.

3.4.3 Fluid Mosaic Model of Membrane Structure

The fluid mosaic hypothesis was formulated by Singer and Nicolson in the early 1970. According to this model, membranes are made up of lipids, proteins and carbohydrates. The main lipid membrane components are phospholipids. These molecules are amphiphilic, i.e., they have one polar part attracted by water (hydrophilic) and one a polar component repelled by water (hydrophobic). When

NOTES

they are diluted in water, amphiphiles spontaneously adopt the most thermodynamically stable molecular structure, namely the one that maximizes both hydrophilic and hydrophobic interactions. These interactions may be affected by several parameters, such as the chemical nature of the molecules, their size, the salinity and pH of the solution. In biological conditions, cell phospholipids form a bilayer in which hydrophobic tails face each other in the core of the structure whereas the hydrophilic heads interact with the surrounding water.

Since proteins are also amphiphilic molecules, the same constraints apply to them. Some proteins (called intrinsic or integral) are embedded in the lipid bilayer matrix where they are able to establish hydrophobic and hydrophilic interactions with their respective lipid counterparts. Other proteins, called extrinsic or peripheral proteins, can also be transiently associated with membrane surfaces through weaker interactions (Figure 3.38). Finally, carbohydrates can be linked to either proteins or lipids, resulting in glycoproteins or glycolipids.

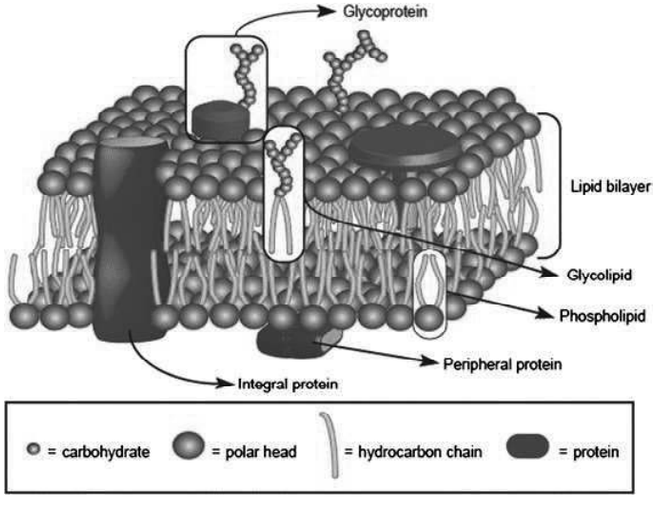


Fig. 3.38 *Fluid Mosaic Model of Plasma Membrane*

(***Source:*** Cooper, G. M. 2000. *The Cell: A Molecular Approach*. United Kingdom: ASM Press.)

The 'mosaic' term of this model means the mixture of lipids and intrinsic proteins in the membrane. These boundaries are also 'fluid' because their components can move, allowing both diffusion of components and local specific gatherings. Other lipids, such as cholesterol, act as membrane fluidity regulators. Phospholipid movements are generally restricted to lateral drift, because the cross of the membrane from one side to the other requires the energetically unfavorable transient contact of their hydrophilic head with the hydrophobic membrane core. Thus, the transfer of molecules from one side of the membrane to the other generally involves the activity of some specific integral membrane proteins, called flippases. For the same reasons, integral proteins can diffuse within the lipid matrix but they seldom switch their polarity from one membrane side to the other. As a result, lipid, protein and carbohydrate composition are different between the two monolayers, a characteristic that is referred to as membrane asymmetry.

NOTES

Membrane associated proteins found on the membrane as deduced by Singer and Nicolson can be distinguished into two classes, a. peripheral and b. integral membrane proteins. Peripheral membrane proteins are the protein which lies on the periphery of the membrane. Because of this arrangement of them, they can dissociate from the membrane by treatments with polar reagents, such as solutions of extreme pH or high salt concentration without damaging the phospholipid bilayer as these proteins are not inserted into the hydrophobic interior of the lipid bilayer but are indirectly bonded to the membranes through protein-protein interactions. After dissociation from the membrane, they are soluble in aqueous buffers. These interactions frequently involve ionic bonds, which are disrupted by extreme pH or high salt.

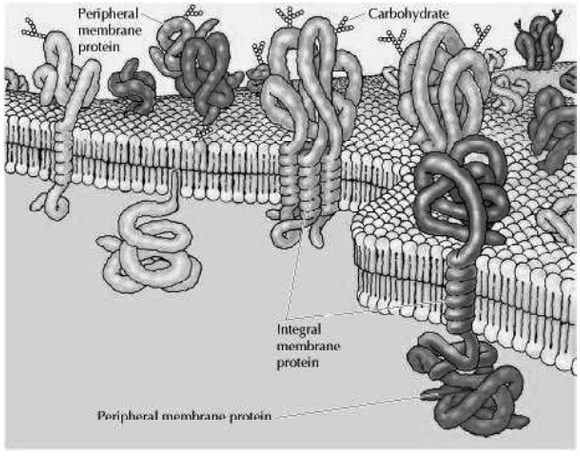


Fig. 3.39 *Arrangement of Integral Membrane Proteins and Peripheral Proteins on Lipid Bilayer Membrane*

(**Source:** Cooper, G. M. 2000. *The Cell: A Molecular Approach*. United Kingdom: ASM Press.)

Peripheral membrane proteins, in contrast to integral membrane proteins can be released only by treatments that disrupt the phospholipid bilayer as their portions of these are inserted into the lipid bilayer, so they can be dissociated only by reagents that disrupt hydrophobic interactions. Detergents are the most commonly used reagents for solubilisation of integral membrane proteins which being the small amphipathic molecules contain both hydrophobic and hydrophilic groups. The hydrophobic portions of detergents displace the membrane lipids and bind to the hydrophobic portions of integral membrane proteins. Because the other end of the detergent molecule is hydrophilic, the detergent-protein complexes are soluble in aqueous solutions.

The membrane-spanning portions of transmembrane proteins are usually á helices of 20 to 25 hydrophobic amino acids that are inserted into the membrane of the endoplasmic reticulum during synthesis of the polypeptide chain. These proteins are then transported in membrane vesicles from the endoplasmic reticulum to the Golgi apparatus, and from there to the plasma membrane. Carbohydrate groups are added to the polypeptide chains in both the endoplasmic reticulum and golgi apparatus, so most transmembrane proteins of the plasma membrane are glycoproteins with their oligosaccharides exposed on the surface of the cell.

NOTES

Studies of red blood cells have provided good examples of both peripheral and integral proteins associated with the plasma membrane. The membranes of human erythrocytes contain about a dozen major proteins, which were originally identified by gel electrophoresis of membrane preparations. Most of these are peripheral membrane proteins that have been identified as components of the cortical cytoskeleton, which underlies the plasma membrane and determines cell shape. For example, the most abundant peripheral membrane protein of red blood cells is spectrin, which is the major cytoskeletal protein of erythrocytes the two major integral membrane proteins of red blood cells, glycophorin provide well-studied examples of transmembrane protein structure. Glycophorin is a small glycoprotein of 131 amino acids, with a molecular weight of about 30,000, half of which is protein and half carbohydrate. Glycophorin crosses the membrane with a single membrane-spanning á helix of 23 amino acids, with its glycosylated amino-terminal portion exposed on the cell surface. The other major transmembrane protein of red blood cells is well understood. This protein, originally known as band 3, is the anion transporter responsible for the passage of bicarbonate (HCO_3^-) and chloride (Cl^-) ions across the red blood cell membrane. The band 3 polypeptide chain is 929 amino acids and is thought to have 14 membrane-spanning á-helical regions. Within the membrane, dimers of band 3 form globular structures containing internal channels through which ions are able to travel across the lipid bilayer.

Check Your Progress

10. Define amphipathic molecules.
11. How are micelle aggregates formed?
12. Mention any two functions of lipid bilayer.
13. Why are sugars molecules used as markers in biological membrane?

3.5 LIPID METABOLISM: β-OXIDATION OF FATTY ACIDS

Lipids are any of a group of organic compounds that include fats, oils, waxes, sterols and triglycerides, which are insoluble in water. They are the important components of the plant and animal life on earth. Both plants and animals lipids are stored in large amounts as neutral, highly insoluble triacylglycerols. They can be rapidly mobilized and degraded to meet the cells' demands for energy. In the complete combustion of a typical fatty acid, palmitic acid, there is a large negative free energy change, which is shown by the following equation:

$$C_{16}H_{32}O_2 + 23O_2 \rightarrow 16CO_2 + 16H_2O$$

$$AG = -2340 \text{ kcal/mole}$$

NOTES

This negative change is due to the oxidation of the highly reduced hydrocarbon radical attached to the carboxyl group of the fatty acid. Of all the common food stuffs, only the long chain fatty acid possesses this important chemical feature. Thus, lipids have quantitatively the best caloric value of all foods, which is 9.3 kcal/gm compare to 4.1 kcal/gm for carbohydrates and proteins. Lipids are fatty or oily substances present in almost all organisms including viruses.

They have two important characteristics, which are as follows:

1. They are insoluble in water.
2. They contain a large proportion of carbon-hydrogen bonds and release a lot of energy on breakdown.

Lipids are soluble in non-polar organic solvents such as acetone, alcohol, chloroform, benzene and ether. The alkaline hydrolysis (saponification reaction) of lipids yields alcohol and fatty acids, which may be water soluble. Fat is the principal form in which lipids are used for storage. Cells synthesize fats from sugars. Chemically, lipids are composed of three fatty acids joined to a glycerol molecule. As with the polysaccharides, these bonds are formed by the removal of a molecule of water. Glycerol, therefore, serves as a binder or carrier for fatty acids. Lipids include, in general, compounds such as fats, oils and waxes. Besides, there are hybrid molecules such as glycolipids (carbohydrate: lipids) and lipo-proteins (proteins: lipids).

The chylomicrons are finally deposited either in the liver, or in the fat storage depots (adipose tissue) which in health from about 15 per cent of the body weight. Lipoprotein lipase in the adipose tissue is responsible for the clearance of chylomicrons. Plasma lipoprotein lipase is also responsible for the clearance of a small amount of chylomicrons in plasma. The triglycerides undergo hydrolysis by an intracellular lipase to form free fatty acids (FFA) and glycerol. The released free fatty acids are carried in the unesterified state in plasma as albumin-FFA complex. Many tissues such as liver, heart, Kidney, muscle, lung, testis, brain and adipose tissue, have the ability to oxidize long chain fatty acids by beta-oxidation. By this way, the long chain fatty acids are degraded completely to acetyl-CoA, which can be oxidized to CO_2 and water via-the citric acid cycle. The glycerol, which has been released by the hydrolysis of fat, enters general carbohydrate metabolism via glyceraldehydes and is either converted to glycogen or oxidized. The triglyceride stores in adipose tissue continually undergo lipolysis (hydrolysis) and re-esterification.

This is influenced by nutritional, metabolic and hormonal factors. When the availability of glucose in adipose tissue is reduced (as in starvation or diabetes mellitus) rate of lipolysis exceed the rate of esterification with subsequent accumulation of FFA and their release into plasma. Table 3.3 gives the summary of digestion

NOTES

Table 3.3 *Summary of Digestion*

Source of Enzyme and stimulus for secretion	Enzymes and other substances required for	Method of activation and substances required for	Substrate	End products
Pancreas: Secretin and pancreozymin stimulate the flow of pancreatic	Lipase	May be activated by bile salts pH 8.0	Primary ester linkage of lipids	Fatty acids, mono-glycerides, diglycerides, glycerol
Liver and gall bladder	Bile salts and alkali	Cholecystokinin and hepatocrinin hormones and intestine stimulate gall bladder and	Fats	Fatty acid - bile salt conjugates and finely emulisfied neutral fat
Small intestine: Enterocrimn stimulates the secretion of	Lecithinase		Lecithin	Glycerol, fatty acids, choline, phosphoric acid

Lipids function as important insulators of delicate internal organs. Nerve tissue, plasma membrane and membranes of subcellular particles such as mitochondria, endoplasmic reticulum and nuclei have complex lipids as essential components. In addition, the vital electron transport system in mitochondria and the intricate structures found in chloroplasts, the site of photosynthesis, contain lipids derivatives in their basic architecture. The chief storage form of available energy in the animal cell is the lipid molecule. When the caloric intake exceeds utilization excess food is invariably stored as fat, the body cannot store any other form of food in such a large amounts. Fat is a concentrated source of energy and occur widely as reserve material, particularly in higher plants; for example, in seeds. The products of their degradation can be again subjected to biological oxidation and thus be utilized for extra ATP production (in addition to that produced in the Krebs cycle). Besides, they can also be used, as in the glyoxylate cycle, for the synthesis of glucose.

β-Oxidation

All enzymes associated with the β-oxidation system are localized in the inner membranes and the matrix of liver and other tissue mitochondria. Since the inner membrane is also the site of the electron transport and oxidative phosphorylation systems, this arrangement is of fundamental importance to the efficient release and conservation of the potential energy stored in the long chain fatty acid. When acetyl-CoA is produced in the breakdown of fatty acids, it may be subsequently oxidized to CO_2 and H_2O by means of tricarboxylic acid cycle enzymes, which are localized as soluble enzymes in the matrix. An unusual property of liver and other tissue mitochondria is their inability to oxidize fatty acids or fatty acyl-CoA unless carnitine (3 hydroxy - 4 trimethyl ammonium butyrate) is added in catalytic

amounts. Evidently, free fatty acids or fatty acyl-CoA's cannot penetrate the inner membranes of liver and other tissue mitochondria, whereas acyl-carnitine readily passes through the membrane and is then converted to acyl-CoA in the matix. Figure 3.40 shows the translocation of acyl-CoA from outside the mitochondrion to internal site of the p-oxidation system. The key enzyme is carnitine acyl-CoA transferase.

NOTES

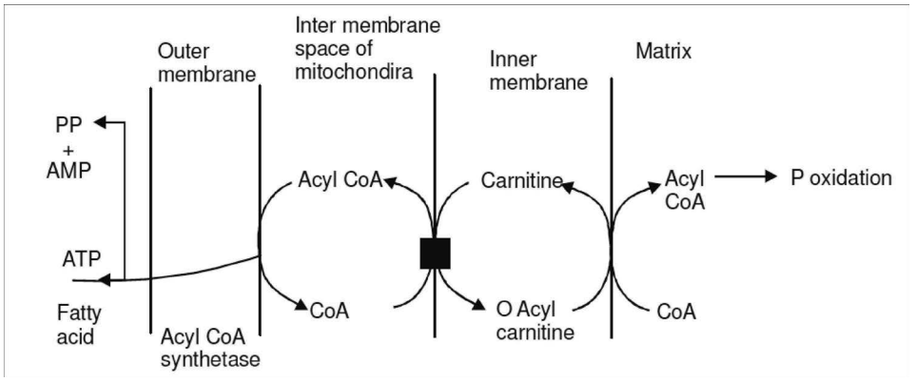


Fig. 3.40 *Transport Mechanism for Fatty Acids from the Cytosol to P-oxidation Site in the Mitochondria*

α-Oxidation

This system was first observed in seed and leaf tissues of plants and is also found in brain and liver cells. In this system, only free fatty acids serve as substrates and molecular oxygen is indirectly involved. The products may be either a D –á hydroxy fatty acid or a fatty acid containing one less carbon atom. This mechanism explains the occurrence of á-hydroxy fatty acids. The latter may, in nature, also be synthesized de novo from propionate. The á-oxidation system has been shown to play a key role in the capacity of mammalian tissue to oxidize phytanic acid, the oxidation product of phytol, to CO_2 and water.

Co-Oxidation

Microsome in hepatic cells rapidly catalyze the oxidation of hexanoic, octanoic, decanoic and lauric acids to corresponding dicarboxylic acids via a cytochrome P450 co-oxidation system. In addition, a number of aerobic bacteria have been isolated from oil soaked soil; these rapidly degrade hydrocarbons or fatty acids to water soluble products. The reaction involves an initial hydroxylation of the terminal methyl group to a primary alcohol and subsequent oxidation to carboxylic acid. Thus, straight chain hydrocarbons are oxidized to fatty acids and fatty acids in turn are p-oxidized to acetyl CoA. The mechanism of the oxidation of oils is primarily by the co-oxidation mechanism.

The overall β-oxidation is presented in Figure 3.41. One molecule of ATP is required to activate a fatty acid for its complete degradation to acetyl CoA, regardless of the number of carbon atoms in its hydrocarbon chain.

NOTES

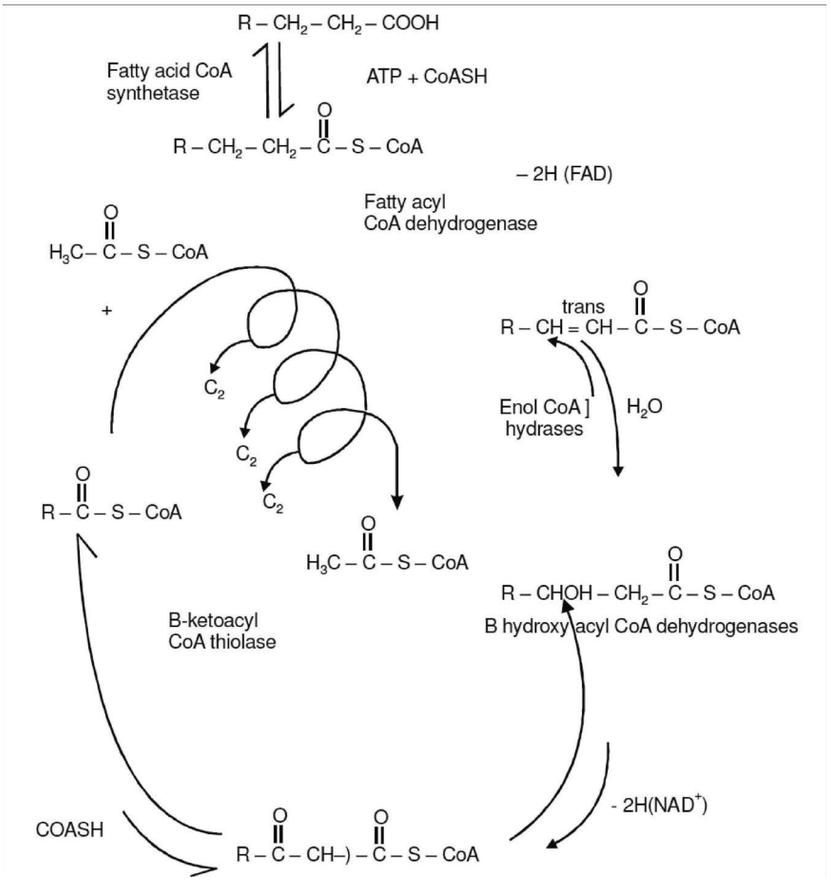


Fig. 3.41 *The p-Oxidation Helical Scheme*

Five enzymes are required for this cycle. These are discussed in following steps:

1. Formation of Acyl-S-CoA's by Acyl-CoA Synthetase

The overall type reaction is depicted as follows:

$$RCOOH + ATP + CoASH \xrightarrow{Mg^{++}} RCO - S - CoA + AMP + PPi$$

There is good evidence that the reaction actually takes place in the following two steps:

1. ROOCH + ATP + Enzyme $\longleftrightarrow$ E-acyl adenylate + PPi
2. Enzyme - acyladenylate + CoASH $\longleftrightarrow$ Enzyme + Acyl S-CoA + AMP

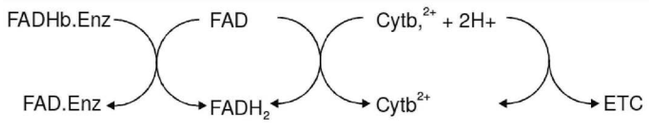
Three different synthetases occur in the cell; one activates acetate and propionate to the corresponding thioesters, another activates medium chain fatty acids from C4 to Cn and the third activates fatty acids from C10 to C20. The first two mentioned are located in the outer membranes of mitochondria and the third synthetase is associated with endoplasmic reticuli membranes.

2. α, α-Dehydrogenation of Acyl-CoA

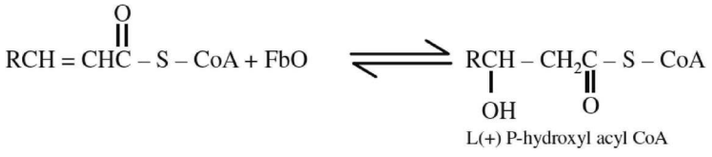
The degradation of fatty acids by oxidation at the α-carbon and beta-carbon was first made by Franz Knoop in 1904. However, the elucidation of the series of

reactions occurring during fl-oxidation has been described in detail by David Green, Severo Ochoa and Feodor Lynen. The initial step in the oxidation processes is a dehydrogenation where a double bond is formed beta to the site of attachment of CoA. This reaction is catalysed by an enzyme called fatty acyl-CoA dehydrogenase, which contains FAD as a prosthetic group. When this acyl-CoA is oxidized, FAD is simultaneously reduced and enoyl-CoA is formed. Reduced FAD, in turn, passes electrons to a second flavopnotein, then ubiquinone and, finally, through the remainder of the respiratory chain. The passage of the electron along the chain is coupled to the formation of ATP.

Three acyl-CoA dehydrogenases are found in the matrix of mitochondria. They all have FAD as a prosthetic group. The first has a specificity ranging from C_2 to Cb acyl-CoAs, the second from C & to C-4 and the third from Ce to Cis. The $FADH_2$ is not directly oxidized by oxygen but follows the path, which is as follows:



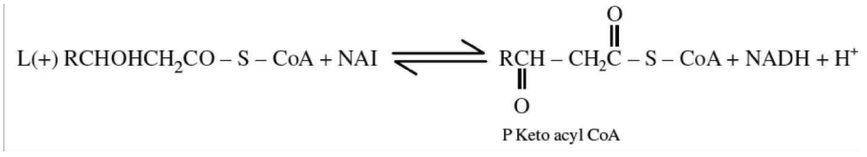
3. Hydration of α, β-unsaturated Acyl-CoA:



The enzyme enoyl-CoA hydrase, catalyzes this reaction; it possesses broad specificity.

4. Oxidation of β-Hydroxyl CoA

The second step in fatty acid oxidation is the addition of water across the double bond formed in the foregoing reaction. The reaction is catalyzed by the enzyme enolhydratase and the product formed is hydroxy-acyl-CoA. This is followed by a second dehydrogenation and the hydroxyacyl CoA is oxidized to the corresponding ketone. This time NAD+ is the hydrogen acceptor. As mentioned earlier, in the case also electrons removed from the fatty acid derivative are passed onto the respiratory chain with the help of NADH dehydrogenase and accompanying ATP synthesis. The â-carbon atom now bears a carbonyl function as shown below.



Finally, there is cleavage of the acyl-CoA molecule in the beta-position and transfer of the two carbon fragments to a second CoASH molecule. The reaction is catalyzed by an enzyme called thiolase and leaves one molecule of acetyl-CoA and the CoA thioester of a fatty acid. The fatty acid derivative produced is now shortened by two carbon units. This molecule is now able to undergo once again

NOTES

the cycle of dehydrogenation, hydration, oxidation and cleavage to yield a product shortened by two more carbons. This process is repeated till only acetyl-CoA is left over. This molecule of acetyl-CoA is also able to contribute towards energy production by entering the citric acid cycle through condensation with oxaloacetate. Stepwise reactions of the breakdown of the fatty acids are summarized in Table 3.4.

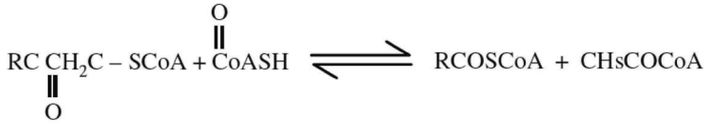
Table 3.4 *Breakdown of Fatty Acid*

Reaction	Enzyme	Site
Activation of fatty acid and formation of acetyl CoA	Thiokinase	
Formation of acyl carnitine and group transfer to CoASH	Carnitine fatty acid transferase	Outside (formation of acyl carnitine) and inside (group transfer) mitochondria
Dehydrogenation and formation of enmyl-CoA FAD acceptor	Fatty acyl-CoA dehydrogenase	Inside mitochondria
Hydration and formation of hydroxyacyl CoA	Enolhydratase	Inside mitochondria
Dehydrogenation and formation of keto acyl-COA (NAD+acceptor)	Hydroxyacyl CoA dehydrogenase	Inside mitchondria
Cleavage and formation of CoA thioester of fatty acid.	Thiolase	Inside mitochondria

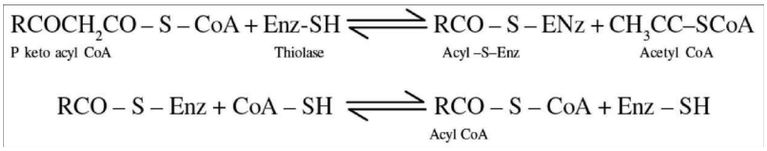
A broadly specific L p-hydroxyacyl-CoA dehydrogenase catalyzes this reaction and it is specific for the L form. Some variations are recorded for the oxidation of unsaturated fatty acids, and, besides the various enzymes described above, two additional enzymes including isomerase and epimerase are required for the degradation of unsaturated fatty acids. Likewise, the degradative pathway of odd chain fatty acids is also with minor modifications in the thiolysis step. The products of thiolysis here are propionyl CoA and acetyl-CoA. The propionyl-CoA is converted to succinyl-CoA and then enters the Krebs cycle. If there is excess production of acetyl coenzyme A due to continuous fat oxidation and unavailability of sufficient oxaloacetate (to allow acetyl-CoA to enter the Krebs cycle), then acetyl- CoA breaks down to liberate ketone bodies including acctoacctate, D-3 hydroxybutyrate and acetone. The conversion is catalyzed by several enzymes.

5. Thiolysis

The enzyme thiolase carries out a thiolytic cleavage of the p-ketoacyl CoA, as shown below.



The enzyme protein has a reactive SH group on a cysteinyl residue that is involved in the following series of reactions:



NOTES

Energy Released by α–Oxidation

The energy released by fat oxidation is quite significant. In general, three molecules of ATP are formed with each NAD+-linked oxidation and two with each flavoprotein linked oxidation. As explained above, in every reaction completed, one molecule each of FADH, NADH and acetyl-CoA are formed. Thus, there are five molecules of ATP for each acetyl-CoA produced. In addition, the oxidation of acetyl-CoA by way of Krebs cycle produces an additional 12 ATP's per acetyl-CoA, thus totalling 17 ATP molecules.

A typical fatty acid may have a total of 16 to 20 carbons. Therefore, a total of about 136–170 ATP molecules may be produced. The net gain from the oxidation of palmitic acid, for example, is 129 ATP. Energy conservation in fatty acid (undergoing oxidation) turns out to be about 40 per cent. In the total combustion of palmitic acid, considerable energy is released, which is shown below:

$$C_6H_{32}O_2 + 23O_2 \rightarrow 16CO_2 + 16H_2O$$

$$AG' = -2340 Kcal/mole$$

Palmitic acid

$$C_5H_{32}COOH + 8COASH + ATP + 7FAD + 7NAD^+ 7H_2O$$

$$\downarrow$$

$$8CH3CO \sim SCoA + AMP + PPi + 7FADH_2 + 7NADH + H^+ \text{ Acetyl-CoA}$$

$$8CH_3CO \sim SCoA + I6O_2 \xrightarrow[\text{Cycle}]{\text{TCA}} I6CO_2 + I6H_2O + 8C0ASH$$

When palmitic acid is degraded enzymatically, one ATP is required for the primary activation and eight energy rich acetyl-CoA thioesters are formed. Each time the helical cycle is traversed, 1 mole of FAD-H_2 and 1 mole of NADH are formed. They may be re-oxidized by the electron transport chain. Since in the final turn of the helix, 2 moles of acetyl-CoA are produced, the helical scheme must be traversed only 7 times to degrade palmitic acid completely. In this process, 7 moles each of reduced flavin and pyridine nucleotide are formed. The sequence can be divided into two steps.

Step 1:

Palmitic acid $\longrightarrow$ P acetyl SCoA + 14 e pair

7 e pairs via flavin system at 2 ~ P/one electron pair = 14 ~ P

7 e- pairs via NAD^H system at 3 ~ P/one electron pair = 21 ~ P

Total = 35 ~ P

Net = 35 ~ P - 1 ~ P

= 34 ~ P

NOTES

Step 2:

$$\beta \text{ Acetyl-CoA} + 16O_2 \xrightarrow[\text{Cycle}]{\text{TCA}} 16CO_2 + 16H_2O + 8 \text{ CoA - SH}$$

Assuming that for each oxygen atom consumed 3 ~ P are formed during oxidative phosphorylation then 32 x 3 = 96 ~ P.

(96 ~ P) = 130 ~ P and 130 × 8000 × 100/2,340,00 = 48 percent

In the complete oxidation of palmitic acid to CO_2 and H_2O, 48 per cent of the available energy can theoretically be conserved in a form (ATP) that is utilized by the cell for work. The remaining energy is lost, probably as heat. It, hence, becomes clear why as a food fat is an effective source of available energy. This is quite close to the values obtained from oxidative phosphorylation and glycolysis. Consequently, the breakdown of fat (which is a storage product in several organisms and seeds) is certainly an additional source of energy to drive the cellular biochemistry.

Fatty Acid Oxidation and Energy Coupling

Many eukaryotic cells utilize fatty acid oxidation as a major route for energy coupling and many store fats in large amounts as an energy source. Glycerol (as shown below) is an example of a polyhydroxy (trihydric) alcohol which can esterify with fatty acids. Fatty acids are organic acids with alkyl side chains of different lengths and variably unsaturated (with double bonds):

CH_2OH
|
CHOH
|
CH.OH
Glycerol

The degradation of fat begins with the breakdown of the neutral fat into glycerol and fatty acids. The reaction is catalyzed by lipases and proceeds with hydrolysis of triglycerides, as shown below.

$$\begin{array}{l} CH_2OOC\,(CH_2)\,CH_5 \\ CH\text{-}OOC\,(CH_2).CH_3 \\ CH.OOC\,(CH\,\hat{}\,CH_3) \end{array} \xrightarrow[\text{Lipases}]{+H_2O} \begin{array}{l} CHOH \\ CHOH + 3\text{-}OOC\,(CH_3)\,CH_2 \\ CH_3OH \end{array}$$

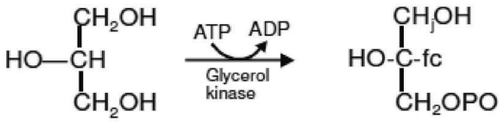
The fatty acids released are immediately subjected to oxidation. This is very important for the cell also, because the free fatty acids produce a number of toxic effects when their intracellular concentration becomes too high. They are uncouples of oxidative phosphorylation, and the high concentration of fatty acids leads to a diminished level of energy coupling. Fatty acids are released in the cytoplasm and are oxidized in the mitochondria. This was discovered by Eugene Kennedy and Albert Lehninger in 1949.

Glycerol Kinase Reaction

The glycerol produced, as a result of the action of lipases on the fat molecule, can be converted to triose phosphate which, in turn, can either be degraded by glycolysis or used for the synthesis of hexoses, as show below:

glycerol + ATP T ⟵⟶ glycerol phosphate + ADP

The glycerol is phosphorylated before entering the glycolytic pathway and the reaction is catalyzed by glycerol kinase:



Formation of Ketone Bodies

Free Fatty Acids (FFA) enter liver cells from chylomicrons and from FFA - albumin complexes originating from adipose tissue fat cells. Fatty acids formed de novo from glucose in the liver also are major contributors to this dynamic pool. Under normal nutritional conditions, these fatty acids have several fates. Figure 3.42 shows the route of fatty acids in animals.

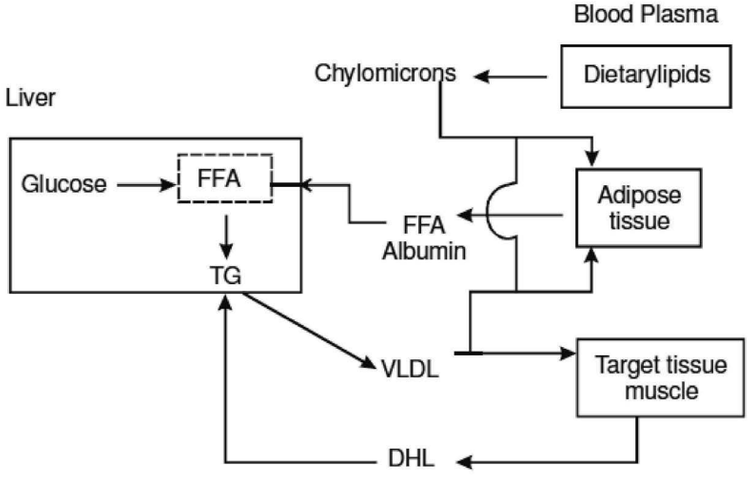


Fig. 3.42 *Route of Fatty Acids in Animals*

The acids are etherified to triacylglycerols. However, the liver has a limited capacity for triacylglycerol storage, any excess combines with HDL, cholesterol esters and phospholipids to form VLDL particles. These are now excreted into the blood system and are transported via the vascular system to target tissues such as muscle and adipose tissue. Here, lipoprotein lipase removes and converts triacylglycerol to free fatty acids, which are then absorbed by trie tissues and utilized. The residual VLDL particles in the meantime convert to HDL particles that presumably return to the liver via the blood system to pick up excess triacylglycerols and repeat the cycle.

The free fatty acids enter the mitochondria to be 3- oxidized and then converted by the TCA cycle to CO_2 and H_2O. The free fatty acids are converted to ketone bodies in the mitochondrion and then transported from the liver to target tissues like red muscle, brain and cardiac muscles to be burned to CO_2 and H_2O. Recent evidence strongly suggest that ketone bodies are major fuels for peripheral muscles and become important sources of energy in muscles involved in prolonged muscle exercises such as long distance running, etc.

NOTES

Ketone bodies are D-cc hydroxy butyric acid, acetoacetic acid and acetone. They are formed by a series of unique reactions, primarily in the liver and kidney mitochondria. The biosynthesis of the acids is summarized. The enzymes involved in the synthesis of ketone bodies are localized primarily in liver and kidney mitochondria. Ketone bodies cannot be utilized in the liver since the key utilizing enzyme 3-oxoacid-CoA transferase is absent in the tissue but is present in all tissue metabolizing ketone bodies, namely red muscle, cardiac muscle, brain and kidney.

Thus, ketone bodies are alternative substrates to glucose for energy sources in muscle and brain. The precursor of ketone bodies, namely free fatty acids are toxic in high concentrations, have a very little solubility and readily saturate the carrying capacity of the plasma albumin. Ketone bodies, on the other hand, are very soluble, low in toxicity, tolerated at high concentrations, diffuse rapidly through membranes and are rapidly metabolized to CO_2 and H_2O.

Ketosis

The fatty acids undergo excessive oxidation in liver under certain metabolic conditions producing large quantities of keto acids - acetoacetic acid and b-hydrobutyric acid, which pass into the blood by diffusion. Acetoacetic acid then undergoes spontaneous decarboxylation to produce acetone. These three substances - acetoacetate p - hydroxybutyrate and acetone are collectively known as ketone bodies (acetone bodies or ketones).

Normally, the blood of mammals contains ketone bodies not exceeding 1 mg/100 ml. The concentration is little higher than this in ruminants. Daily excretion of ketone bodies of normal person is less than 1 mg higher than normal quantities in the blood or urine constitute ketonemia or ketonuria respectively. The conditions in which there is a high concentration of ketone bodies in tissues and blood is called ketosis. Acetoacetic acid and p-hydroxybutyric acid are moderately strong acids and are buffered in blood or tissues. Their excretion in large quantities admits some loss of buffer cation, which depletes the alkali reserve causing ketoacidosis.

Ketosis generally occurs in severe diabetes mellitus prolonged starvation, glycogen storage disease, toxemia of pregnancy, infective hepatic disease and continued fever. Experimentally, it occurs in the oral administration of fatty acids, high fat diet, low carbohydrate diet, pancreatectomy and administration of growth hormone or ACTH. Under these conditions, there is diminished carbohydrate utilization and increased fat mobilization.

Effects of Ketosis

Both acetoacetate and P-hydroxybutyrate are moderately strong acids. They neutralize bicarbonates resulting depletion of alkali of the body and produce metabolic acidosis. In case of severe ketosis, death may ensure from acidosis. The excretion of ketone body in the urine involves the loss of Na^+ in particular leading to total electrolyte and Na^+ deficiency. The severe diabetic patient excretes large quantities of both ketone bodies and glucose in the urine with a large quantity of water developing dehydration. In diabetic acidosis, there is severe alteration in cationanion balance in the plasma.

Prevention of Ketosis

In case of diabetic ketosis, carbohydrate diet intramuscular infection of insulin and antiketogenic substance (aspartic acid) which may provide oxaloacetate by transamination.

In case of prolonged starvation ketosis, carbohydrate diet and antiketogenic substance which may provide oxaloacetate by transamination should be given. The electrolytes and fluids of the body must be restored by intravenous injection of isotonic solution of sodium salts such as NaCl, $NaHCO_3$ or sodium lactate.

3.5.1 Influence of Hormones in Lipid Metabolism

Insulin administration is followed by a fall in circulating plasma FFA. It enhances lipogenesis and increases the oxidation of glucose through hexose monophosphate shunt. Both glucose oxidation and lipogenesis are reduced in diabetes mellitus. Other hormones, such as Adrenocorticotropic Hormone (ACTH), Growth Hormone (GH), Thyroid Stimulating Hormone (TSH), epinephrine, nor-epinephrine and glucagon raise the plasma FFA level by increasing the rate of lipolysis of the triglyceride store.

Phospholipid synthesis and degradation occur within each cell. Most of the lecithin in the plasma is derived from synthesis in the liver, and the normal plasma phospholipid concentration is 150-250 mg/dl as lecithin. Lecithin is synthesized from phosphatidic acid by the transfer of a choline group from a nucleotide, Cytidine Diphosphate Choline (CDP-Choline) in the presence of a transferase enzyme and megnesium ions. Other phospholipids are synthesized by similar reactions.

Cholesterol metabolism

About 1.5 to 2 g of cholesterol is synthesized by body and about 0.3 g/day is provided by the average diet. Cholesterol is eliminated by two main pathways (a) conversion to bile acids and (b) by excretion of neutral steroids in the faeces. The liver is the main site of cholesterol synthesis and other tissues known to be capable of synthesizing cholesterol include the testis, aorta, skin, intestine and the adrenal cortex. Acetyl-CoA is the source of all the carbon atoms in cholesterol.

Biosynthesis of cholesterol

Major site of cholesterol synthesis is liver (Denovo synthesis). Other sites are adrenal cortex, testis, ovaries and intestine. Liver is responsible for 80 per cent of endogenous cholesterol synthesis. Enzymes involved are partly located in endoplasmic reticulum and partly in cytoplasm

The process has five major steps: (1) Acetyl - CoA's are converted to 3-hydroxy-3- methyl glutaryl-CoA (HMG-CoA). HMG-CoA is converted to mevalonate (3) Mevalonate is converted to isopentenyl pyrophosphate (IPP), with the concomitant loss of CO_2 (4) IPP is converted to squalene (5) Squalene is converted to cholesterol (Figure 3.43).

NOTES

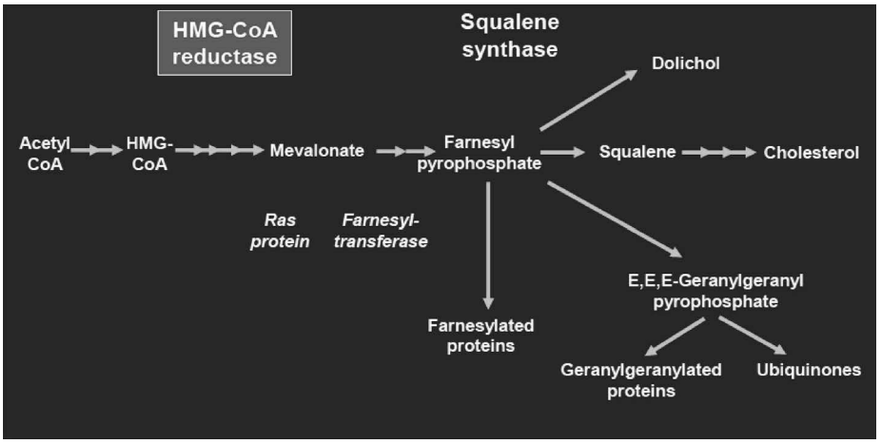


Fig. 3.43 *Biosynthetic Pathway of Cholesterol*

The first steps involve the synthesis of the important intermediate Mevalonic Acid from acetyl-CoA and acetoacetyl-CoA, in two enzymatic steps. The first enzyme, hydroxymethyl-glutaryl (HMG) - CoA synthase. The second enzyme HMG-CoA reductase is the rate-limiting step in the overall synthesis of sterols. Enzymes are membrane -bound and are located in the endoplasmic reticulum. The next sequence of reactions involves phosphorylation of mevalonic acid by mevalonate kinase to form the 5-monophosphate ester. Followed by a further phosphorylation to yield an unstable pyrophosphate, which is rapidly decarboxylated to produce 5-Isopentenyl Pyrophosphate (IPP). An isomerase converts part of IPP to 3, 3-dimethylallyl pyrophosphate.

IPP and 3, 3-dimethylallyl pyrophosphate condense readily with the elimination of pyrophosphoric acid to form Geranylpyrophosphate. This reacts with another molecule of 5-isopentenyl pyrophosphate to produce Farnesyl Pyrophosphate. Two molecules of farnesyl pyrophosphate, condense to yield Presqualene Pyrophosphate. The last is reduced by squalene synthase and Nadph to produce a further key intermediate Squalene. Squalene is first oxidized by squalene mono oxygenase to Squalene 2, 3 Epoxide. Squalene 2, 3-epoxide undergoes cyclization catalyzed by squalene epoxide lanosterolcyclase to form the first steroidal intermediate Lanosterol.

In this remarkable reaction, there is a series of 1, 2- methyl group and hydride shifts along the chain of the squalene molecule to bring about the formation of the four rings.

Finally, lanosterol is converted to cholesterol by multiple reactions that involve the removal of three methyl groups, hydrogenation of the double bond in the side-chain, and a shift of the double bond from position 8, 9 to 5, 6 in ring B.

Cholesterol can be exported and transported to other tissues in the form of lipoprotein complexes for uptake via LDL receptors. The enzyme Lecithin Cholesterol Acyl Tranferase (LCAT) are responsible for transport and elimination of cholesterol from the body. LCAT is synthesized by the liver. Degradation of Cholesterol leads to synthesis of bile acids, steroid hormones and vitamin D.

Regulation of Cholesterol Synthesis

NOTES

Regulatory enzyme is HMG – CoA reductase (HMGR) which has a feedback regulation by cholesterol amount of dietary cholesterol determines the cholesterol synthesis. Short term regulation is by covalent modification of the enzyme. Insulin and T3 increase activity of HMGR, whereas, Cortisol and glucagon decreases the activity of HMGR. Regulation of HMGR activity is the primary means for controlling the level of cholesterol biosynthesis. The enzyme is controlled by three distinct mechanisms: control of gene expression, rate of enzyme degradation and phosphorylation-dephosphorylation.

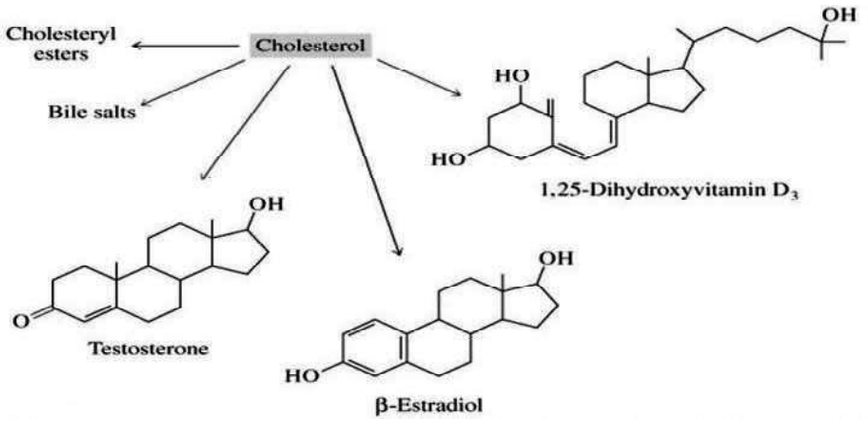


Fig. 3.44 *Cholesterol Metabolism*

Clinical Significance of Cholesterol

Although most of us are more or less familiar with the term 'cholesterol', the world of sterols is far more complicated and interesting. Apart from cholesterol, many noncholesterol sterols can be found in human plasma and these sterols serve many important functions in human organism. They are either derived from endogenous biosynthesis of cholesterol or they come from dietary sources (phytosterols). The sole cholesterol molecule is used for keeping our cell membranes fit, for signalization purposes as well as a precursor for bile acids and steroid hormones. The compounds prior to cholesterol in its biosynthetic pathway were identified as vitamin D3 precursor, meiosis activating sterols and nowadays, it seems that they could play a role in cholesterol homeostasis. The sterols from ingested vegetable sources, the phytosterols, are expelled from enterocytes and thus, indirectly help our gut in coping with abundant cholesterol in the lumen.

Higher plants synthesize many phytosterols, but in marine organisms, you can find other innumerous sterol molecules. The diversity of sterol molecules produced and resistance of their tetracyclic core to enzymatic activities implies crucial importance of sterols during the ontogenesis of multicellular organisms. First oxygen appeared on the earth approximately 3.7 billion years ago and since that time, every new life form took the advantage of oxygen needed also for build-up of sterol molecules. The last decades changed our view to the sterol molecules on almost at all levels of their appearance in human body. In the gut, the absorption of sterols was proven to be protein dependent and the quest for the transporter

NOTES

was successful. The general concepts of intracellular homeostasis of cholesterol have been described including the covalent interaction unbelievable so far - cholesterol and a protein. The clinical importance of non-cholesterol sterols rises with the effort to discover underlying facts about the causes of atherosclerosis.

3.5.2 Role of Liver in Lipid Metabolism

Liver plays the following roles in the lipid metabolism:

- Synthesis and oxidation of fatty acids
- Synthesis of triglycerides
- Synthesis of cholesterol and its derivative such as bile salts
- Synthesis of phospholipids
- Synthesis of VLDL and HLD (plasma lipoproteins)
- Formation of ketone bodies

Catabolism of Chylomicrons

Apoprotein B-48 is required for the assembly and secretion of chylomicrons from the intestine. Chylomicrons are responsible for the transport of all dietary lipids into the circulation. The clearance of chylomicrons is rapid. It is carried out by the enzyme lipoprotein lipase (LPL) in the walls of blood capillaries. Both phospholipids and apoprotein C-II are required for lipoprotein lipase activity. The action of LPL results in the loss of 90 per cent of triglycerides. The resultant free fatty acids are taken up by the extrahepatic tissue. The remnant chylomicrons are taken up by liver. Cholesterol esters and triglyceride core gets hydrolysed and metabolized in the liver.

Catabolism of VLDL

Apolipoprotein B-100 is responsible for the assembly of and secretion of VLDL by liver. Smaller and denser particles having the physical characteristics of VLDL are also found in chyle. The intermediate forms formed by the action of LPL are IDL and LDL. The released free fatty acids are taken up by extra hepatic cells.

Catabolism of LDL

Fibroblasts, lymphocytes, arterial smooth muscle cells and liver have special receptors for LDL. Approximately 50 per cent of LDL is degraded in extrahepatic tissue and 50 percent in the liver.

Metabolism of HDL

HDL is synthesized and secreted from both liver and intestine. Apolipoprotein (apo) C and E are synthesized in the liver and transferred to intestinal HDL. HDL acts as a repository for apo C and E. The plasma enzyme Lecithin Cholesterol-Acyl Transferase (LCAT) converts surface phospholipids into lysolecithin and cholesterol into esterified cholesterol. Lysolecithin is transferred to plasma albumin. Apo D transfers esterified cholesterol to chylomicrons and VLDL. Through the chylomicron remnants and LDL, the esterified cholesterol is taken up by the hepatic cells and these can be converted to bile salts. Thus, the excess unesterified

cholesterol is removed from the circulation by the aforementioned mechanism involving LCAT. Apo E in the presence of LCAT is also responsible for the reverse cholesterol transport from extrahepatic tissue to liver through HDL. The liver seems to be the final site of degradation of HDL apoproteins. The net effect of the maturation of HDL is the removal and transport of cholesterol and phospholipids from peripheral cells back to liver for eventual excretion.

3.5.3 Biosynthesis of Fatty Acids

Fatty acid synthesis is the process of creating fatty acids from acetyl-CoA and malonyl-CoA precursors through the action of enzymes called fatty acid synthases. It is an important part of the lipogenesis process, which together with glycolysis stands behind creating fats from blood sugar in living organisms.

The synthesis of fatty acid involves the following steps :

1. Two-carbon units derived from acetyl-CoA are added to form fatty acid chains.
2. Malonyl-CoA is formed, resulting in activation of the acetate units.
3. Due to decarboxylation of malonyl-CoA, two-carbon units are added to the growing chain. These elongation reactions get repeated until the growing chain reaches 16 carbons in length (palmitic acid).
4. Double bonds and additional carbon units are added to the chain by other enzymes.

Fatty Acid Synthesis in Animals

Eukaryotic cells are sometimes unable to provide suitable amounts of substrate for fatty acid synthesis. For fatty acid synthesis, sufficient quantities of acetyl-CoA, malonyl-CoA, and NADPH must be generated in the cytosol. Since Malonyl-CoA is made by carboxylation of acetyl-CoA, the problem relates to generation of only sufficient acetyl-CoA and NADPH. The main sources of acetyl-CoA are amino acid degradation, fatty acid oxidation and glycolysis.

Generally, the acetyl-CoA derived from amino acid degradation is insufficient for fatty acid biosynthesis. Also the acetyl-CoA produced by pyruvate dehydrogenase and by fatty acid oxidation is unable to participate directly in fatty acid synthesis as it cannot cross the mitochondrial membrane. Instead, acetyl-CoA gets linked with oxaloacetate and forms citrate, which is transported from the mitochondrial matrix to the cytosol. In the cytosol, the citrate can be converted back into acetyl- CoA and oxaloacetate by ATP-citrate lyase. This makes the mitochondrial acetyl-CoA the substrate for cytosolic fatty acid synthesis. Oxaloacetate returns to the mitochondria in the form of either pyruvate or malate, which is then reconverted to acetyl-CoA and oxaloacetate, respectively. Figure 3.45 shows the various pathways for fatty acid synthesis in eukaryotic cells.

NOTES

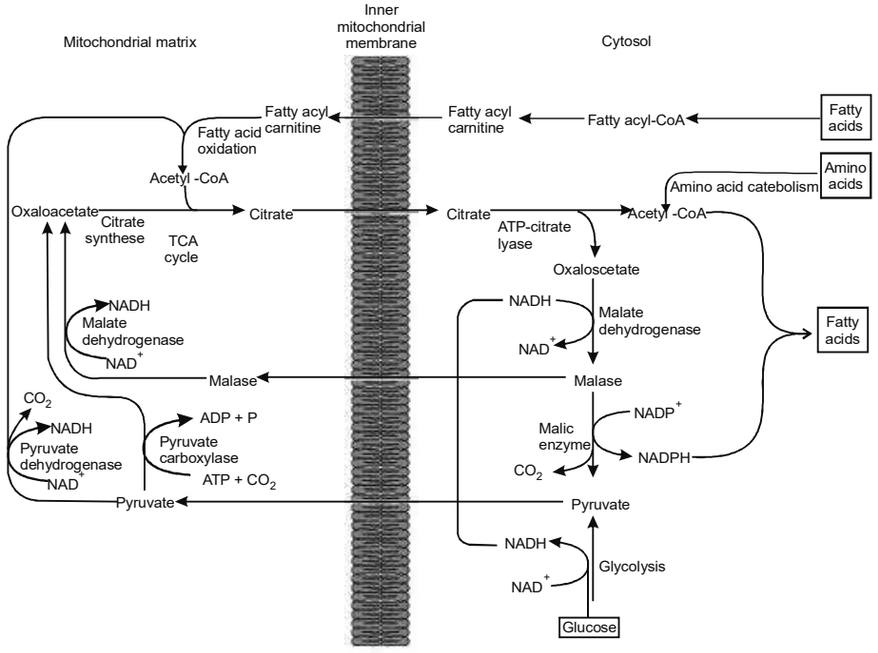


Fig. 3.45 *Pathways for Fatty Acid Synthesis in Eukaryotic Cells*

As shown in Figure 3.45, NADPH can be produced by the malic enzyme as well as in the pentose phosphate pathway. Using the combined action of malate dehydrogenase and malic enzyme, it is also possible to transform the reducing equivalents (electrons) derived from glycolysis in the form of NADH into NADPH.

$$\text{Oxaloacetate} + \text{NADH} + \text{H}^+ \rightarrow \text{malate} + \text{NAD}^+$$

$$\text{Malate} + \text{NADP}^+ \rightarrow \text{-pyruvate} + \text{CO}_2 + \text{NADPH} + \text{H}^+$$

Depending on the status of malate, the number of NADPH that can be made this way varies. Each citrate, which enters the cytosol, produces one malate and one acetyl-CoA. Each malate, which is oxidized by malic enzyme, produces one NADPH, at the expense of a decarboxylation to pyruvate. This means that when a malate is oxidized, for each acetyl-CoA, one NADPH is produced. This way, conversion of eight acetyl-CoA units to one palmitate involves production of eight NADPH. The other six NADPH, of the total 14 required, are provided by the pentose phosphate pathway.

In the late 1940s, Rittenberg and Bloch showed that acetate units are the building blocks of fatty acids. Later on, as it was discovered that bicarbonate is required for fatty acid biosynthesis, it became clear that the carboxylation of acetyl-CoA to form malonyl-CoA is a committed step in the synthesis of fatty acids. The reaction is then catalyzed by acetyl-CoA carboxylase, which contains a biotin prosthetic group and is not part of fatty acid synthase, the multi-enzyme complex in animals. Figure 3.46 shows the steps involved in acetyl-CoA carboxylase reaction.

NOTES

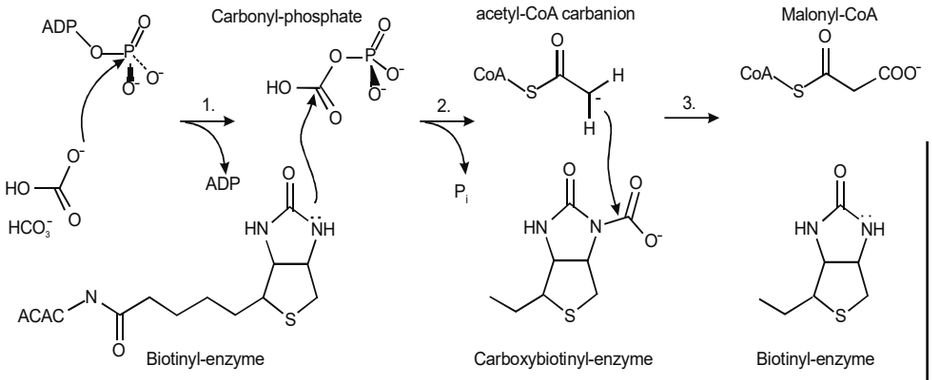


Fig. 3.46 *Acetyl-CoA Carboxylase Reaction*

The pyruvate carboxylase and the biotin prosthetic group of acetyl-CoA carboxylase are covalently linked to the α-amino group of an active-site lysine, similarly. The reaction mechanism of the biotin prosthetic group is also analogous to that of pyruvate carboxylase, that is, ATP-driven carboxylation of biotin is followed by transfer of the activated CO_2 to acetyl-CoA to form malonyl-CoA.

The enzyme from Escherichia coli has three subunits; a biotin carboxyl carrier protein, biotin carboxylase, which adds CO_2 to the prosthetic group and transcarboxylase, which transfers the activated CO_2 unit to acetyl-CoA. The activated carboxyl group can be transported between the biotin carboxylase and the transcarboxylase using the long, flexible biotin-lysine chain (biocytin). Figure 3.47 shows the biotin ring as part of the acetyl-CoA carboxylase reaction.

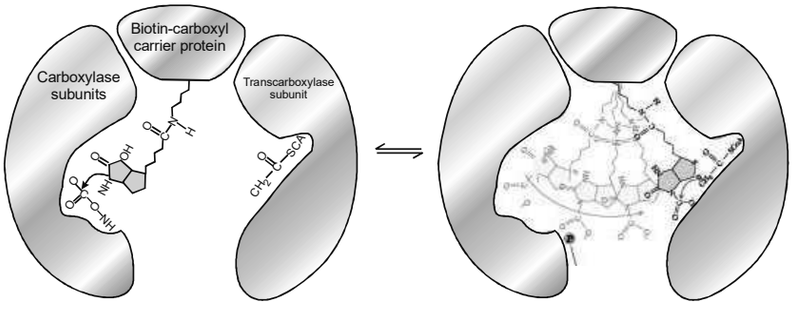


Fig. 3.47 *The Biotin Ring in the Acetyl-CoA Carboxylase Reaction*

The basic building blocks of fatty acid synthesis, acetyl and malonyl groups are not able to reach the fatty acid chain directly from CoA. They first reach the Acyl Carrier Protein (ACP). In *E. coli*, this protein consists of a single polypeptide chain of 77 residues. A phosphopante-theine group is attached to this chain. This is the same group that forms the business end of coenzyme A. Thus, ACP is a somewhat larger version of coenzyme A, specially used in fatty acid biosynthesis. Figure 3.48 highlights the conjugation of fatty acids, with both coenzyme A and ACP.

NOTES

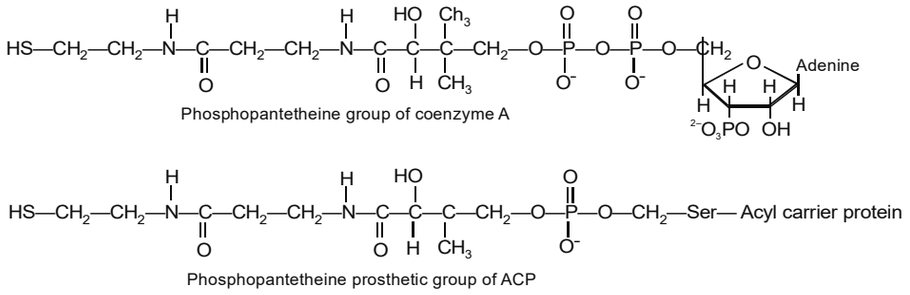


Fig. 3.48 *Conjugation of Fatty Acids with both Coenzyme A and ACP*

Different organisms have different organization of the enzymes that catalyze formation of acetyl-ACP and malonyl-ACP and the subsequent reactions of fatty acid synthesis. Figure 3.49 shows the pathway of palmitate synthesis from acetyl-CoA and malonyl-CoA.

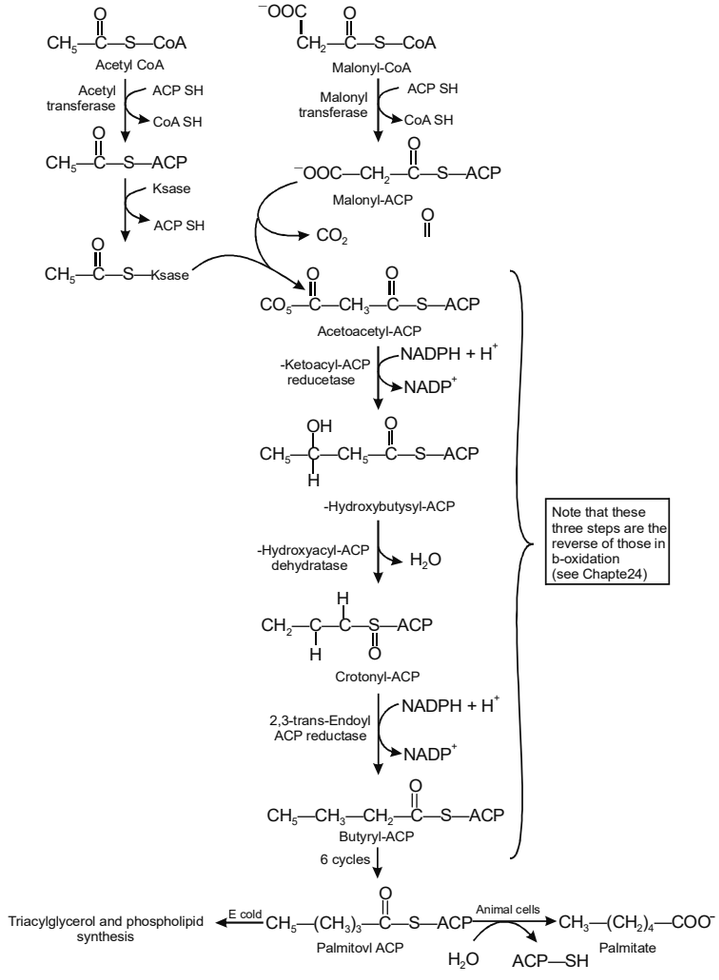


Fig. 3.49 *Pathway of Palmitate Synthesis from Acetyl-CoA and Malonyl-CoA*

Sometimes, the degradation and the synthesis of a class of biomolecules are carried out in diffrent ways. For example, glycolysis vs gluconeogenesis and glycogen or starch breakdown vs polysaccharide synthesis. This holds true for fatty acids and other lipid components as well. The strategies used by fatty acid

synthesis reactions are quite different from the corresponding degradative process.The primary differences between fatty acid synthesis and breakdown, are as follows:

1. While the intermediates in breakdown are bound to the -SH group of coenzyme A, the intermediates in synthesis are linked covalently to the sulfhydryl groups of special proteins, that is , the acyl carrier proteins.
2. While the synthesis occurs in the cytosol, degradation takes place in mitochondria.
3. In animals, while the enzymes of synthesis are part of a long polypeptide chain, called the fatty acid synthase, the degradative enzymes have no such association.
4. While the oxidation-reduction reactions of synthesis involve the coenzyme, $NADP^+$/NADPH, degradation involves the NAD^+/NADH couple.

Fatty Acid Synthesis in Bacteria and Plants

The steps involved in the elongation of the fatty acid chain are quite similar in bacteria, fungi, plants and animals. The formation of acetyl-ACP and malonyl-ACP initiate the elongation reactions. Acetyl-ACP and malonyl-ACP are formed by acetyl transacylase (acetyl transferase) and malonyl transacylase (malonyl transferase), respectively. While the acetyl transacylase enzyme is not highly specific as it can transfer other acyl groups, such as the propionyl group, malonyl transacylase is highly specific.

The acetyl group from ACP gets transferred to β-keto-acyl-ACP synthase (KSase), also known as acyl-malonyl-ACP condensing enzyme by another transacylase reaction. The first actual elongation reaction involves the condensation of acetyl-ACP and malonyl-ACP by the β-ketoacyl-ACP synthase to form acetoacetyl-ACP. The decarboxylation that accompanies the reaction with malonyl-ACP drives the synthesis of acetoacetyl-ACP. ATP is responsible for the condensation reaction to form acetoacetyl-ACP. Malonyl-CoA can be viewed as a form of stored energy for driving fatty acid synthesis and it can be established that all the carbons of acetoacetyl-ACP are derived from acetate units of acetyl-CoA.

The biosynthetic cycle finally results in the synthesis of a four-carbon unit, a butyryl group, from two smaller building blocks. This butyryl-ACP condenses with another malonyl-ACP, in the next cycle of the process, making a six-carbon β-ketoacyl-ACP and CO_2. A six-carbon saturated acyl-ACP is yielded by a subsequent reduction to a β-alcohol, dehydration and another reduction. Until the chain is 16 carbons long, this cycle continues with the net addition of a two-carbon units in each turn. Hydrolysis of the C_{16}-acyl-ACP yields a palmitic acid and the free ACP.

In the end, seven malonyl-CoA molecules and one acetyl-CoA yield a palmitate (shown here as palmitoyl-CoA):

$$\text{Acetyl-CoA} + 7\ \text{malonyl-CoA}^2 + 14\ \text{NADPH} + 14\ H^+ \rightarrow \text{palmitoyl-CoA} + 7\ HCO_3^2 + 14\ NADP^+ + 7\ \text{CoASH}$$

NOTES

The formation of seven malonyl-CoA molecules requires:

$$7\ \text{Acetyl-CoA} + 7\ HCO_3^{2} + 7\ ATP^{42} \rightarrow$$
$$7\ \text{malonyl-CoA}^{2} + 7\ ADP^{32} + 7\ P_i^{22} + 7\ H^{1}$$

Thus, the overall reaction of acetyl-CoA to yield palmitic acid is:

$$\text{palmitoyl-CoA} + 14\ NADP^{+} + 7\ \text{CoASH} + 7\ ADP^{32} + 7P_i^{22}$$

Thus, it can be said that the fatty acid synthase complex contains activities of seven different enzymes and an acyl carrier protein (ACP). In summary, the following reactions occur in FAS:

1. The two carbons of acyl CoA are transferred to ACP of FAS complex by the enzyme acetyl CoA –ACP transacylase. The acetyl group attached to ACP is transferred to cysteine.
2. The malonyl group of malonyl CoA is transferred to ACP by enzyme malonyl CoA – ACP transacylase
3. The enzyme β –ketoacyl ACP synthase transfers acetyl group from cysteine to malonyl, attached to ACP with loss of one CO2. This results in the synthesis of β-ketoacyl ACP.
4. The β –ketoacyl ACP is converted to β-hydroxyacyl- ACP by reduction reaction in the presence of β-ketoyacyl –ACP reductase. The reducing equivalent is NADPH.
5. A dehydration reaction of β-hydroxyacyl- ACP introduces a double between á and β carbons. This reaction is catalysed by β-hydroxyacyl- ACP dehydratase. This results in the formation of Trans e^{2} –Enoyl CoA.
6. The enzyme enoyl CoA reductase reduces Trans Δ^{2} –Enoyl CoA to Acyl – ACP(butyryl-ACP). The reduction is done with the help of NADPH.
7. The steps repeat till all the 16 carbons of palmitate are added.
8. The enzyme palmitoyl thioesterase separates palmitate from fatty acid synthase.

Additional Elongation

Though the primary product of the fatty acid synthase is palmitate, cells synthesize many other fatty acids. If the chain is released before reaching 16 carbons in length, many shorter chains can be easily made. Special elongation reactions, taking place both in the mitochondria and at the surface of the endoplasmic reticulum are used to make longer chains. The ER reactions of adding two-carbon units at the carboxyl end of the chain by means of oxidative decarboxylations involving malonyl-CoA provide the thermodynamic driving force for the condensation reaction. The mitochondrial reactions involve addition (and subsequent reduction) of acetyl units. These reactions are essentially a reversal of fatty acid oxidation, with the exception that NADPH is utilized in the saturation of the double bond, instead of $FADH_2$.

In a newly synthesized fatty acid, both prokaryotes and eukaryotes can introduce a single *cis* double bond. While eukaryotes need an O_2-dependent pathway to carry out this process, bacteria like *E. coli* can need an O_2-independent pathway. While the O_2-independent reaction requires some other

means to activate the desired bond toward dehydrogenation, O_2-dependent reaction can occur anywhere in the fatty acid chain.

In E. coli, four normal cycles of elongation begin the biosynthesis of a monounsaturated fatty acid to form a 10-carbon intermediate, β-hydroxydecanoyl-ACP. At this point, β-hydroxydecanoyl thioester dehydrase forms a double bond 3,4 to the thioester and in the cis configuration. This is followed by three rounds of the normal elongation reactions to form palmitoleoyl-ACP. Elongation may terminate at this point or may be followed by additional biosynthetic events. cis-vaccenic acid, the principal unsaturated fatty acid in E. coli is formed by an additional elongation step, using palmitoleoyl-ACP as a substrate.

It is only after the fatty acyl chain has reached its full length, does the addition of double bonds to fatty acids in eukaryotes occur. Even if no useful functional group exists on the chain to facilitate activation, dehydrogenation of stearoyl-CoA occurs in the middle of the chain:

$$CH_3O(CH_2)_{16}CO\text{-}SCoA \rightarrow CH_3O(CH_2)_7CH = CH(CH_2)_7CO\text{-}SCoA$$

stearoyl-CoA desaturase, a 53-kD enzyme containing a non-heme iron center catalyzes this reaction. Other requirements include, NADH, and oxygen (O_2) and two other proteins; cytochrome b_5 reductase (a 43-kD flavoprotein) and cytochrome b_5 (16.7 kD). A pair of electrons is transferred by cytochrome b_5 reductase from NADH through FAD to cytochrome b_5. Reduction of nonheme Fe^{3+} to Fe^{2+} in the desaturase occurs with oxidation of reduced cytochrome b_5. The Fe^{3+} accepts a pair of electrons (one at a time in a cycle) from cytochrome b_5 and creates a cis double bond at the 9, 10 position of the stearoyl-CoA substrate. O_2 is the terminal electron acceptor in this fatty acyl desaturation cycle. Animals can synthesize fatty acids with double bonds at positions beyond C-9 only by this method.

This single desaturation reaction may lead to additional chain elongation. Two carbons can be used to elongate the oleoyl-CoA produced to form a 20:1 *cis*-Δ^{11} fatty acyl-CoA. Reactions similar to the preceding scheme yield palmitoleoyl-CoA, if the starting fatty acid is palmitate. This can subsequently be elongated to yield *cis*-vaccenic acid. Similarly, C_{16} and C_{18} fatty acids can be elongated to yield C_{22} and C_{24} fatty acids.

Differences occur in organisms with respect to formation, processing and utilization of polyunsaturated fatty acids. For example, while Eukaryotes synthesize different types of polyunsaturated fatty acids, E.Coli does not have any such acids. While plants manufacture double bonds between the Δ^9 and the methyl end of the chain, mammals cannot. However, they can introduce double bonds between the double bond at the 8- or 9-position and the carboxyl group. Additional double bonds to unsaturated fatty acids can be added by mammals in their diets. They can make arachidonic acid from linoleic acid. This fatty acid serves as the precursor for prostaglandins and other biologically active derivatives like leukotrienes. Synthesis involves formation of a linoleoyl ester of CoA from dietary linoleic acid and introduction of a double bond at the 6-position. This is followed by elongation of a triply unsaturated product by malonyl-CoA with a decarboxylation step so as to yield a 20-carbon fatty acid with double bonds at the 8, 11, and 14-positions.

Finally, a 20-carbon fatty acid with double bonds at the 5-, 8-, 11-, and 14-positions is liberated by a second desaturation reaction at the 5-position followed by an acyl-CoA synthetase reaction. These steps are shown in Figure 3.50.

NOTES

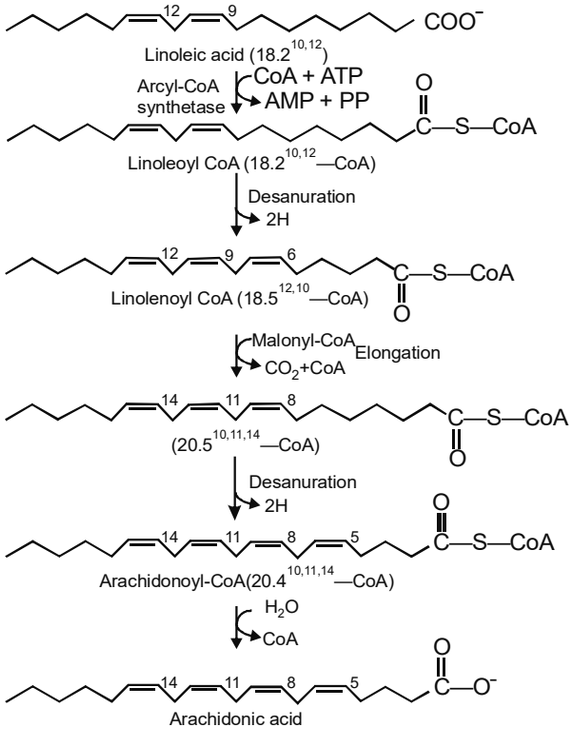


Fig. 3.50 *Coupling of Fatty Acid Synthesis and Fatty Acid Oxidation*

Check Your Progress

14. What are important characteristics of liver?
15. Which adipose tissue is responsible for clearance of chylomicrons?
16. Which enzyme carries out the cleavage and formation of CoA thioester of fatty acids inside mitochondria?
17. What are ketone bodies?
18. What are two pathways by which cholesterol is eliminated?
19. What are the main sources of acetyl-CoA?
20. What is fatty acid synthase?

3.6 ANSWERS TO 'CHECK YOUR PROGRESS'

1. Lipids are considered as important constituents of the diet because of their high energy value and also because they are precursors of the fat-soluble vitamins.
2. Fatty acids are long-chain organic acids having generally carbon atoms from 4 to 30 and they have a single carboxyl group with a long, nonpolar hydrocarbon tail.

NOTES

3. The names of the saturated fatty acids end with the suffix 'anoic' acid.
4. Saponification refers to the hydrolysis of fats by alkali.
5. The two best sources of carotene are carrots and alfalfa.\
6. The chain length of the constituent fatty acid and the degree of unsaturation regulate the melting point of fats.
7. Triacylglycerols are the simplest lipids constructed from fatty acids and they are composed of three fatty acids each in ester linkage with a single glycerol molecule.
8. The two examples of biological waxes are lanolin and beeswax.
9. Soyabean and yeasts are the two important sources of lecithins.
10. Lipid molecules are composed primarily of hydrocarbon groups which make them hydrophobic (water insoluble) and hydrophilic groups (water soluble). Such molecules which contain both hydrophobic (a nonpolar tail) and hydrophilic (a polar head) fractions are known as amphipathic molecules (amphi-both, pathos-passion).
11. When the concentration of the amphiphilic molecule reaches a given concentration called critical micelle concentration (cmc), it led to the formation of micelle aggregates.
12. The lipid bilayer performs various biological functions, like:
    a. Lipid bilayer structure is the foundation of all the membranes found in the biological system, together with the several membrane proteins. Inner and outer monolayers have different lipid compositions, which make the two faces of bilayer functionally different.
    b. They show prominent role in separation. It forms the barrier of the aqueous compartment from their neighbouring surroundings. The existence of a cell without a membrane cannot be even imagined. It differentiates the cells, provide the knowledge of 'self' and 'non-self'. The amphipathic property of lipid bilayer does not allow hydrophilic molecules to cross the hydrophobic bilayer core.
13. The sugars molecules are used as markers in biological membrane because of structural diversity of their chains. These molecules are bonded to the lipids and proteins.
14. The important characteristics of liver are: (i) They are insoluble in water. (ii) They contain a large proportion of carbon-hydrogen bonds and release a lot of energy on breakdown.
15. Lipoprotein lipase in the adipose tissue is responsible for the clearance of chylomicrons.
16. Thiolase carries out the cleavage and formation of CoA thioester of fatty acids inside mitochondria
17. The fatty acids undergo excessive oxidation in liver under certain metabolic conditions producing large quantities of keto acids-acetoacetic acid and b-hydrobutyric acid, which pass into the blood by diffusion. Acetoacetic acid then undergoes spontaneous decarboxylation to produce acetone. These

NOTES

three substances-acetoacetate p, hydroxybutyrate and acetone are collectively known as ketone bodies (acetone bodies or ketones).

18. Cholesterol is eliminated by two main pathways (a) Conversion to bile acids and (b) By excretion of neutral steroids in the faeces.
19. The main sources of acetyl-CoA are amino acid degradation, fatty acid oxidation and glycolysis.
20. Fatty acid synthesis is the process of creating fatty acids from acetyl-CoA and malonyl-CoA precursors through the action of enzymes called fatty acid synthases.

3.7 SUMMARY

- Lipids are a heterogeneous group of compounds, including fats, oils, sterols, waxes and related compounds which are associated more by their physical and by their chemical properties.
- Fatty acids are long-chain organic acids having carbon atoms and they are broadly classified into saturated and unsaturated fatty acids.
- Unsaturated fatty acids may be further classified into monoethenoid, diethenoid, triethenoid and tetraethenoid acids based on the composition of the double bonds.
- Fatty acids have both physical and chemical properties. The physical properties of fatty acids include colour, odour, taste, solubility, melting point, specific gravity, etc.
- The chemical properties of fatty acids include reactions involving –COOH group, double bonds, -OH groups, etc.
- Reactions involving –COOH group are hydrolysis, saponification, hydrolytic rancidity.
- Reactions involving double bond include hydrogenation, halogenation, oxidation and oxidative rancidity.
- Reactions involving –OH groups include Acrolein test, quantitative tests such as acid value, saponification number, iodine value, polenske number, Reichert-Meissl number, etc.
- Lipids perform a wide variety of functions such as providing food material of high calorific value, acting as food reserve, having high insulating capacity, acting as sources of natural fat-soluble vitamins, etc.
- Lipids are classified into three broad categories which are as follows: Simple Lipids, Conjugated Lipids, and Derived Lipids.
- Simple lipids can be again classified into Triglycerides and Waxes.
- Triaclyglycerols are the simplest lipids that are constructed from fatty acids and they are composed of three fatty acids each in ester linkage with a simple glycerol molecule.
- Conjugated lipids are further classified into phospholipids, glycolipids, sulfolipids and lipoprotein,

NOTES

- Derived lipids are classified into fatty acids, glycerol, cholesterol, vitamins A, E, K, etc.
- Lipid molecules are composed primarily of hydrocarbon groups which make them hydrophobic (water insoluble). Some of the lipid possess hydrophilic groups too which make them water soluble.
- Lipids may form the stable aggregated like micelles and thermo thermodynamically unstable aggregates like bilayer sheets.
- Micelles are the aggregates of amphipathic molecules formed when the molecules are mixed with an aqueous solution.
- Lipid Bilayer is a membranous structure consisting of two layers of amphipathic lipid molecules.
- The structure of liposomes was deduced for the first time by British haematologist, Alec D. Bangham in 1961**.** Liposome consists of words 'Lipos' and 'Soma', meaning fat and body, respectively.
- Liposomes can be made with various different combinations of structure, size, shape, and surface properties as they show a very wide diversity in terms of size and shape.
- The biological membrane covers the cells, shield and separate it from the surroundings. Biological membrane is a semi-permeable membrane which allows the selective movement of molecules only from outside into the cell and from cell to outside.
- The fluid mosaic hypothesis was formulated by Singer and Nicolson in the early 1970. According to this model, membranes are made up of lipids, proteins and carbohydrates. The main lipid membrane components are phospholipids.
- The 'mosaic' term of this model means the mixture of lipids and intrinsic proteins in the membrane. These boundaries are also 'fluid' because their components can move, allowing both diffusion of components and local specific gatherings.
- Lipids are any of a group of organic compounds that include fats, oils, waxes, sterols and triglycerides, which are insoluble in water.
- Lipids are soluble in non-polar organic solvents such as acetone, alcohol, chloroform, benzene and ether. The alkaline hydrolysis (saponification reaction) of lipids yields alcohol and fatty acids, which may be water soluble
- Fat is a concentrated source of energy and occur widely as reserve material, particularly in higher plants; for example, in seeds.
- All enzymes associated with the p-oxidation system are localized in the inner membranes and the matrix of liver and other tissue mitochondria.
- The a-oxidation system has been shown to play a key role in the capacity of mammalian tissue to oxidize phytanic acid, the oxidation product of phytol, to CO_2 and water.
- The mechanism of the oxidation of oils is primarily by the co-oxidation mechanism.

NOTES

- The degradation of fatty acids by oxidation at the a-carbon (beta-carbon) was first made by Franz Knoop in 1904.
- Ketone bodies are alternative substrates to glucose for energy sources in muscle and brain. The precursor of ketone bodies, namely free fatty acids are toxic in high concentrations, have a very little solubility and readily saturate the carrying capacity of the plasma albumin.
- Ketosis generally occurs in severe diabetes mellitus prolonged starvation, glycogen storage disease, toxemia of pregnancy, infective hepatic disease and continued fever.
- The liver is the main site of cholesterol synthesis and other tissues known to be capable of synthesizing cholesterol include the testis, aorta, skin, intestine and the adrenal cortex.
- Fatty acid synthesis is the process of creating fatty acids from acetyl-CoA and malonyl- CoA precursors through the action of enzymes called fatty acid synthases.

3.8 KEY TERMS

- **Lipids:** These are soluble biomolecules in nonpolar solvents. These heterogeneous group of compounds include fats, oils, steroids, waxes, and other compounds.
- **Fatty Acid:** It is a hydrocarbon that ends with carboxylic acid groups. It can be either saturated or unsaturated.
- **Oxygenated Fatty Acids:** These refer to fatty acids that may have hydroxyl group in the chain.
- **n-6 Fatty Acids:** It is the name given to Lin and longer chain fatty acids, which are derived from Lin.
- **Lipases:** These are enzymes that catalyse the hydrolysis of stored triacylglycerols, releasing fatty acids for export to sites where they are required as fuel.
- **Reichert-Meissl Number:** It refers to the number of millilitres of 0.1N KOH required to neutralize the soluble, volatile fatty acids derived from 5g of fat.
- **Metabolism:** It is the chemical processes that occur within a living organism in order to maintain life.

3.9 SELF ASSESSMENT QUESTIONS AND EXERCISES

Short-Answer Questions

1. State the important functions of lipids.
2. Briefly explain the process of hydrolytic rancidity.

3. What are the various functions of micelles?
4. Write a short note on the fluid mosaic model of membrane structure.
5. What is the clinical significance of cholesterol?

Long-Answer Questions

1. Describe the physical and chemical properties of fatty acids.
2. Discuss the various types of conjugated lipids.
3. Explain the structure and functions of liposomes.
4. Discuss the metabolism of lipids in detail.
5. Describe the α-oxidation and β oxidation of fatty acids.

NOTES

3.10 FURTHER READING

Blackstock, James C. 2014. *Guide to Biochemistry*. Oxford: Butterworth-Heinemann.

Park, K. H. 2008. *Carbohydrate-Active Enzymes: Structure, Function and Applications*. Amsterdam: Elsevier.

Bhardwaj, Uma. 2012. *Biochemistry for Nurses*. Noida: Pearson Education India.

Fromm, Herbert J. and Mark Hargrove. 2012. *Essentials of Biochemistry*. Berlin: Springer.

Fearon, William Robert. 2014. *An Introduction to Biochemistry*. Amsterdam: Elsevier.

UNIT 4 AMINO ACIDS, PEPTIDES AND PROTEINS

Structure

4.0 INTRODUCTION

Amino acids are important for living beings as they perform several roles in the metabolic activity occurring in biological organisms. One of the most important functions is performing the role of being the building blocks of proteins. Proteins constitute the main class of molecules studied in biochemistry. They are also an important constituent of all living organisms and supply the majority of the molecular mechanism of cells. Many proteins form enzymes or function as sub-units of

enzymes. There are also other proteins that have structural or mechanical functions, such as forming the joints of the skeleton. A molecule of protein is composed of an un-branched polymer or chain of amino acids. Protein polymers have extremely large molecular mass that is made of one or more than one polypeptide chains. The monomers are amino acids, joined together by peptide bonds. The human body needs proteins for specific purposes, such as to make haemoglobin and build cardiac muscles. Proteins also act as nutrient sources for organisms that do not produce their own energy from sunlight. In this unit, we will discuss the concepts, formation, and structure of amino acids, peptides and proteins. It will also focus on the amino acids metabolism, and chemistry of oxytocin and Thyrotropin Releasing Hormone (TRH).

4.1 OBJECTIVES

After going through this unit, you will be able to:

- Describe the concepts, formation, and structure of amino acids, peptides and proteins
- Explain the amino acids metabolism
- Analyze the sequence determination
- Discuss the chemistry of oxytocin and Thyrotropin Releasing Hormone (TRH)

4.2 AMINO ACIDS AND THEIR SEQUENCING

The molecules containing an amine group, a carboxylic acid group and a side chain (R) containing carbon, hydrogen, oxygen, and nitrogen having general formula is -H_2NCHRCOOH, where R is an organic substituent are called as amino acid (Figure 4.1). In an amino acid, the amino and carboxylic groups are attached to the same carbon atom, which is called the α–carbon and R-groups varies in different types of amino acids.

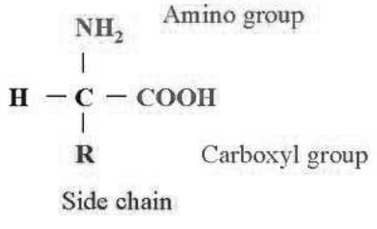


Fig. 4.1 *Structure of Amino Acid*

Amino acids are critical to life as they perform many functions in metabolic activity occurring in biological organism. One of the most important functions is as the building blocks of proteins. The unique sequence in linear fashion of amino acid residues chemically defines primary structure of protein, which in turn also defines the three-dimensional structure of the protein. Just like as the letters of the alphabet can be combined to form an almost endless variety of words, amino acids can

also be joined together in varying sequences to form a vast variety of proteins. Amino acids are also plays crucial role in biological organism such as forming parts of coenzymes or as precursors for the biosynthesis of molecules such as hemoglobin. Due to having central role in biochemistry, amino acids are very important in nutrition. Amino acids are commonly used in food technology and industry. An application of amino acid in industry includes the production of biodegradable plastics, drugs and chiral catalysts.

4.2.1 Physical Properties of Amino Acids

Amino acids are colorless, crystalline substances. The crystal shape may vary from slender needles as found in tyrosine to thick hexagonal plates as presented in Cysteine. They may be either sweet (glycine and alanine) or bitter (arginine) and tasteless (tyrosine) in nature. Amino acids have high melting points approximately 200°C and often result in decomposition. They are usually soluble in polar solvents (ethanol and water but are insoluble in non-polar solvents (ether and benzene). Some amino acids like tyrosine, histidine tryptophan, and phenylalanine absorb ultraviolet radiation between 300-320 nm wavelengths but most of amino give greatest absorbance at 280 nm. These characteristics are useful in recognition of not only these amino acids but also the proteins which include them. The polymers of amino acids forms peptides and proteins range in size from small to very large, consisting of two or three to thousands of linked amino acid residues as presented in living organism. Two amino acid molecules can be covalently joined through a substituted amide linkage, termed as peptide bond**,** to yield a dipeptide. Such a linkage is formed by removal of the elements of water (dehydration) from α -carboxyl group of one amino acid and α -amino group of another amino acid molecule (Figure 4.2).

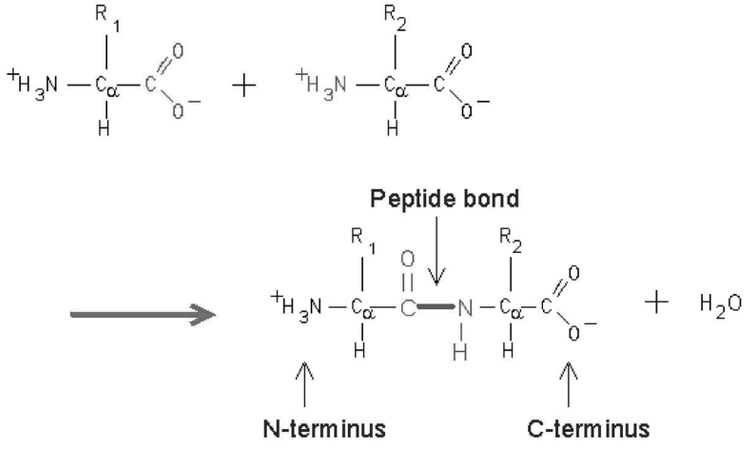


Fig. 4.2 *Formation of Peptide Bond*

Peptide bond formation is an example of a condensation reaction, a common class of reactions in living cells. Similarly three amino acids can be joined by two peptide bonds to form a tripeptide. Hence, amino acid can form tetrapeptides, pentapeptides, and so forth in similar order. When a few amino acids up to 20 are joined in this fashion, the structure is called an oligopeptide**.** When more than 20 amino acids are joined, the product is called a polypeptide. Proteins may have

thousands of amino acid residues. Although the terms "protein" and "polypeptide" are sometimes used interchangeably, molecules referred to as polypeptides generally have molecular weights below 10,000, and those called proteins have higher molecular weights. An amino acid unit in a peptide is often called a residue which can be defined as the part left over after losing a hydrogen atom from its amino group and the hydroxyl moiety from its carboxyl group. In a peptide, the amino acid residue at the end with a free α-amino group is the amino-terminal (*N*-terminal) residue and the residue at the other end, which has a free carboxyl group, is the carboxyl-terminal (*C*-terminal) residue (Figure 4.2). Although hydrolysis of a peptide bond is an exergonic reaction, it occurs slowly because of its high activation energy. As a result, the peptide bonds in proteins are quite stable, with an average half-life ($t_{1/2}$) of about 7 years under most intracellular conditions.

The 20 naturally occurring amino acids which are commonly found in proteins can be divided into several groups based on their properties which include charge, hydrophilicity or hydrophobicity, size and functional groups. These properties are important for protein structure and protein–protein interactions. The first amino acid to be discovered was asparagine in 1806. The last of the 20 amino acid to be found is threonine in 1938. All the amino acids have common trivial names but in some cases, they are named on the basis of the source from which they were first isolated. For example asparagine was first found in asparagus and glutamate in wheat gluten, tyrosine was first to isolated from cheese and glycine from sweet was so named because of its sweet taste.

4.2.2 Chemical Properties of Amino Acids

The chemical properties of amino acids are mainly determined by the presence of two functional groups: carboxyl group and amino group. The properties due to the amino group are as follows:

- **Basic Nature:** Since the amino group is basic in nature, it reacts with acids to form salts.
- **Transamination Reaction:** In amino acid metabolism, the amino group of an amino acid is transferred to a keto acid to form a new amino acid. This reaction is known as a transamination reaction.
- **Deamination Reaction:** In amino acid metabolism, free ammonia is removed from amino acid by oxidative deamination reaction.
- **Reaction with Ninhydrin:** After reacting with ninhydrin, amino acid forms a purple, blue or pink colour complex, known as Ruhemann's purple. This colour complex can be used to estimate amino acid and protein levels quantitatively.

4.2.3 Classification of Amino Acids

The classification of amino acids can be done in the following discussed ways.

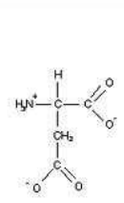
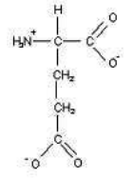
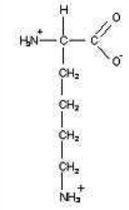
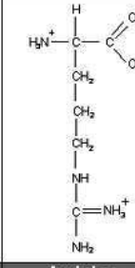
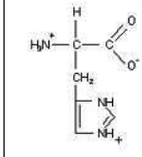
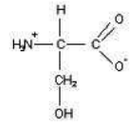
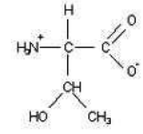
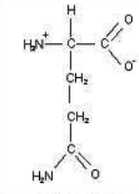
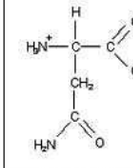
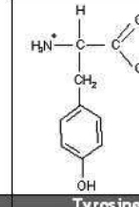
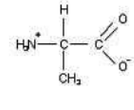
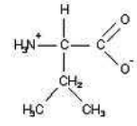
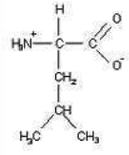
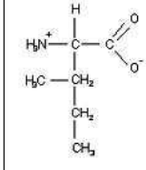
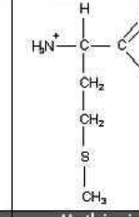
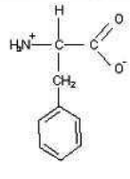
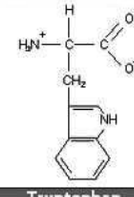
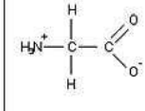
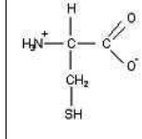
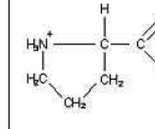
A. On the Basis of their R Groups

The knowledge of the chemical properties of the common amino acids is fundamental to an understanding of biochemistry. The area under discussion can be made easy by grouping the amino acids into five major classes based on the

NOTES

properties of their R groups in particular, their polarity, or tendency to interact with water at neutral pH near 7.0 (Table 4.1). The polarity of the R groups varies widely, from Nonpolar and hydrophobic (water-insoluble) to highly polar and hydrophilic (water-soluble). Within each class there are gradations of polarity, size, and shape of the R groups as described in the below texts.

Table 4.1 *Showing the Structures of 20 Amino Acids*

Polar		Non-polar		"Special"
				
Aspartic Acid (asp or D)	Glutamic Acid (glu or E)	Lysine (lys or K)	Arginine (arg or R)	Histidine (his or H)
				
Serine (ser or S)	Threonine (thr or T)	Glutamine (gln or Q)	Asparagine (asn or N)	Tyrosine (tyr or Y)
				
Alanine (ala or A)	Valine (val or V)	Leucine (leu or L)	Isoleucine (ile or I)	Methionine (met or M)
				
Phenylalanine (phe or F)	Tryptophan (trp or W)	Glycine (gly or G)	Cysteine (cys or C)	Proline (pro or P)

1. Non-Polar, Aliphatic R Groups

This class of amino acids has nonpolar and hydrophobic R groups. The side chains of alanine, valine, leucine, and isoleucine have a tendency to cluster together within proteins, stabilizing protein structure by means of hydrophobic interactions (Figure 4.3). Glycine has the simplest structure as they have only H-atom in R-group position so; it's very small side chain makes no real contribution to hydrophobic interactions. We can say glycine formally as nonpolar. Methionine is one of the two sulfur-containing amino acids, has a nonpolar thio-ester group in its side chain. Proline has an aliphatic side chain with a typical cyclic structure. The secondary

NOTES

amino (imino) group of proline residues is detained in a rigid conformation that reduces the structural flexibility of polypeptide regions containing proline.

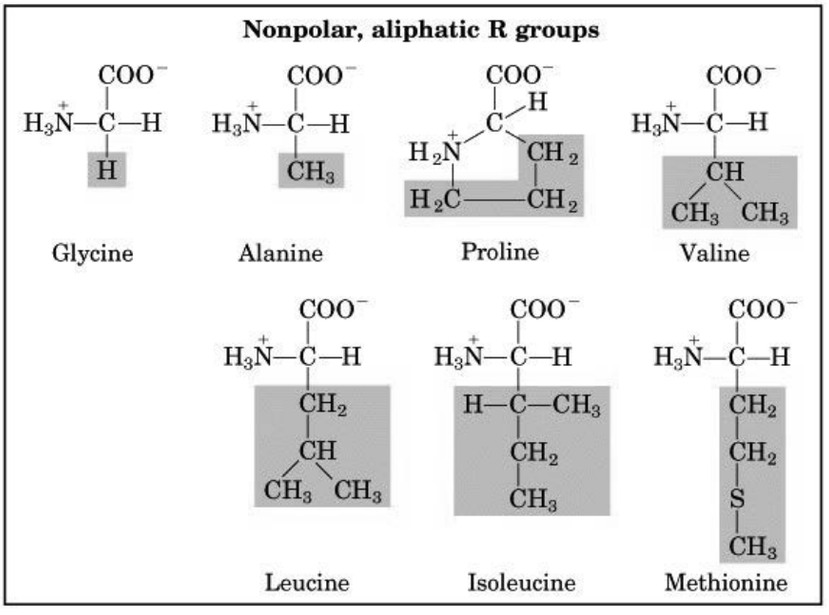


Fig. 4.3 *Structure of Nonpolar, Aliphatic R Groups Amino Acid*

2. Aromatic R Groups

The aromatic side chains containing amino acids include Phenylalanine, tyrosine, and tryptophan (Figure 4.4) which are relatively nonpolar (hydrophobic). All these amino acids can participate in hydrophobic interactions. The hydroxyl group of tyrosine can form hydrogen bonds and some enzyme contains this group as an important functional group. Tyrosine and tryptophan are significantly more polar than phenylalanine due to presence of hydroxyl group in tyrosine and nitrogen in tryptophan indole ring. Tryptophan and tyrosine, and to a much lesser extent phenylalanine, absorb ultraviolet light which accounts for the characteristic strong absorbance of light by most proteins at a wavelength of 280 nm, an important characteristic oppressed by researchers in the characterization of proteins.

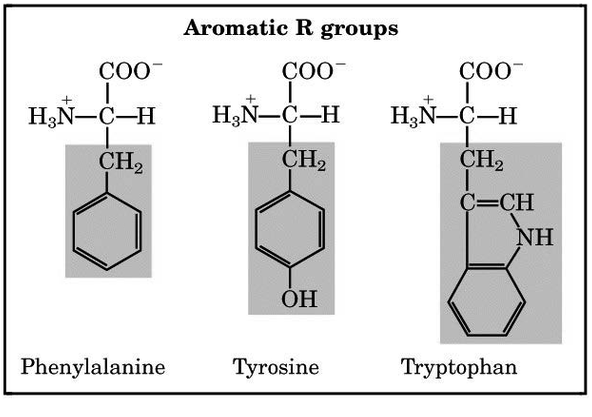


Fig. 4.4 *Structure of Aromatic R Groups Amino Acid*

NOTES

3. Polar, Uncharged R Groups

The R groups of polar, uncharged amino acids includes serine, threonine, cysteine, asparagine, and glutamine (Figure 4.5) are more soluble in water than those of the nonpolar amino acids because they contain functional groups that form hydrogen bonds with water. The polarity of amino acids depends on functional group presented in their side chain such as in serine and threonine is contributed by their hydroxyl groups, sulfhydryl group in cysteine and amide groups in asparagine and glutamine. Asparagine and glutamine are the amides of two other amino acids which are also found in proteins called as aspartate and glutamate formed by the hydrolysis of asparagine and glutamine by an acid or base. Oxidation of two molecule of cysteine form a covalently linked dimeric amino acid called cystine, in which cysteine molecules or residues are joined together with the help of disulfide bond. The disulfide bonds are strongly hydrophobic (nonpolar) in nature and play a specific role in maintain the structures of several proteins by forming covalent bonding between two different polypeptide chains or between parts of a protein molecule.

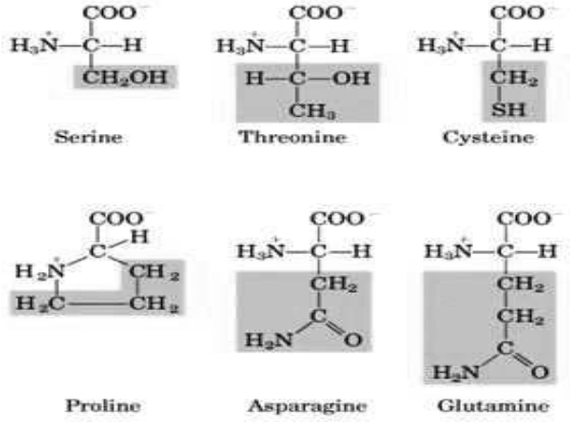


Fig. 4.5 *Structure of Polar, Uncharged R Groups Amino Acid*

4. Positively Charged (Basic) R Groups

The most hydrophilic R groups are those that are either positively or negatively charged. Lysine are amino acids in which the R groups have considerable positive charge at pH near neutrality and has a second primary amino group at the ε position on its aliphatic chain (Figure 4.6). Arginine has a positively charged guanidino group while histidine has an imidazole group. Histidine is the only common amino acid having an ionizable side chain with a pKa near neutrality. A histidine residue facilitates the reaction by serving as a proton donor/acceptor n many enzyme-catalyzed reactions. Lysine and arginine are basic amino acids which serve in forming the structure of histone protein.

NOTES

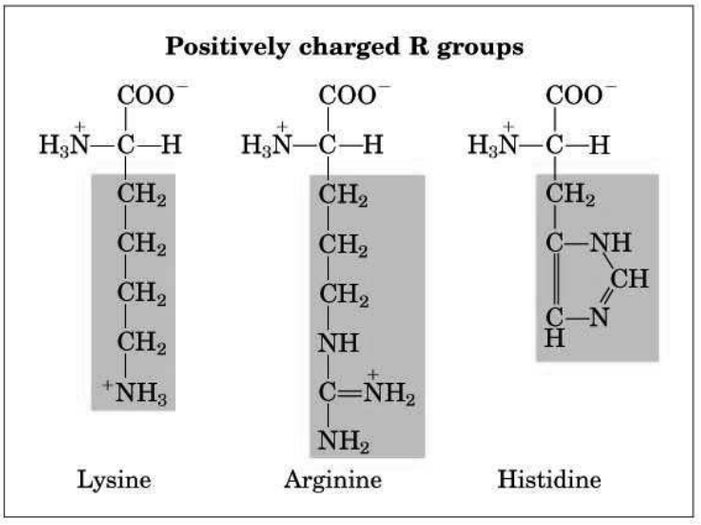


Fig. 4.6 *Structure of Positively Charged R Groups Amino Acid*

5. Negatively Charged (Acidic) R Groups

The R group amino acids having a net negative charge at neutral pH includes aspartate and glutamate. Both of these amino acids contain two carboxyl (COOH) groups in their side chain (Figure 4.7).

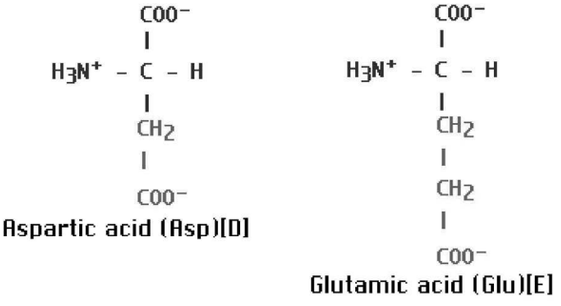


Fig. 4.7 *Structure of Negatively Charged R Groups Amino Acid*

B. On the Basis of Synthesis

In reality, all the 20 amino acids are necessary to the organism in the common sense that all must be present for the synthesis of protein and continuing of life processes. A certain number of living organisms (e.g., plants and bacteria) are able to produce all these 20 amino acids from amphibolic intermediates, whereas, other forms together with human and other animals can biosynthesize only half of those amino acids as required. On that basis, amino acids are classified into two categories described below.

1. Non-Essential Amino Acids

The amino acids which are formed in our body and are not required in the course of any diet, they are referred as non-essential amino acids (Table 4.2). Out of 11 nutritionally non-essential amino acids, eight are formed from amphibolic intermediates and three (cysteine, tyrosine and proline) are formed from nutritionally

essential amino acid. The detailed properties of all non-essential amino acids are described separately.

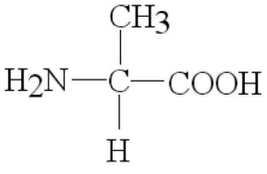
Table. 4.2 *Showing the List of Essential and Non-essential Amino Acid*

Essential	Nonessential
Histidine	Alanine
Isoleucine	Aspartate
Leucine	Cysteine
Lysine	Glutamate
Methionine	Glutamine
Phenylalanine	Glycine
Threonine	Proline
Tryptophan	Serine
Valine	Tyrosine
Arginine	Asparagine

NOTES

(i) Alanine

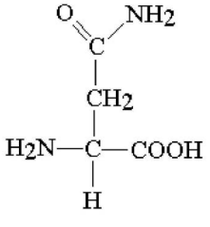
Alanine is formed by transamination of pyruvate and can be defined as α-amino acid linked with a methyl group (-CH_3). Its codons are GCU, GCC, GCA, and GCG. Some bacterial cell walls and few peptide antibiotics contain alanine in their D- isomeric forms. In liver and muscle tissue, it is used for maintaining blood glucose levels by converting stored glycogen into glucose which is needed by the body to satisfy its energy needs.



alanine

(ii) Asparagine

Asparagine has carboxamide group in their side chain as functional group. Its codons are AAU and AAC. The conversion of aspartate to asparagine is catalyzed by asparagine synthetase. It participates in the metabolic control of cells functions in the brain and nervous system, thus used in treatment of brain and nervous system disorders. It is also involved in the formation of ammonia.

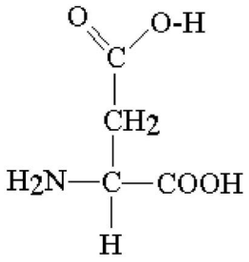


asparagine

NOTES

(iii) Aspartate

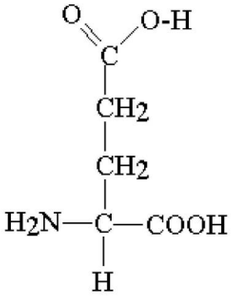
Aspartate is formed by transamination of oxaloacetae. It is also known as aspartic acid. Its codons are GAC and GAU. Ribonucleotides are formed by nitrogen presented in the amide group of this amino acid, which acts as precursors to RNA and DNA. It facilitates the removal and detoxification of ammonia in liver of the body. Resistivity to fatigue can be enhanced by aspartate.



Aspartic Acid

(iv) Glutamate

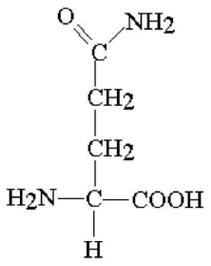
It is formed by amination of glutamate catalyzed by the enzyme glutamine synthetase. The codons for glutamate include GAA and GAG and are also known as glutamic acid. It constitutes 50% of the total amino acid composition of protein found in the brain. The transportation of potassium across the blood brain barrier is facilitated by this amino acid. It also acts as an excitatory neurotransmitter in the central nervous system, so it is regarded as important fuel for brain. It is presented in stomach for maintaining the physiological environment.



glutamic acid

(v) Glutamine

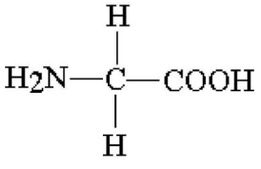
Glutamine acts as brain barrier when eagerly passing through the blood stream. It delivers ammonia to kidneys for deanimation by picking it from central nervous system. It is used in the treatment of schizophrenia and senility. It protects against the poisonous effects of alcohol. They are also present in the intestinal tract. Its codons are CAA and CAG.



glutamine

(vi) Glycine

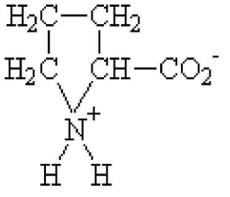
Glycine is the simplest and smallest amino acid. It can be present inside or outside of the protein molecule. In aqueous solution at or near neutral pH, Glycine will found predominantly as zwitter ion. Molecules that contain an equal number of ionizable groups of opposite charge and therefore bear no net charge are termed as Zwitter ions. Its codons are GGC, GGU, GGA and GGG .Glycine could be considered a derivative of amino-ethane.Glycine amino transferase can catalyze the synthesis of glycine from glyoxylate and glutamate or alanine. Additional important mammalian routes for glycine formation are from Choline and from serine. It mediates the formation of the purine framework which is utilized in manufacturing of DNA and RNA strands. It acts as inhibitory neurotransmitter in the central nervous system. It also promotes the release of growth hormone and facilitates the synthesis of hemoglobin. It may also used as food additive for making the taste sweeter and is also effective in neutralizing hyperacidity occurring in the stomach.



glycine

(vii) Proline

Proline is considered as an imino acid rather than amino acid. Its codons are CCC, CCU, CCA and CCG. It is the major component of collagen, the chief fibrous protein presented in bone, cartilage and other connective tissue. That's why it is used in the repairing and safeguarding of ligaments and tendons.



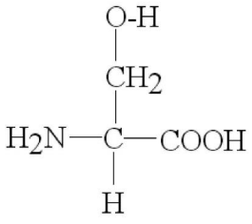
L-proline

pro

NOTES

(viii) Serine

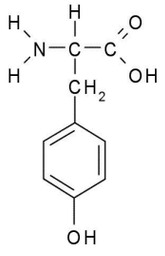
The oxidation of α-hydroxyl group of 3-phosphoglycerate which is a glycolytic intermediates converts it to an oxoacid, whose successive transamination and dephosphorylation leads to production of serine. Its codons are UCC, UCU, UCA and UCG. Serine assists in the manufacture of the organic acid creatine, found in blood while also present within muscle and brain tissue as an energy source for muscle contraction. It also participant in biosynthesis of purine, pyrimidine and porphyrin. It acts as natural moisturizing agent used in cosmetics.



serine

(ix) Tyrosine

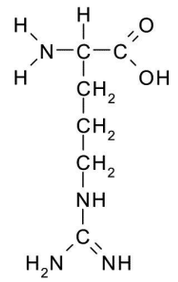
Tyrosine is formed by the hydroxylation of phenylalanine. Tyrosine is hydrophobic that's why it is significantly more soluble that phenylalanine and more acidic than serine or threonine. Its codons are ACC, ACU, ACA and ACG. Tyrosine amino acid is able to absorb UV radiation which contributes to the absorption of protein. The extinction of tyrosine is just around 1/ 5th of that of tryptophan at 280 nm, which is the primary contributor to the UV absorbance of proteins depending upon the number of tyrosine amino acid residues in the protein. For avoiding nutritional conflicts to occur, it must be kept in balance with other amino acids through dietary intake. In the liver, it acts as a lipotropic agent to prevent fat formation. It is an important component of collagen, elastin and enamel proteins.



Tyrosine

(x) Arginine

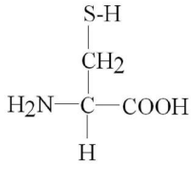
Arginine is a basic amino acid and its codons are CGC, CGU, CGA and CGG. It enhances the immune system for accelerating wound curing and blocking the development of tumors. Arginine causes growth hormones release in the anterior pituitary gland. It increases the formation of sperms (spermatogenesis) in sex glands. It also mediates the detoxification of ammonia and is concerned with liver restoration.



Arginine

(xi) Cysteine

Cysteine is formed from methionine, which is nutritionally essential amino acid. Its codons are UGC and UGU. Cysteine is used to stimulate the activity of white blood corpuscles (WBC) in concerned to the immune system. It serves as essential constituent for the formation of skin, and promotes healing from cuts and burns. It can also acts as effective antioxidant and scavenger of free radicals. It is the only amino acid which form connections with the protein involved in folding process.



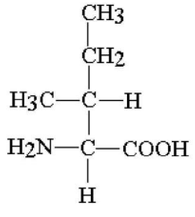
cysteine

2. Essential Amino Acids

An essential amino acid is an amino acid that cannot be synthesized by the human beings and other animals and must be supplied through diets. Essential amino acids are 'essential' because the body does not synthesize them, making it essential to include them in one's diet in order to obtain them. Out of twenty amino acids, eight are nutritionally essential amino acid for humans and other animals whose names are presented in Table 4.2. The detailed properties of all essential amino acids are described separately in the below stated texts.

(i) Isoleucine

Isoleucine is one of the three amino acid (the other two being leucine and valine), major Branched-Chain Amino Acids (BCAA). Muscle tissues contain high concentrations of isoleucine as an energy source which provides strength and stamina to muscles. It is also required in the formation of hemoglobin. Its codons are AUC, AUU and AUA.

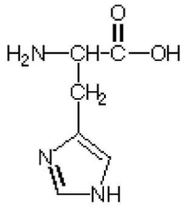


isoleucine

NOTES

(ii) Histidine

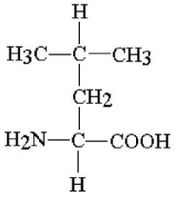
Histidine is found in hemoglobin in high concentration. So, it is useful in treating anemia and rheumatoid arthritis due to relationship to hemoglobin. It acts as precursor to histamine associated with allergic response and has been used to treat allergy. It helps in maintaining proper blood pH. High histidine levels are associated with low zinc levels and low Histidine levels are associated with high zinc levels. Thus, abnormal Histidine levels are an indicator that zinc levels should be tested. It also helps to remove heavy metals from tissues by chelating trace minerals. Its codons are CAC and CAU.



Histidine

(iii) Leucine

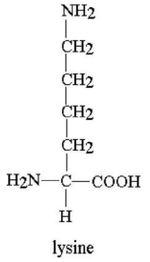
Leucine acts as persuasive stimulator of insulin. It helps promoting in bone and skin healing. It modulates the release of encephalin, which are natural pain-reducers. Its codons are CUC, CUU, CUA and CUG.



leucine

(iv) Lysine

Lysine inhibits the viral growth and as a result, it is used in the treatment of Herpes Simplex, as well as the viruses associated with Chronic Fatigue Syndrome, such as- Epstein-Barr Virus, Cyto-megalo Virus etc. Lysine in association with vitamin C leads to the formation of L-Carnitine. Its codons are AAA and AAG. It helps to form collagen, the connective tissue present in bones, ligaments and tendons. It also assists in the absorption of calcium and production of hormone. It is critical for bone formation, so it is essential in to take in diet for children.

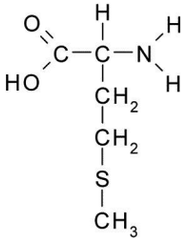


lysine

NOTES

(v) Methionine

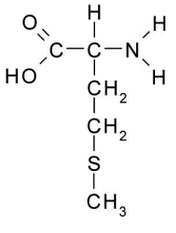
Methionine is the sulfur-containing amino which assists in breakdown of fats to reduce blood cholesterol levels and the removal of toxic wastes from the liver. It acts as precursor for the production of amino acids cysteine. It acts as anti-oxidants to neutralize free radicals and assist to prevent disorder of hair, skin, and nails occurring due to sulfur deficiency. It serves as precursor for the formation of carnitine, melatonin (the natural sleep supplements) and choline (part of the neurotransmitter as in Acetylcholine). Methionine is concerned to the breakdown of nicotinic Acid, histamine and epinephrine. It is also required for production of DNA and RNA molecules. It acts as natural chelating agent for heavy metals such as, mercury (Hg) and lead (Pb). Its codon is AUG which mostly serves as initiation codon in protein synthesis.



Methionine

(vi) Phenylalanine

Phenylalanine serves as precursor for the formation of Tyrosine is precursor to neurotransmitters such as, dopamine and the excitatory neurotransmitters epinephrine and norepinephrine. It also serves as precursor to the Thyroxin hormone. It is responsible for enhancing thought, concentration, mood, memory and and represses appetite. It is the key component in collagen formation. The L-form of all of the other amino acids is mostly beneficial to humans but the D and DL forms of Phenylalanine have been useful in treating pain including arthritic pain. It may be used as strong anti-depressant. It is used in the treatment of Parkinson's disease, a disease of brain.

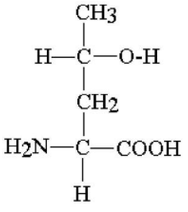


Phenylalanine

(vii) Threonine

Threonine is required for the formation of collagen, elastin and enamel proteins. It acts as a lipotropic agent which helps to prevent fatty deposits in the liver. It also aids in production of antibodies. Threonine can be converted to glycine in the central nervous system. It can acts as detoxifier. It is needed by the gastrointestinal (GI) tract for normal functioning. It provides symptomatic relief in ALS

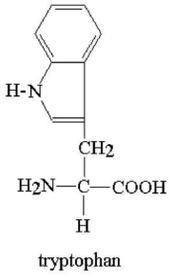
(Amyotrophic Lateral Sclerosis) and Lou Gehrig's disease. Threonine increases the weight of thymus as it was experimentally proved on animals. The level of threonine is quite low in patients with depression and so, in depressive patients, threonine is helpful in treating depression. Their codon includes ACC, ACU, ACA and ACG.



threonine

(viii) Tryptophan

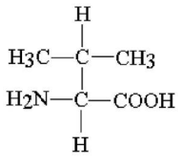
Tryptophan is the main precursor of neurotransmitter serotonin, which exerts a reassuring effect. It stimulates the transmission of impulses between nerve cells, contraction of the smooth muscles and in the regulation of cyclic body processes. It may be used as effective sleep supplement and reduces anxiety or nervousness. Tryptophan is very effective in treating some forms of depression, migraine headaches and insomnia. It also stimulates the production of growth hormone. Along with Lysine and Carnitine, it is also effective in lowering cholesterol levels. It can be converted into niacin (Vitamin B3). Tryptophan is the only plasma amino acid that is found in linkage to protein. Tryptophan must compete with 5 other amino acids to pass through the blood-brain barrier and enter the brain and those 5 includes-tyrosine, phenylalanine, leucine, isoleucine, and valine which are termed as Large Neutral Amino Acids (LNAA). It requires pyridoxal-5-phosphate (P5P) a form of vitamin B6 (pyridoxine) as a cofactor to be converted into serotonin. P5P deficiency will lower serotonin levels, even if Tryptophan levels are normal. Its codon is UGG.



tryptophan

(ix) Valine

Valine is one of the branched chain amino acids found in high concentration in muscle tissue which is absorbed and used directly by muscle as an energy source. It is processed by the liver after entering into the blood stream. Any acute physical stress including surgery, sepsis, fever, trauma and starvation requires higher amounts of Valine, Leucine and Isoleucine than any of the other amino acids. During phase of Valine deficiency, all of the other amino acids and protein are less well absorbed by the gastro-intestinal (GI) tract. It is used to treat severe amino acid deficiencies caused by addictions of drugs. Their codons include GUC, GUU, GUA and GUG.



valine

Check Your Progress

1. What is α-carbon?
2. Give an example of condensation reaction in amino acids.
3. Which functional groups determine the chemical properties of amino acids?
4. What is leucine?
5. What are the uses of tryptophan?

4.3 PROTEINS AND PEPTIDES

The polymer of amino acids is referred as protein. The term 'protein' was primary suggested by a Berzelius, a Swedish chemist and Mulder, a Dutch chemist in 1838. Another definition of protein can be as complex organic nitrogenous substances found in the cells of the living beings. Carbon, hydrogen, oxygen, nitrogen are the key constituent elements of proteins. Phosphorus and sulfur occurs as well in certain complex proteins. The fundamental composition of proteins in plants and animals presents has a great deal of variation. Protein comprise over half of the dry weight of most living organisms and they are most abundant intracellular macro-molecules. Proteins take up a primary position in the constructional framework and functioning of living matter. They are systematically connected with all phases of chemical and physical activity, that constitute the life of the cell. Some proteins play as essential structural elements of the body. For example-wool, hair and collagen which serves as an important constituent of connective tissue. The other proteins may be in form of enzymes, hormones or as oxygen-carriers. Enzymes are the proteins with catalytic activity which are mostly responsible for determining the phenotype or properties of a cell in a particular environment. The genotype constituent of living organism governs the production of different type of protein in the cell.

Proteins are a unique class of biomolecules in that they can recognize and interact with diverse substances. They contain complimentary clefts and surfaces which are designed to bind to specific molecules. Often only a single molecule or even a single stereoisomer can bind to a complimentary protein surface. Once this binding takes place, a complex is formed. This induces a conformational change which may act as a signal within the cell, or may serve to activate an enzyme. Many proteins, for example the enzymes ribonuclease A and chymotrypsinogen, contain only amino acid residues and no other chemical constituents; these are considered simple proteins. However, some proteins contain permanently associated chemical components in addition to amino acids; these are called

NOTES

conjugated proteins. The non–amino acid part of a conjugated protein is usually called its prosthetic group. Conjugated proteins are classified on the basis of the chemical nature of their prosthetic groups. For example, lipoproteins contain lipids, glycoproteins contain sugar groups, and metalloproteins contain a specific metal. A number of proteins contain more than one prosthetic group. Usually the prosthetic group plays an important role in the protein's biological function.

Proteins can be associated with membranes, and in fact carry out almost every membrane function. Interestingly, membrane proteins have special characteristics that allow them to exist in this lipid environment. Proteins that sit on the inner or outer surface of the membrane are called extrinsic or peripheral, and have a large percentage of hydrophobic amino acids in the portion of the molecule that is close to the hydrophobic membrane structure. The amino acids on the outer portion of the protein (facing the aqueous environment of the cytoplasm or extracellular fluid) are mostly hydrophilic, allowing the protein to be compatible with water. Proteins can also traverse the membrane, and in this case they are called intrinsic or integral. The portion of the protein that passes through the membrane is composed of hydrophobic amino acid residues, while the inner and outer portions exposed to water are largely hydrophilic. Transmembrane proteins can move laterally in the membrane, but cannot flip-flop.

Some proteins play a part in contraction of muscles and some are remain associated with the genes. The proteins structure has information which instructs them to regulate the performance of cell or tissue by the signaling pathway, intracellular organization, and catalytic activity and to control the function through interactions with other proteins including activators and inhibitors. The gram negative bacterium Escherichia coli (E.coli) are expected to contain about 5,000 various kinds of biomolecules which include around 3,000 different types of proteins and 1,000 different types of nucleic acid molecules. In humans, there may be around 1, 00, 000 different kinds of proteins, each with a distinctive structure. None of the proteins found in E.coli shows similarity with any of the human proteins. Thus, in about 1.5 million species diversity of living organisms, there are almost certainly 1010 to 1012 different kinds of protein molecules and about 1010 different kinds of nucleic acids.

Distribution of Amino Acids in Proteins

The distribution of all the twenty amino acids in all proteins is not similar. Fibroin, the protein presented in silk is rich in alanine nearly 30 per cent by weight. Glycine predominately accounts for almost 40% by weight in silk fibroin protein and 25% by weight in collagen protein. Casein and phosvitin protein have more content of Serine and threonine. The proteins which are rich in proline amino acid include collagen presented in connective tissue, gliadin in wheat and zein in corn. Human serum albumin consists of 585 amino acid residues has only one tryptophan moiety. The pulse contain good amount of basic amino acid, lysine but notably they lack sulfur-containing amino acid cysteine methionine, whereas, Cereals lack lysine but have sufficient quantity of methionine. When combination of these are taken in diet, they make good nutritive value lacking deficiency of required amino acid through mutual supplementation and are better utilized in human body.

NOTES

4.3.1 Chemical and Enzymatic Hydrolysis of Proteins to Peptides

Protein hydrolysis is carried out by various chemical and enzymatic methods. In this section we will discuss both these methods in detail.

Chemical Hydrolysis: Chemically, the proteins can be hydrolysed by the following methods:

1. **Acid Hydrolysis:** Acid hydrolysis is the most common method to hydrolyse a given protein sample. Acid hydrolysis can be carried out in both gaseous and liquid states. A wide range of acids can be used to carry out the acid hydrolysis but the most preferred acid for this process is 6 M HCl because the HCl is highly evaporative and due this, it can recover the hydrolytes even in the smaller portions of buffer. Also, the HCl being so versatile, it may be used both in the liquid as well as gaseous phase. The acid-hydrolysis reaction with 6 M HCl results in the addition of water to each covalent peptide bond, yielding the desired individual amino acids.
2. **Alkaline Hydrolysis:** Although, the acid hydrolysis is the most commonly used method of protein hydrolysis, alkaline or base hydrolysis is often employed for the hydrolysis of proteins. This method gives accurate quantification of tryptophan and is widely used for a variety of samples, from foods and feeds to peptides and proteins. The most widely used reagents for the chemical hydrolysis of proteins are NaOH and KOH. Alkaline hydrolysis is typically used only for tryptophan.

Enzymatic Hydrolysis: The enzymes used for protein hydrolysis are derived commonly from resources like animal sources (such as pancreatin and pepsin), plant sources (such as papain from papaya, ficin from fig, and bromelain from pineapple), and microbial sources (such as alcalase).

Proteolytic enzyme hydrolyses the proteins at the optimum temperature and pH and usually target specific peptide cleavage bonds, resulting in digestion consisting of amino acids and peptides of varying size. Enzymes from animal sources are more specific to their site of action compared to plant enzymes, which are more broadly specific in their action.

For example, pepsin cleaves the protein at the phenylalanine or leucine bond. Papain has a comparatively broad specificity, cleaving the bonds at phenylalanine, arginine, and lysine. Pancreatin cleaves the proteins at tryptophan, arginine, tyrosine, leucine, phenylalanine, and lysine bonds.

Enzymes from animal sources are more specific to their site of action compared to plant enzymes, which are more broadly specific in their action.

4.3.2 Proteins Containing other Chemical Groups

Many proteins such as the enzymes ribonuclease A and chymotrypsinogen contain only amino acid residues and no other chemical constituents, so they are referred as simple proteins. On the other hand, some proteins contain undyingly attached chemical components in addition to amino acids, these are called conjugated proteins**.** The non–amino acid fraction of a conjugated protein is generally referred as prosthetic group. On the basis of the chemical nature of their prosthetic groups,

NOTES

conjugated proteins are classified into various categories. For example, glycoproteins contain sugar groups, lipoproteins contain lipid molecules and metalloproteins contain a specific metal. A number of proteins contain more than one prosthetic group. Usually the prosthetic group plays an important role in the protein's biological function.

Check Your Progress

6. Who suggested the term 'protein'?
7. Give examples of proteins which are rich in proline amino acid.

4.4 STRUCTURAL ORGANIZATION OF PROTEIN

The tasks of describing and understanding structure of large macromolecules such as, proteins are approached at several levels of complexity which are therefore set in a category of theoretical chain of command. The four levels of protein structure were first defined by Linderstrom and Lang which includes primary structure, secondary structure, tertiary structure and quaternary structure. In mathematical term, these are also depicted as 1º, 2º, 3º and 4º, respectively.

4.4.1 Primary Structure of Protein

The linking of different amino acid residues in a polypeptide chain by covalent bonds mainly peptide bonds and disulfide bonds constitute the primary structure of protein**.** The fundamental primary structure of a protein is relatively simple and consists of one or more linear chains of a number of amino acid units (Figure 4.8). An example of primary structure is angiotensin II, which is the hypertensive octapeptide and has the sequence Asp-Arg-Val-Tyr-Ile-His-Pro-Phe. The peptide bond links α -carboxyl group of one amino acid residue to the α -amino group of the other amino acid is the major type of linkage of the amino acids found in proteins . The proteins may consist either of one or of more peptide chains.

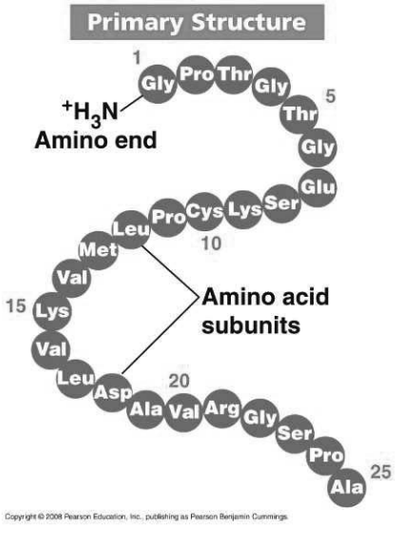


Fig. 4.8 *The Primary Structure of Protein*

Linus Pauling and Robert Corey in 1930, confirmed that α -carbons of adjacent amino acids are separated by three covalent bonds, arranged in C_α–C-N- C_α form. The studies of crystals of amino acids, simple dipeptides and tripeptides by X-ray diffraction studies confirmed that C-N bond present in the peptide is somewhat shorter than the C-N bond present in a simple amine. With this intention, they confirmed that the atoms associated with the peptide bond are coplanar due to a resonance or partial sharing of two pairs of electrons between the carbonyl oxygen and the amide nitrogen (Figure 4.9). The nitrogen atom has a partial positive charge and the oxygen has a partial negative charge, providing setting up a small electric dipole. The six atoms of the peptide group be positioned in a single plane, with the oxygen atom of the carbonyl group and the hydrogen atom of the amide nitrogen remaining trans to each other. Pauling and Corey from these findings concluded that the peptide C-N bonds are unable to rotate freely for the reason that of their partial double-bond nature, whereas rotation is allowed only about the N- C_α and the C- C_α bonds. The backbone of a polypeptide chain can be thus pictured as a sequence of rigid planes with consecutive planes sharing a common point of rotation at C_α. The rigid peptide bonds limit the range of conformations that can be assumed by a polypeptide chain.

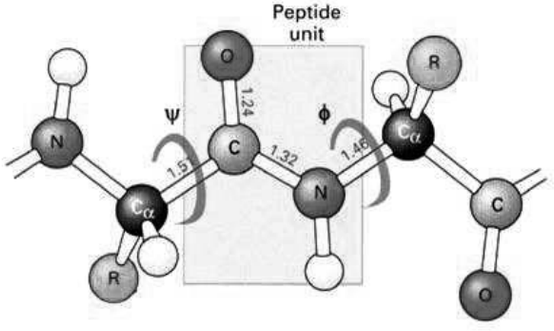


Fig. 4.9 *Rigid and Planar Peptide Bond*

In 1963, G. N. Ramachandran recognized that an amino acid residue in a polypeptide chain cannot have just any pair of values of ϕ and ψ. The bond angles resulting from rotations at C_α are labeled by convention as ϕ (phi) for the N- C_α bond and ψ (psi) for the C-C_α bond. When the polypeptide is in its fully extended conformation and all peptide groups are in the same plane, both ϕ and ψ are defined as 180° again by convention. In principle, the values of ϕ (phi) and ψ (psi) can lies between -180° and +180°. Many values are prohibited by steric interference between atoms in the polypeptide backbone and amino acid side chains. The conformation of the main polypeptide chain can be completely determined if the values ϕ and ψ for each amino acid residue in the chain are known.

By assuming that atoms behave as hard spheres, allowed ranges of ϕ and ψ can be predicted and visualized in steric contour diagram called Ramachandran plots (Figure 4.10). The values for ϕ and ψ are graphically revealed when ψ is plotted versus ϕ. One of them contains ϕ–ψ values that generate the antiparallel α sheet, the parallel Beet and the collagen helix. A second region has ϕ–ψ values that produce the right-handed α helix, the left-handed α helix. Left-handed helices are not found in proteins because they are much less favored energetically.

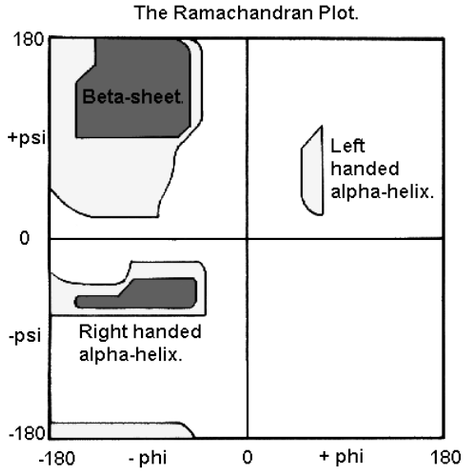


Fig. 4.10 *Ramachandran Plot*

4.4.2 Secondary Structure of Proteins: Forces Responsible for Holding of Secondary Structure- α-Helix, β-Pleated Sheets, Super Secondary Structure, and Triple Helix Structure of Collagen

Secondary structure of protein is formed by twisting or folding of polypeptide chain (primary structure of protein). These twisting and folding are because of interactions between atoms of polypeptide backbone which includes polypeptide chain (carboxyl group, amino group) but not R groups. The interaction found in the secondary structure of proteins are stabilized by the hydrogen bonding. The most common types of secondary structures are the α helix and the β pleated sheet. Both structures are held in shape by hydrogen bonds, which form between the carbonyl O of one amino acid and the amino H of another.

The hydrogen bonds form between the partially negative oxygen atom and the partially positive nitrogen atom. It is significant to point out that the only hydrogen bonds involved in secondary structure do not include any involving amino acid side chains. Most proteins have segments of their polypeptide chains that are either coiled or folded in patterns that contribute to the protein's shape. Many of these coils and folds repeat so often that they have been given names. Two folds that are extremely common in biochemistry are the alpha-helix and the beta-pleated sheet.

α-Helix Structure

α-helices are regular right-hand turns of amino acids 3.6 residues long, i.e., 5.4 Å. Hydrogen bonding between the first backbone carbonyl oxygen atom and the fourth residue NH group stabilizes the structure, van der Walls interactions across the axis further stabilize the structure.

There are some rare exceptions to this general scheme where hydrogen bonding can occur between three residues (3_{10}-helix) or between five residues (p-helix). These structures are much less stable than general α -helices and are not normally favoured.

Any of the 20 amino acids can participate in α -helix but some are more favoured than others. Ala, Glu, Leu, and Met are most often found in helices whereas, Gly, Tyr, Ser, and Pro are less likely to be seen. Proline, for instance, is rarely seen as its backbone nitrogen is bonded to its cyclic side group and cannot participate in hydrogen bonding. When prolines are found in α -helices, they tend to cause the helix to bend due to steric hindrance caused by its side group. They can be found as the first or last residue in the helix where they do not cause bending.

The side groups of the other amino acids point out and down relative to the helix. In globular proteins, those that are hydrophobic tend to be on one side of the helix and interact with other amino acids of the protein, and those on the other side are generally hydrophilic and interact with the solvent. For this reason, α -helices of globular proteins are predominantly found on the protein surface and have polar, hydrophobic, and hydrophilic amino acids. On average, α -helices in globular proteins have 11 residues, ~17Å long. Some α -helices have mainly hydrophobic residues, which are found buried in the hydrophobic core of a globular protein, or are transmembrane proteins.

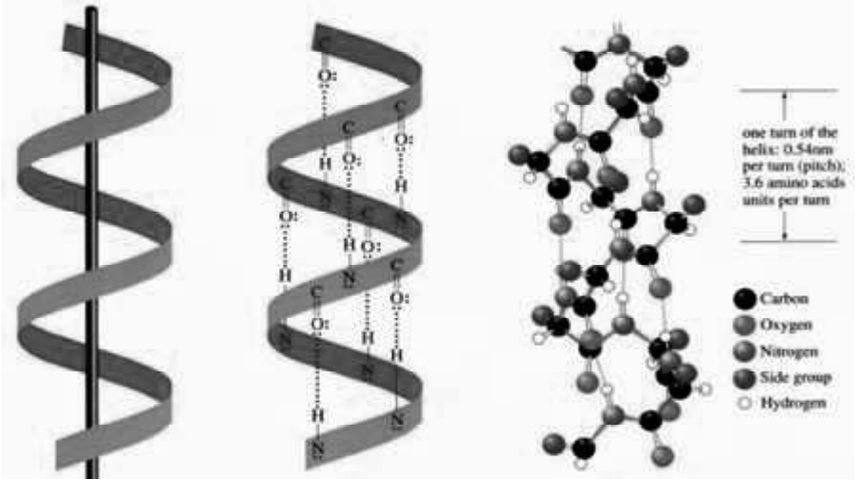


Fig 4.11 *Geometry of α-Helix Structure*

The α-helix adapted by homopeptides can be described as a structure element of a rather persistent rod, which assembles driven by structural anisotropy and dipole moment. The organization behaviour is broadened by peptides with defined monomer sequences.

Although the rigid-rod character is preserved, the helices resulting from sequence-defined peptides can be longitudinal polarized, laterally amphiphilic, or can have hydrophobic patches or sticky ends.

β-Pleated Sheet

β-pleated sheet secondary structure are generally fibrous, such as, silk, but pleated sheet is observed as a significant part of secondary structure in other proteins. They generally have rod-like shapes and are not so soluble in water. β sheets consist of β strands connected laterally by at least two or three backbone hydrogen bonds, forming a generally twisted, pleated sheet. Unlike the α-helix, β-sheets can involve

NOTES

one or more polypeptide chains–interchain or intrachain interactions. A β-strand is a stretch of polypeptide chain typically 3 to 10 amino acids long with the peptide backbone almost completely extended the R-groups stick up and down from β-sheets of alternating on either side of the strand.

Usually, small compact side chains like Gly, Ser, Ala. In an amphipathic β-sheet the amino acids alternate with polar/charged and non-polar (hydrophobic) amino acids. Thus, the hydrophobic residues are one side and the polar/charged are on the other. Stabilized by Hydrogen bonds (near perpendicular to direction of peptide backbone). Carbonyl of each amino acid is H-bonded to the NH of another amino acid. Sheets are formed by the interactions between parallel regions of a protein chain. These either run in the same direction, parallel; or in the opposite direction, antiparallel. These structures are stabilized by hydrogen bonds between backbone carbonyl oxygen atom and the hydrogen of the amino group.

In parallel β-sheets, the distances between the carbon and nitrogen involved in binding on one strand and on the other differ. This means that the hydrogen bonds are at an angle in relation to the protein strand. This is thought to make parallel b-sheets less stable than antiparallel b-sheets.

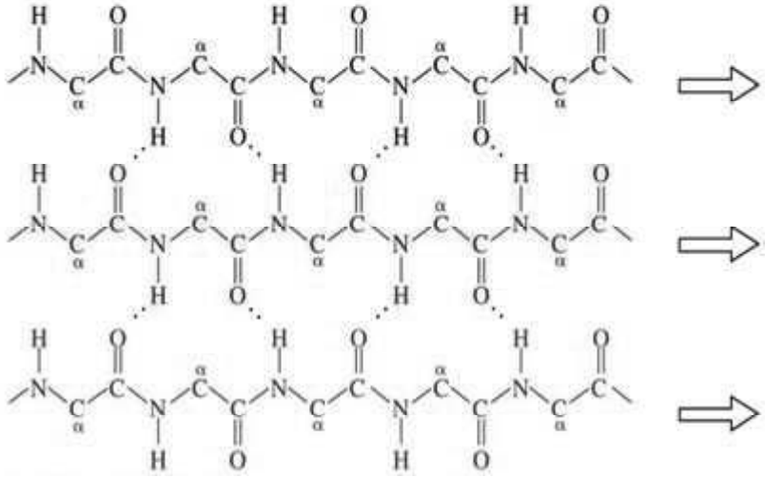


Fig 4.12 *Parallel β-Pleated Sheets*

In antiparallel b-sheets, the atoms on opposite strands involved in hydrogen binding are the same distance so that hydrogen bonds are at 90 degrees to the strand.

B-Sheets are not flat but have a pleated appearance due to the C atoms being successively above and below the plane of the sheet. The side groups are also successively above and below the plane of the sheet and them, therefore, cannot interact with each other. They do have significant interactions with neighbouring side chains and with their backbone. Proteins can contain parallel b-sheets, antiparallel b-sheets, or a mixture of both, although mixed proteins account for only 20% of proteins with b-sheets. The strands are typically 5–10 amino acids long and b-sheets contain 2–15 strands. The strands in b-sheets always have a right-handed twist.

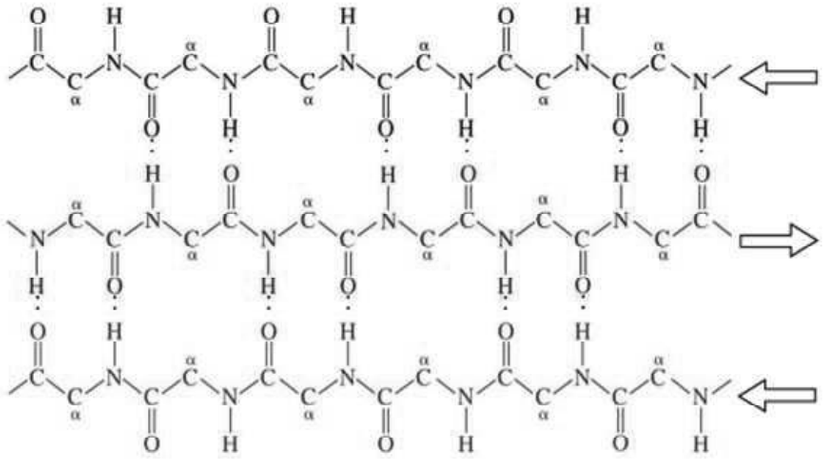


Fig. 4.13 *Antiparallel β-Pleated Sheets*

Super Secondary Structure

Many globular proteins contain combination of α-helix and β-pleated secondary structures. When this specific geometries of α-helix and β-pleated secondary structures are connected through loops are called super secondary structures, which are also called motifs. The various combinations of structure which can be formed out of these structures are αα- when two α-helices are linked with a loop, ββ- when two β-strands are linked by a loop, βαβ-two parallel β-strands connected by single α-helix, or more complex structures like Greek motif structure, beta barrel structure, etc.

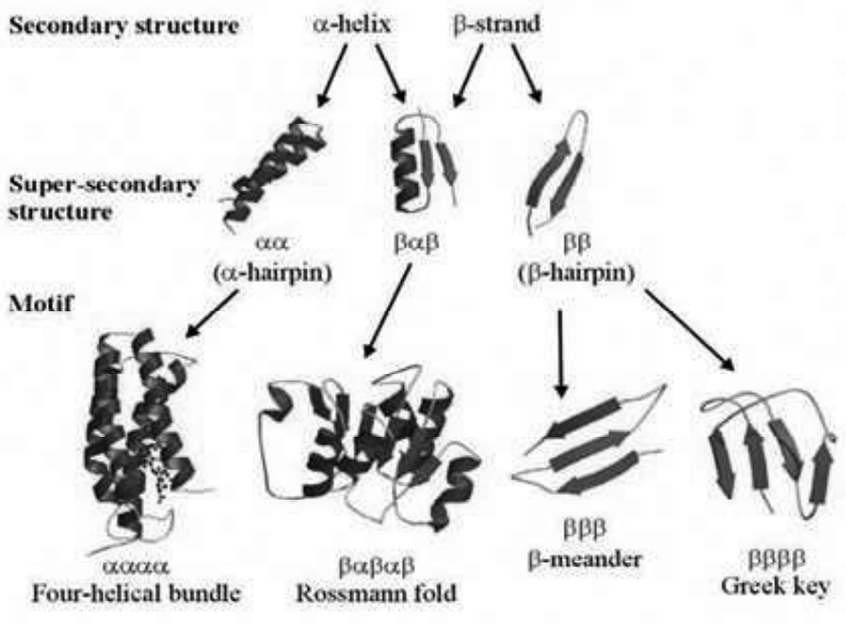


Fig. 4.14 *Examples of Super Secondary Structures in Hierarchal Form*

NOTES

Triple Helix Structure of Collagen

Collagen is the fibrous, structural protein found abundantly in the animals. It is the prime constituent of their various connective tissues like skin, tendons, ligaments, cartilage, bone, teeth, basement membranes, blood vessels, etc. They are almost one quarter of the total protein content found in most animals. The word 'collagen', is derived from a Greek word which means glue as primarily it was used to define that constituent of connective tissue which gives gelatin on boiling. But, as the research progressed, it was found that collagen is that component of connective tissue which cannot be extracted just by heating. The reason of this observation is the strong bonding in the form of crosslinking or covalent bond of collagen molecules with stable structures of some tissues. Due to this stable bonding, they form long rope like structure in tendons and are responsible for the high tensile strength present in the tendons. They also provide hard rigid structure to the bones and teeth by calcification of the interstitial space present between these molecules. The crosslinking between the fibers of them give two-dimensional flexible sheets in skin whereas the more complex arrangement of them in three dimension is found in cartilage.

Astbury and Bell in 1940 suggested the structure of collagen molecule as a single extended polypeptide chain with all amide bonds in the *cis* conformation. Its advancement was given in 1951 issue of the *Proceedings of the National Academy of Sciences* by him and his co-workers in the form of α-helix and β-sheet. Then, Pauling and Corey proposed the structure for collagen which is composed of three polypeptide strands held in a helical conformation by hydrogen bonds. Four to six main chain heteroatoms of the amino acid triplet are connected by hydrogen bonds and further, they are connected by two of the three peptide bonds to be in the cis conformation. The triple helix molecular structure of collagen was discovered using the early amino acid composition and sequence data and by the application of techniques like fiber diffraction analysis and model building.

The ordered arrangement of collagen molecules in the stretched tail tendon gave the highly focused fibre x-ray diffraction patterns of collagen, from this pattern, indexing lines in combination with the distinguishing amino acid features, directed Ramachandran & Kartha in 1954, to propose the advanced structure of collagen triple helix. Their proposed structure consisted of three left-handed PPII helices containing peptide bonds all in the trans conformation and two hydrogen bonds within each triplet which together form a right-handed triple helix. The structure received further refinement by Rich and Crick in 1955 and by North and co-workers in the form of the supercoiled triple-helical structure. This is the structure which is accepted today containing a single interstream $N–H_{(Gly)}...O=C_{(Xaa)}$ hydrogen bond per triplet and a tenfold helical symmetry with a 28.6-Å axial repeat (10/3 helical pitch).

Fraser et al. in 1979 further advanced the structure using linked atom least squares refinement on improved x-ray data. The fiber diffraction pattern characterizes an average over the whole collagen molecule, and the model obtained provides coordinates for an average Gly-Pro-Hyp tripeptide unit. The analysis

showed that the structure of collagen molecule consists of three left-handed supercoiled α-polypeptide chains which copy the conformation of polyproline-II chains coiled around a common axis. The structural arrangement of collagen into a triple helix is known as tropocollagen (length H" 300 nm).This polypeptide helices are held together by interchain hydrogen bonds. Type I, type II and III represent 90% of collagens. The three chains can be identical or different; in some molecule of collagen, the polypeptide chains are found to be same, while some may contain two or even three different chains represented as a1, a2 and a3. The difference is also seen in the amino acids present at X and Y positions of the triplets.

In a collagen molecule, three polypeptide chains each having the length of more than 1000 residue are present which consist of a characteristic triplet repeat sequence of $(Gly\text{-}X\text{-}Y)_n$. The reason of glycine placed permanently at every third position in the collagen chain is that, the centre of triple helix structure is very small and hydrophobic, close packing of the chains at the common axis put steric constraints on every third position and every third residue of the helix must have contact with the centre The contact is possible with glycine only as very little space is present in the centre and because of its smaller size its hydrogen bonds can easily interact with the centre in compare to bigger amino acids. So, glycine take place at every third position without chain distortion.

X and Y are commonly proline or hydroxyproline and 3- and 4-hydroxyproline respectively. These imino acids are found in a large proportion of about 20 %. The preference of imino acids on these positions is favorable because the fixed angle and restricted angle values are close to that found in the triple helix structure. The post translationally modification of Proline residue to hydroxyproline (Hyp) occurs on Y position as it increases the stability of the structure. It is carried out by prolyl hydroxylase and lysyl hydroxylase. Ascorbate or vitamin C act as a cofactor in this reaction. Hydroxylation is an important step which facilitate the inter hydrogen bonding. This also allows glycosylation of hydroxylysine residues. Deficiency of ascorbic acid causes scurvy, the disease which affects the structure of collagen. This occurs due to the impaired synthesis of collagen due to prolyl and lysyl hydroxylase deficiency.

A-chains are synthesized on the membrane bound ribosomes and then enter into the lumen of Endoplasmic reticulum. There, the selected proline and lysine residues are hydroxylated to form hydroxyproline and hydroxylysine respectively and some of the hydroxylysine residues are glycosylated. This α-chains then combine with two other chains by hydrogen bonds to form a complete triple- stranded helix molecules called tropocollagen or collagen molecule. Tropocollagen is secreted to the extracellular space to form collagen fibrils.

The position of the hydroxyl group hinders the thermal stability. Gly-Pro-Hyp sequence provide maximum stability to the collagen triple helix structure, while differences in the residues on the X and Y positions control global thermal stability and moderate local stability and energetics that are mandatory for self-association, recognition, and binding. Type I collagen has an equal amount of Pro and Hyp, while most collagens have more Hyp than Pro residues. As glycine and proline are

NOTES

present abundantly in the structure, the regular α-helix and β-sheet structure is not formed.

So, to form a right-handed triple helix, three left-handed helical strands twist around an axis. In a collagen triple helix rise of the 2.9 Å (0.29 nm) per residue is there and 3.3 residues per turn are present. The three chains are bonded by hydrogen bonds between NH group of glycine which act as a donor groups and CO group on the other chain which act as hydrogen bond acceptors. Hydrogen bonds are essential component of this structure as they hold the triple helix together and their absence in some natural collagen can lead to various pathological condition.

Covalent crosslinks are formed within a tropocollagen molecule and between different molecules. The intramolecular crosslinks are formed by lysyl oxidase, which is a copper dependent enzyme that oxidatively deaminates the [-amino groups of lysine residues, yielding reactive aldehydes of allysine residues. Then, such aldehydes of two side chains link covalently in a spontaneous non-enzymatic aldol condensation. Histidine may also be involved in some cross links. The intermolecular cross-linking of tropocollagen molecules involves the formation of unique hydrooxypyridinium structure from one lysine and two hydroxylysine residues.

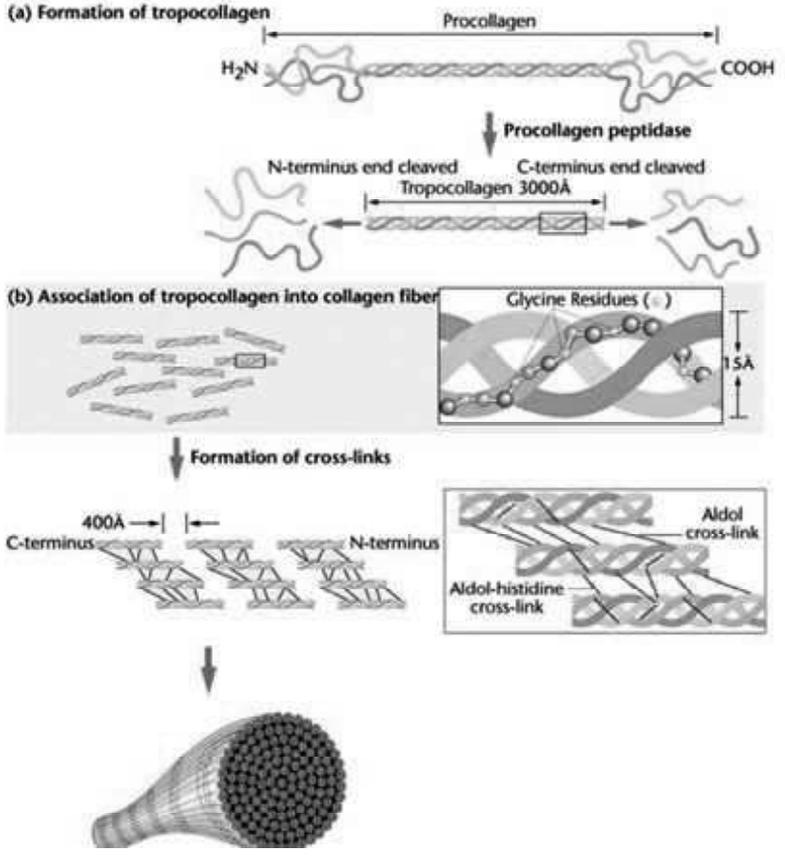


Fig. 4.15 *Separate Triple Helices Arranged to Form Fibrils*

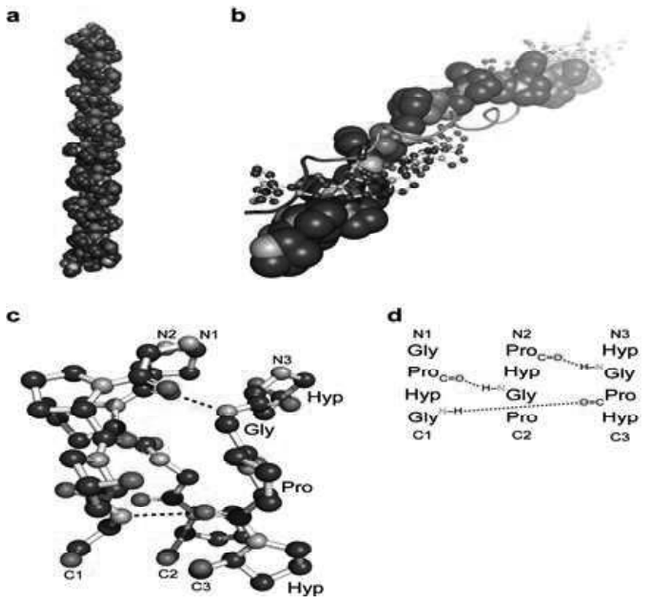


Fig. 4.16 *Collagen Triple Helix*

Figure 4.16 shows:

a) Crystal structure of a collagen triple helix, formed from $(ProHypGly)_4$–(ProHypAla)–$(ProHypGly)_5$

b) Axis of a $(ProProGly)_{10}$ in triple helix structure with the three strands depicted in space-filling, ball-and-stick, and ribbon illustration.

c) Ball-and-stick picture of a segment of collagen triple helix with highlighted interstrand hydrogen bond ladder.

d) Stagger of the three strands in the segment in panel *c.*

Hydroxyproline's OH group does not participate in hydrogen bonding but stabilize the entire complex by stabilizing trans isomer of proline. The chains are stabilized due to the steric repulsion between pyrrolidine rings of proline and hydroxyproline residues. The pyrrolidine rings retain out of each other's way once polypeptide chain assumes long helical form, which is comparatively more open than the tightly coiled form of the alpha helix. So, the whole process of crosslinking and formation of fibrile can be summarized as:

- Synthesis of polypeptide and its entry into the lumen of endoplasmic reticulum.
- Hydroxylation of prolyl and lysyl residues.
- Glycosylation of the residues.
- Formation of tropocollagen.
- Packaging into transport vesicle.
- Exocytosis.
- Lateral covalent crosslinking of tropocollagens.
- Aggregation of fibrils.

NOTES

As per the conducted studies, there are at least 27 collagen types in vertebrates with 42 different polypeptide chains (designated as type I to XXVII)) and more than 20 other proteins with collagen like domains. Different types of collagens carry out different specialized functions in various tissues and have different supramolecular organization. The molecules can be homotrimers, heterotrimers with two or three distinct chain types.

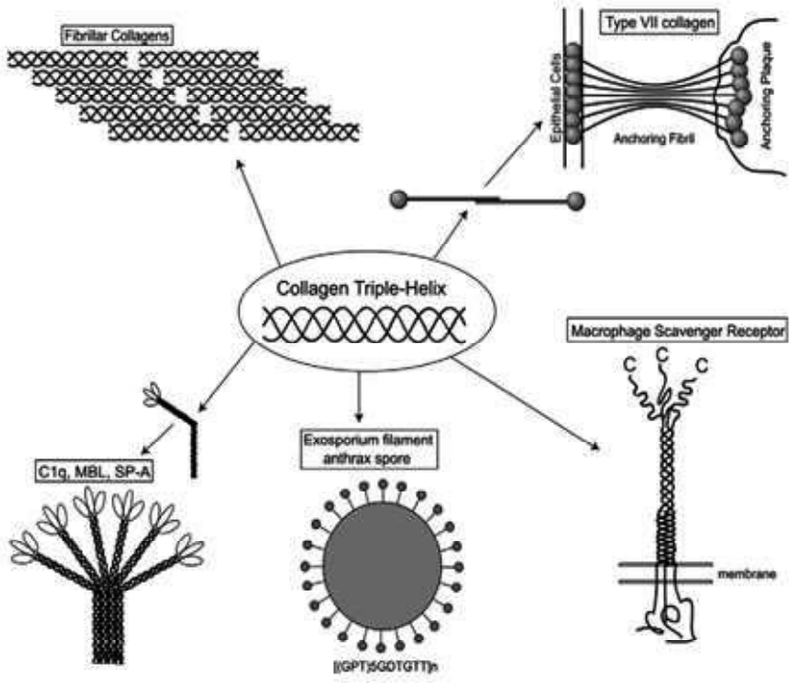


Fig. 4.17 *Some Biological Forms of Collagen Triple Helix Domain*

Fibrils are the most commonly found collagen with major types I, II, III, and minor types V and XI. They make the structural basis for skin, tendon, bone, cartilage, and other tissues. Collagens found on the surface of fibrils are FACIT types IX, XII, XIV, XVI, XIX, XX, XXI, XXII, XVI; and within the basement membrane networks are type IV collagen, in hexagonal networks types VIII and X, as beaded filaments type VI, in the anchoring fibrils of skin type VII, or as membrane proteins types XIII, XVII, XXIII, XXV.

4.4.3 Tertiary Structure of Protein-Folding Domain Structure and Quaternary Structures of Protein

Tertiary structure of protein refers to the arrangement of amino acids that are far apart in the chain. Each protein ultimately folds into a three dimensional shape with a distinct inside and outside. The interior of a protein molecule contains a preponderance of hydrophobic amino acids, which tend to cluster and exclude water. The core is stabilized by Van der Waals forces of interaction and hydrophobic bonding. By contrast, the exterior of a protein molecule is largely composed of hydrophilic amino acids, which are charged or able to form hydrogen bond with water. This allows a protein to have greater water solubility**.** It indicates in three-dimensional space, how secondary structural features such as sheets, helices, bends,

turns and loops assemble to form domains and how these domains relate spatially to one another. A domain is a section of protein structure sufficient to perform a particular chemical or physical task such as binding of a substrate or other ligands. Other domains may anchor a protein to a membrane or interact with the regulatory molecule that modulates its function.

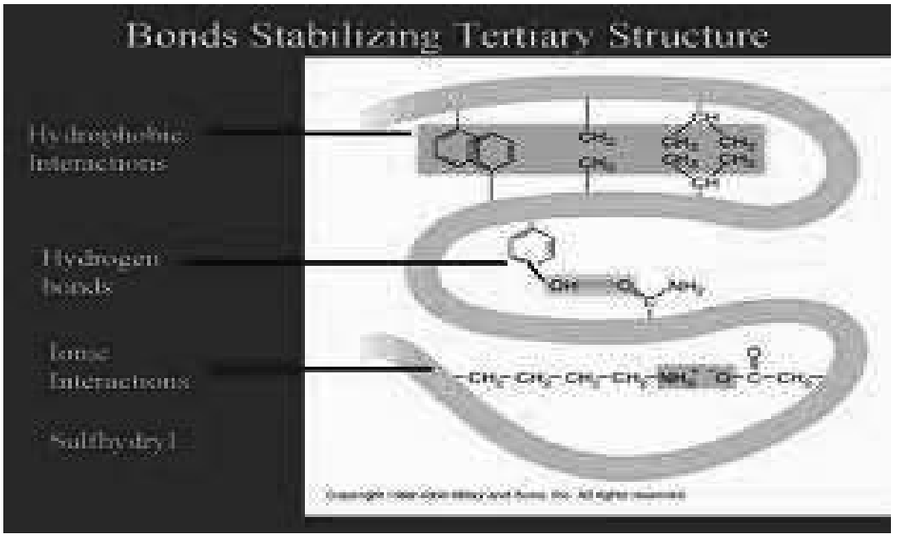


Fig. 4.18 *Tertiary Structure of Protein*

Amino acids that remain far apart in the polypeptide chain can exist in different types of secondary structure which may intermingle within the completely folded structure of a protein. The site of bends in the polypeptide chain including α-turns plays a crucial role in stabilizing the protein structure. The direction and angle of these bends are determined by the number and location of specific bend-producing amino acid residues, such as proline, threonine, serine and glycine. The interacting segments of polypeptide chains are held in their characteristic tertiary positions by different kinds of weak bonding interactions and sometimes by covalent bonds such as disulfide cross-link between the segments.

When a protein has two or more polypeptide subunits, their arrangement in space is referred to as quaternary structure (Figure 4.19).The proteins containing two or more separate polypeptide chains or subunits may be identical or different. .It defines the polypeptide composition of a protein and, for an oligomeric protein, the spatial relationships between its unit or protomers. The relatively large size of proteins reflects their functions. The function of an enzyme requires a stable structure containing a compartment large enough to bind its substrate and catalyze a reaction. Protein size has limitation however, forced by two factors. The first factor includes the genetic coding ability of nucleic acids and the precision of the protein biosynthetic process.

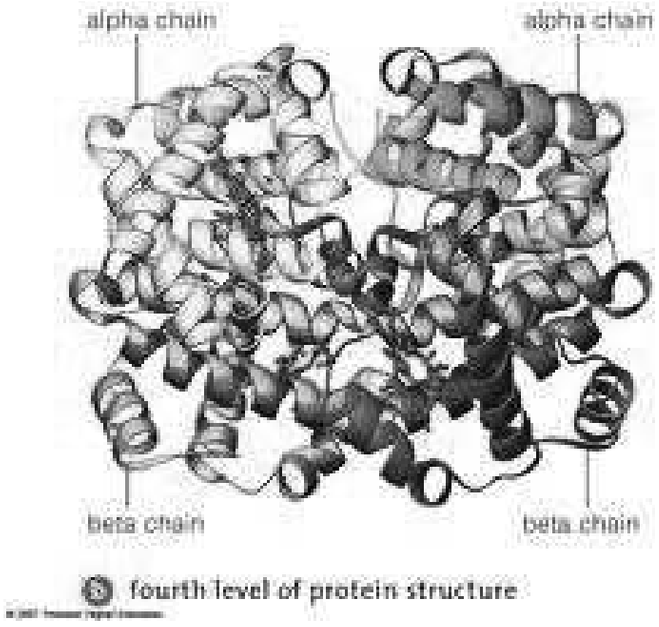


Fig. 4.19 *Quaternary Structure of Protein*

Check Your Progress

8. Who first defined the four levels of protein structure?
9. How is the secondary structure of proteins formed?
10. Where is the word 'collagen' derived from?
11. What is the difference between tertiary and quaternary structure of protein?

4.5 CLASSIFICATION OF PROTEINS

It is useful to classify proteins into two major groups for the considering these higher levels (quaternary) of structure which includes fibrous proteins and globular proteins.

4.5.1 Fibrous Proteins

Fibrous proteins are those proteins containing polypeptide chains arranged in the form of long strands or sheets, whereas, globular proteins are those proteins having polypeptide chains folded into a spherical or globular shape. These two groups of proteins are structurally dissimilar. Fibrous proteins usually consist largely of a single type of secondary structure. α- keratin, collagen, and silk fibroin are some examples of fibrous protein which properly demonstrate the relationship between protein structure and biological function. Fibrous proteins contribute to properties that provide strength or elasticity to the structures in which they occur.

All the fibrous proteins are mostly insoluble in water, a property conferred by a high concentration of hydrophobic amino acid residues both on its surface and in the interior of the protein. These hydrophobic surfaces are mostly covered by

packing numerous similar polypeptide chains together to produce complicated supra-molecular complexes. The fundamental structural simplicity of fibrous proteins makes them predominantly practical for illustrating some of the fundamental principles of protein structure. We start our discussion with fibrous proteins, before turning to the other complex folding patterns founded in globular proteins.

α-Keratin

The crucial structural unit of α-keratin generally consists of three right-handed helical polypeptides that are stabilized by cross linking disulfide bonds in a left-handed orientation. The α-keratins have evolved for providing strength and are found mostly in mammals. These proteins constitute almost the entire dry weight of hair, wool, nails, claws, quills, horns, hooves, tortoise shell and much of the outer layer of skin. The α-keratins are part of a broader family of proteins called intermediate filament (IF) proteins. The α-keratin helix is right-handed helixes which are as the same helix found in many other proteins. Francis Crick and Linus Pauling in the early 1950s suggested both separately that the α-helices of α-keratin were arranged as a coiled coil form. Two strands of α-keratin are oriented in parallel i.e. with their amino termini (N) at the same end, which are wrapped about each other to form a supertwisted coiled coil structure. This supertwisting structure amplifies the strength of the overall structure of α-keratins (Figure 4.20), just as strands are twisted to make a strong rope. The helical path of the supertwists is left-handed, opposite in sense to α helix. The surfaces where the two α helices touch are made up of hydrophobic amino acid residues, their R groups meshed together in a regular interlocking pattern. This permits a close packing of the polypeptide chains within the left-handed supertwist. α-keratin is rich in the hydrophobic amino acid residues which include alanine, valine, leucine, isoleucine, methionine and phenylalanine. The hardest and toughest α-keratin occurs in mammals such as those of rhinoceros horn, constitute up to 18% of sulfur containing amino acid residues such as cysteine which forms disulfide bonds in the polypeptide chain.

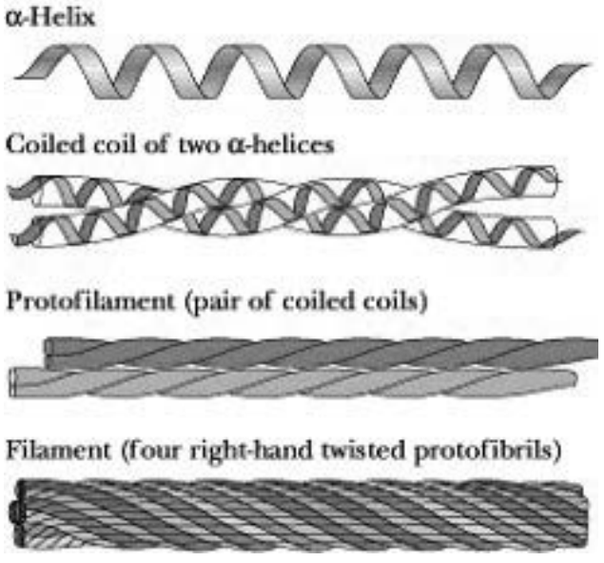


Fig. 4.20 *Structure of α-Keratin*

NOTES

Collagen

As alredy discussed, Collagen like the α-keratins has evolved to provide strength. It is found in connective tissue such as tendons, cartilage, and organic matrix of bone and cornea of the eye. The collagen helix is a unique secondary structure quite distinct from α helix. It is left-handed and has three amino acid residues per turn.

Silk Fibroin

Silk Fibroin is the protein of silk and is chiefly produced by insects and spiders. The polypeptide chains of this protein are predominantly occurs in the α-conformation. Fibroin is a type of fibrillar proteins which is a member of a class called α-keratins. Silk fibroin (Figure 4.21) is composed mainly of glycine, alanine and serine amino acid residues linked together by the peptide bonds. Glycine constitutes approximately 45 per cent of the total amino acid residues whereas alanine in addition to serine composes another 42 per cent of the total amino acid residues. Therefore, the R groups extending above and below the plane of the α-pleated sheet are small and thus, allow the pleated sheets to stack or hoard. At this point, the two chains of this protein run in opposite directions which are linked by hydrogen bonds but it may be noted that each of the two chains can also form hydrogen bonds with another protein and so on. The relatively small R groups of alanine, glycine and serine facilitate the formation of very large protein aggregates. If the R groups are too large, the hydrogen bonding may not occur as the chains are held far away from each other for union.

Fig. 4.21 *Structure of Silk Fibroin*

4.5.2 Globular Proteins

Globular proteins frequently contain several types of secondary structure. Most enzymes and regulatory proteins are types of globular proteins which includes cytochrome *c,* lysozyme, ribonuclease and myoglobin. With illumination of the

tertiary structures of hundreds of other globular proteins by x-ray analysis, it became comprehensible that myoglobin illustrates there is only one of many ways in which a polypeptide chain can be folded. The structures of cytochrome *c,* lysozyme, and ribonuclease are compared and these proteins have different amino acid sequences and different tertiary structures, reflecting differences in their function as discussed below in detail. All of these proteins are relatively small in size and so, are suitable to work with to analyze and predict the structural and functional difference.

Myoglobin

Myoglobin is a heme protein which is found mainly in muscle cells. Myoglobin is first globular protein to have its 3-D structure elucidated by x-ray diffraction studies which was accomplished by John C. Kendrew in 1959. Myoglobin molecule (Figure 4.22) contains a single polypeptide chain of 153 amino acid residues and a single prosthetic iron-porphyrin or heme group, identical with that of hemoglobin. The heme group presented in of myoglobin and as well as in hemoglobin is responsible for providing the deep red-brown color to muscles tissue and blood. Myoglobin is the oxygen binding protein which is especially abundant in the muscles of such as the whale, seal and porpoise. The diving mammals have brown muscles due to presence of this protein in abundant amount. Storage of oxygen by muscle myoglobin permits these diving animals to remain submerged in water for long periods. The chief function of myoglobin is to bind oxygen in the muscles and to enhance transportation to the mitochondria which consume oxygen during respiration.

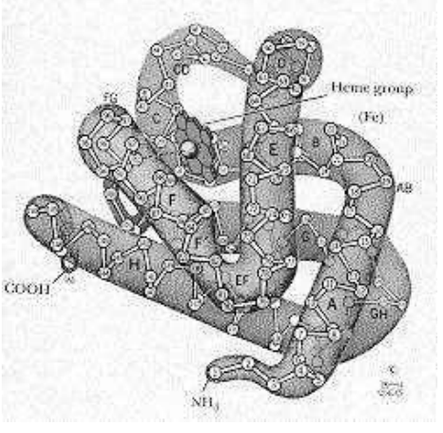


Fig. 4.22 *Structure of Myoglobin*

Cytochrome *C*

Cytochrome *c* is a major constituent of the respiratory chain presented in mitochondria. Cytochrome *c* is a heme protein and it contains a single polypeptide chain of about 100 amino acid residues with a single heme group. The other example of heme protein includes myoglobin and hemoglobin. The protoporphyrin ring structure of the heme group presented in cytochrome c is covalently attached to the polypeptide chain. Only about 40 per cent of the polypeptide is found in α-helical segments while 70 per cent of the polypeptide is found in α-helical

segments in case of the myoglobin chain. The remaining of the cytochrome *c* chain contains α turns structure, irregularly coiled and extended segments. Cytochrome *c* (Figure 4.23) is covalently associated to the protein on two sides, providing significant stabilization to the entire protein structure. The hydrophobic amino acid side chains are oriented toward the interior, i.e., away from water and the hydrophilic amino acid side chains remain on the surface layer. The structures are also stabilized by a multitude of hydrogen bonds and some ionic interactions.

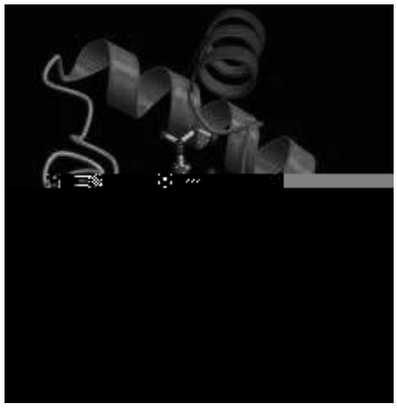


Fig. 4.23 *Structure of Cytochrome C*

Lysozyme

Lysozyme is an enzyme found abundant in human tears and white part of the egg. Lysozyme catalyzes the hydrolytic cleavage of polysaccharides presented in the protective cell walls of some species of bacteria. It can serve as a bactericidal agent as it can lyse or degrade the bacterial cell wall constituents. Lysozymes are also comprised of both alpha and beta pleated structure (Figure 4.24) showing similarity with cytochrome c structure. About 40 per cent of its 129 amino acid residues are in α-helical segments, but the arrangement is different as compared to cytochrome c. The stability to this protein structure is maintained by four disulfide bonds. The helices processions having a long gap in the side of the molecule are termed as active site, which is the site of substrate binding and catalysis. The bacterial polysaccharide that is the substrate for lysozyme fits into this active site crevice.

Fig. 4.24 *Structure of Lysozyme*

NOTES

Ribonuclease

Ribonuclease is another small globular protein. Ribonuclease (Figure 4.25) is an enzyme secreted by the pancreas into the small intestine where it catalyzes the hydrolysis of certain bonds presented in ribonucleic acids (RNA) of ingested food. It is also known as RNase. Two groups of workers Christian Anfinsen and his associates at the National Institute of Health and William Stein and Stanford Moore at Rockfeller Institute) have elucidated the complete structure of this pancreatic protein .Its tertiary structure, determined by x-ray analysis, shows that little of its 124 amino acid polypeptide chain is in an α-helical conformation, but it contains many segments in the α conformation. Ribonuclease has four disulfide bonds between loops of the polypeptide chain. In small proteins, hydrophobic residues are less likely to be sheltered in a hydrophobic interior—simple geometry dictates that the smaller the protein, the lower the ratio of volume to surface area. Small proteins also have less prospective of weak interactions available to stabilize them. This explains why many smaller proteins are stabilized by a number of covalent bonds.

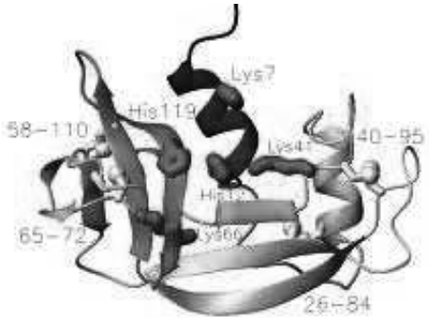


Fig. 4.25 *Structure of Ribonuclease*

Hemoglobin

Hemoglobin is a heme- protein belonging to the myoglobin-hemoglobin family. It plays crucial role as the oxygen transporter in erythrocytes. It constitutes about 90 per cent of the protein of red blood corpuscles or erythrocytes. Hemoglobin is the first oligomeric protein for which the three-dimensional structure was elucidated. It is written in short form as Hb and is a tetrameric protein due to four polypeptide chains content. The x-ray diffraction analysis has revealed that the hemoglobin molecule is nearly spherical in shape with a diameter of about 55 Å (5 .5 nm). Human hemoglobin protein (Figure 4.26) consists of four polypeptide chains of consisting of two types, two α-chains and two α-chains. The polypeptide portion is communally called as globin. The- chain has valine amino acid residue at the N-terminal and arginine at the C-terminal whereas in the α-chain, valine is situated at the N-terminal and histidine at the C-terminal. Each α-chain is in contact with both α chains. In contrast, there are a small number of interactions between the two α-chains or between the two α-chains. Each chain has a heme prosthetic group in a gap near the exterior of the molecule. The heme groups of hemoglobin are concerned in the binding of oxygen. The α-chain has 141 amino acid residues and the α-chain has 146 amino acid residues. The α-chain is more acidic than α-chain.

NOTES

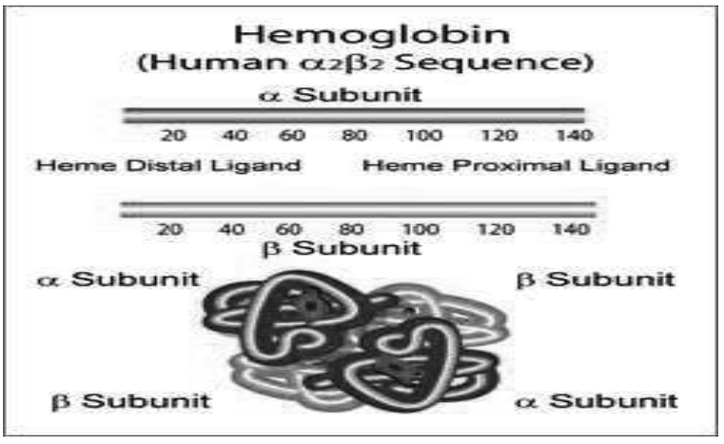


Fig. 4.26 *Structural Organization of Hemoglobin*

Each of the four chains has a characteristic tertiary structure, in which the chain is folded. Like myoglobin, the α- and α-chains of hemoglobin contain numerous segments of α-helix, separated by bends. The α- and α-chains are held together as a pair by ionic and hydrogen bonds. These two pairs are then joined to each other by additional ionic bonds hydrogen bonds and the hydrophobic forces. Thus, the four polypeptide chains fit together almost tetrahedrally to form the characteristic quaternary structure. The hemes are 2.5 nm apart from each other and tilted at different angles. Each heme is partly buried in a pocket lined with hydrophobic R- groups of amino acid. It is bound to its polypeptide chain through a coordination bond of the iron atom to the R group of a histidine residue. The sixth coordination bond of the iron atom of each heme is available to bind a molecule of oxygen. The α and α chains of hemoglobin have nearly the same tertiary structure as myoglobin. Both have well over 70% α-helical nature, both have analogous lengths of α-helical segments and the bends also have near about the same angles.

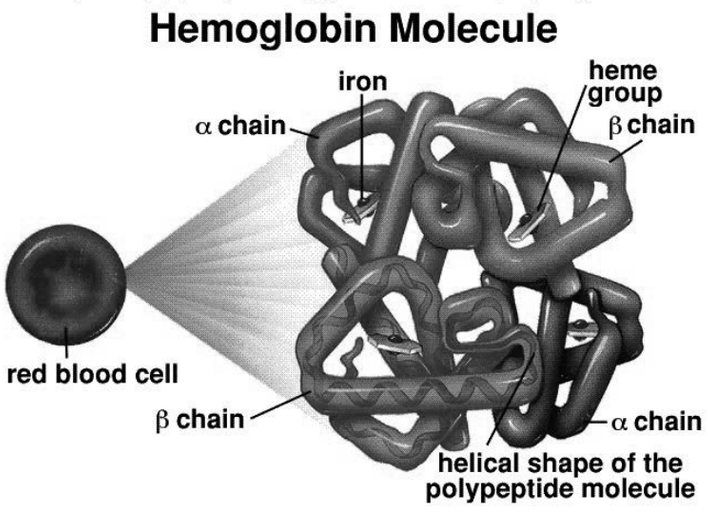


Fig. 4.27 *Constitute of Hemoglobin from RBC*

NOTES

Allosteric Nature of Hemoglobin

The binding of an O_2 molecule to hemoglobin has consequent influence on the subsequent oxygenation of the other heme groups of the molecule which is an example of allosteric interactions. Allosteric interactions are the type of an interaction in which one site on a protein affects another site located distinctly in a different region of the same molecule. In addition to hemoglobin, many enzymes are allosteric proteins. A quaternary structure (Figure 4.28) in common is the feature of allosteric proteins.

Hemoglobin is a much more complex and alert molecule than is myoglobin. Hemoglobin, in addition to transporting oxygen also transports CO_2. CO_2 is a waste product of metabolism and is transported to the lungs to be respired. The ability of hemoglobin to bind H+, (another waste product of metabolism) is also a significant component for sustaining physiological function of the macromolecule, as it is critical for the maintenance of physiological pH. The oxygen-binding properties of hemoglobin are regulated by interactions between separate, nonadjacent sites. Hemoglobin is an allosteric protein, whereas myoglobin is not. This difference is expressed in three ways.

1. The binding of O_2 to hemoglobin enhances the binding of extra O_2 to the same hemoglobin molecule. In other terms, oxygen binds cooperatively to hemoglobin but the binding of O_2 to myoglobin is not supportive.
2. The attraction of hemoglobin for O_2 is pH-dependent, whereas that of myoglobin is independent of pH.
3. The oxygen affinity of hemoglobin is further maintained by organic phosphates such as 2, 3- Bisphosphoglycerate (BPG), whereas that of myoglobin is not. Thus, hemoglobin has a lesser affinity for oxygen than does myoglobin.

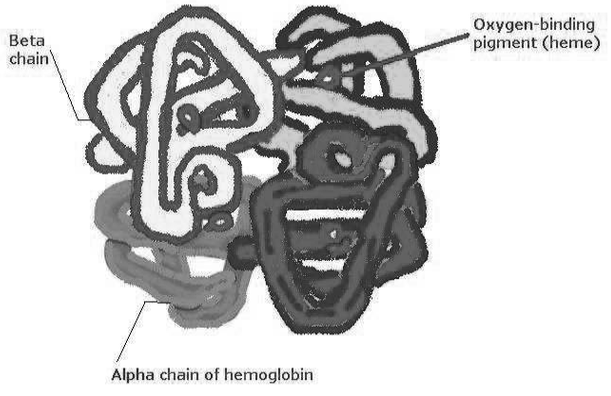


Fig. 4.28 *Oxygen Binding Site on Hemoglobin*

Check Your Progress

12. What do you mean by fibrous proteins?
13. What is the composition of silk fibroin?

NOTES

4.6 AMINO ACID METABOLISM: DEGRADATION AND BIOSYNTHESIS OF AMINO ACIDS

The sources of amino acids in the body are as follows:

- Dietary protein
- Intercellular synthesis
- Tissue protein breakdown

The amino acids are metabolized as pass in the portal system to the liver and determination of portion of these amino acids takes place there. The amino group and carbon residues may be used for the synthesis of other amino acids or other nitrogenous compounds. The reminders are used for protein synthesis in the cells. Amino acids metabolism takes places in following two phases:

1. **Anabolic Phase:** It involves synthesis of protein such as enzymes, hormones, blood proteins, tissue proteins, etc. Synthesis of non-protein nitrogenous substances such as purines, creatine, pyrimidines, choline, etc.
2. **Catabolic Phase:** It involves the following reactions:
   - **Transamination:** The transfer of an amino ($-NH_2$) group from an amino acid to a keto acid by enzyme catalysed reaction known as transamination. At the end of the reaction non-essential amino acids are formed.
   - **Deamination:** The removal of amino group from the amino acid as NH_3 is known as deamination. It results in the formation of keto acids and ammonia.
   - **Oxidative Deamination:** This reaction takes place mostly in liver and kidney. It is liberation of free ammonia from amino group to amino acids coupled with oxidation with the formation of α-Keto acids.

4.6.1 Anobolic Phase: Amino Acid Biosynthesis

Amino acids are the building units of proteins and there are twenty protein amino acids. Besides, there are several nonprotein amino acids. All amino acids consist of an amino group and a carboxyl group bonded- to a carbon atom i.e., the so-called alpha carbon. In addition, there is an alkyl group (R) which varies in structure. The structure of all the amino acids, is basically the same, but they differ with respect to their side groups.

Because of such differences, amino acids may be nonpolar (polar but with two charges balancing one another out so that the amino acid as a whole is unchanged); positively charged (acidic); or negatively charged (basic). The nonpolar molecules are not soluble in water, whereas the charged and polar molecules are soluble; The synthesis of an amino acid involves interesting pathways and it is connected with the Kreb cycle and other pathways of general nitrogen metabolism.

NOTES

Besides, the study of their biosynthesis has revealed regulatory mechanisms of high effectiveness. As a matter of fact, many of the initial studies of end product inhibition of metabolic pathways were performed using bacterial amino acid synthesis.

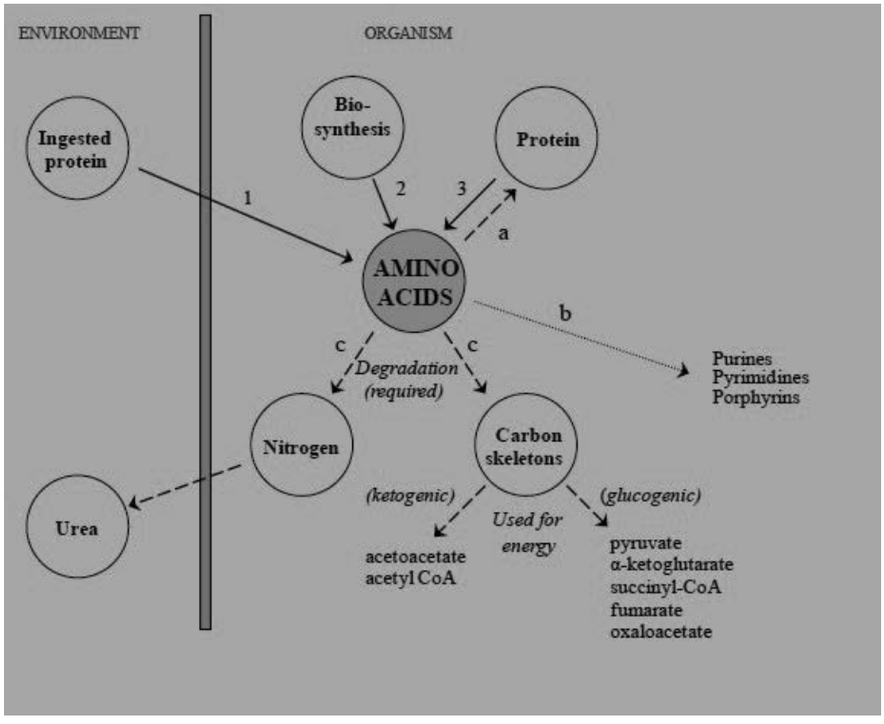


Fig. 4.29 *Overview of Amino Acid Metabolism*

Flow Sheet of Amino Acid Oxidation

Amino acids form the class of biomolecules whose oxidation contributes significantly to the generation of energy: During normal synthesis and degradation of protein, when the diet is rich in protein, amino acids are not stored. Figure 4.30 shows the flow sheet of amino acid oxidation. Most amino acids are metabolized in the liver. Some of the ammonia generated is used in biosynthesis; the excess is excreted. Excess ammonia generated in extra hepatic tissues is transported to liver for conversion to the appropriate excreted form. This process of conversion of excess amino acids to carbohydrates for energy production is known as amino acid pool.

Removal of the α-amino groups occur in the cytosol by transamination reactions catalysed by aminotransferases (transaminases) as follows:

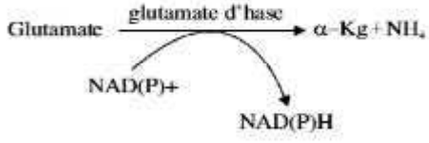
$$\alpha\text{-keto acid}_2 + \text{amino acid}_1 \rightarrow \alpha\text{-keto acid}_1 + \text{amino acid}_2$$

In liver, α-keto acid is usually α-ketoglutarate (AKG); in muscles, it is usually pyruvate. Serum Glutamic Pyruvic Transaminase (SGPT) and Serum Glutamic Oxaloacetic Transaminase (SGOT) are sensitive indicators for a number of disease conditions. During heart attacks, damaged heart cells leak aminotransferases. Damaged liver cells also leak aminotransferases. SGPT and SGOT levels should be monitored in people exposed to industrial chemicals. The

NOTES

effect of transamination is to collect amino groups from many amino acids and convert them into one, glutamate. Glutamate channels amino groups into biosynthetic pathways or into reactions where nitrogenous waste products are formed. Glutamate channels amino groups into biosynthetic pathways. The amino groups are removed from glutamate and prepared for excretion. Glutamate is transported into the mitochondrial matrix where it undergoes oxidative deamination catalysed by glutamate dehydrogenase. Glutamate is transported into the mitochondrial matrix where it undergoes oxidative deamination catalyzed by glutamate dehydrogenase.

In the liver, glutamine is converted back to glutamate by glutamines. Glutamine is the major transport form of ammonia. It is present in blood in much higher concentrations than other amino acids. Alanine also plays a role in transport of amino groups to the liver by glucose-alanine cycle.

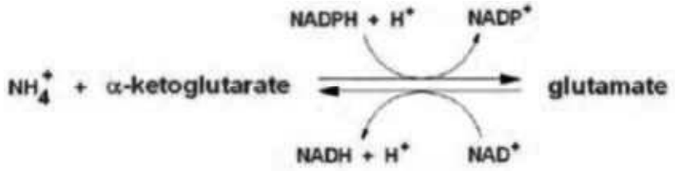


Reductive Amination

Cells exhibit different abilities to synthesize amino acids. For example, some bacteria, including *Escherichia Coli*, can synthesize all the 20 amino acids from simple inorganic sources. On the other hand, animal cells, as well as those of many species of microorganisms must receive about half of the 20 common amino acids in their diet, as they do not possess the enzymatic machinery to synthesize them. Based on this fact, the amino acids could be categorized into the following groups:

1. Essential amino acids
2. Non-essential amino acids

Essential amino acids are those that the organisms are unable to synthesize. Therefore, the organism must be able to get them in a readymade form from external sources in diet. One of the most important nonessential amino acids is glutamic acid which is very close to a Kreb cycle precursor. This amino acid is readily formed from the α-ketoglutaric acid by a single reductive reaction catalysed by the enzyme glutamate dehydrogenase. Besides, experiments with labelled nitrogen have shown that during the early stages of nitrogen assimilation, glutamic acid is by far the most labelled compound. This indicates that ammonia is directly incorporated into α-ketoglutaric acid by a process called reductive amination. The equation for reductive amination is as follows:



The reaction is reversible, so that glutamate can serve as an energy source by entering the Kreb cycle at the level of -ketoglutarate and requires the presence of reduced nicotinamide adenine dinucleotide phosphate ($NADPH^+ H^+$). This reaction of reductive amination is of central importance because of the high proportion of glutamate formed in this manner (Figure 4.30). It is actually the major 'port of entry' of ammonia into the metabolic reactions.

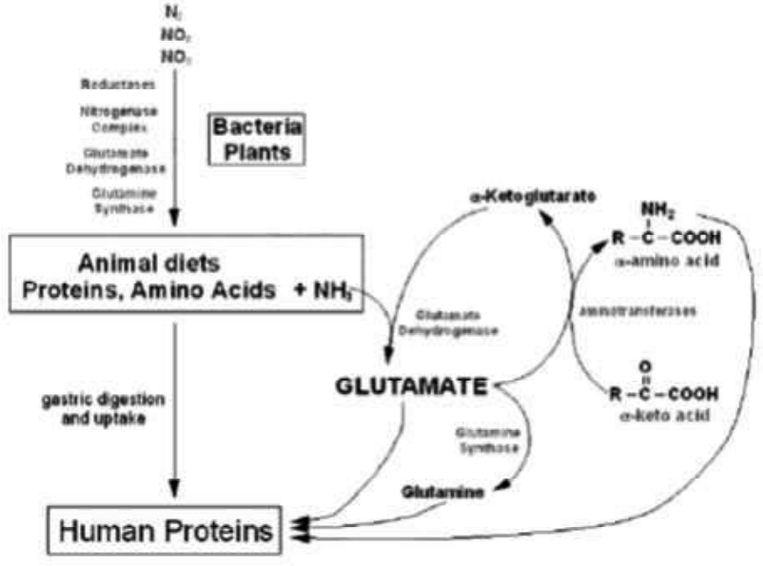


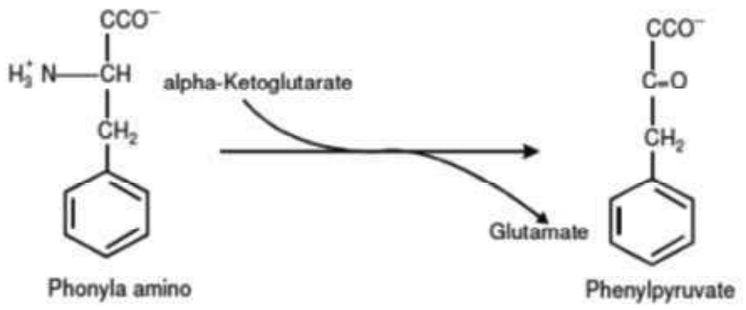
Fig. 4.30 *Reductive Amination*

Besides glutamate, some other amino acids such as aspartic acid and alanine are reportedly formed due to reductive deamination. An enzyme called, aspartase has also been reported to be present in some plants and this helps to form aspartic acid by reductive amination of fumaric acid. However, the most important reductive amination is that of α-keto glutarate only and glutamic acid is the main amino acid formed as a consequence. Glutamate dehydrogenase has been detected in several organisms including mammals, plants and microbes and contains zinc as a prosthetic group.

4.6.2 Catabolic Phase: Transamination Reaction

Besides reductive amination, another major route for the introduction of an amino group into keto acids is transamination. This occurs by means of a reaction where the -amino group from one amino acid changes place with a carbonyl group on another molecule (keto acid). Glutamate is the most important donor of an amino group and, since it is formed in abundance during reductive amination, it is never in short supply. It has been reported that more than eighteen amino acids are formed through transamination reactions with glutamate. The following equation shows transamination of phenylalanine to phenylpyruvate:

NOTES



Transamination reactions are catalysed by a group of enzymes collectively known as aminotransferases or transaminases. There are several transaminases known, and particularly important are aspartate transminase (catalysing the formation of aspartic acid), alanine transminase (catalysing the formation of alanine, leucine transaminase (catalysing the formation of leucine) and tyrosine transaminase (catalysing the formation of tyrosine). In almost all the reactions, glutamate is primarily responsible, for donating an amino group to form other amino acids.

Amino acid nitrogen is removed from the body mainly by the reactions called transamination. These reactions convert nitrogen from all free amino acids into a small number of compounds, which may undergo the following processes:

- Oxidative deamination to produce ammonia
- Conversion of the amine groups into urea by urea cycle

In the transamination reaction, an α-amino group from a donor α-amino acid moves to the keto carbon of an acceptor α-keto acid. These reactions are reversible. A group of intracellular enzymes known as aminotransferases catalyse these reactions. These enzymes usually use covalently bonded pyridoxal phosphate as a cofactor. However, pyruvate can act as a cofactor for some aminotransferases.

Aminotransferases do not catalyse the transamination of threonine and lysine. Glutamate and AKG are the most common compounds involved as a donor/acceptor pair in transamination reactions. They can act as donor/acceptor pairs in the reactions with several different aminotransferases. The types of serum aminotransferases are as follows:

- Aspartate Aminotransferase, AST (also called Serum Glutamate-Oxaloacetate Aminotransferase, SGOT)
- Alanine Transaminase, ALT (also called Serum Glutamate-Pyruvate Aminotransferase (SGPT)

The serum aminotransferases have been utilized as clinical markers of tissue damage with increasing serum levels indicating an enhanced level of damage of the use of enzyme levels in diagnosis. ALT plays a significant role in the delivery of skeletal muscle carbon and nitrogen (in the form of alanine) to liver in the following steps:

Step 1: In the skeletal muscle, pyruvate is transaminated to alanine. This provides an additional route of nitrogen transport from muscle to liver.

Step 2: In the liver, alanine transaminase transfers ammonia to AKG and regenerates pyruvate.

NOTES

Step 3: Pyruvate is diverted into the process of gluconeogenesis. This step is referred to as the glucose-alanine cycle (Figure 4.31).

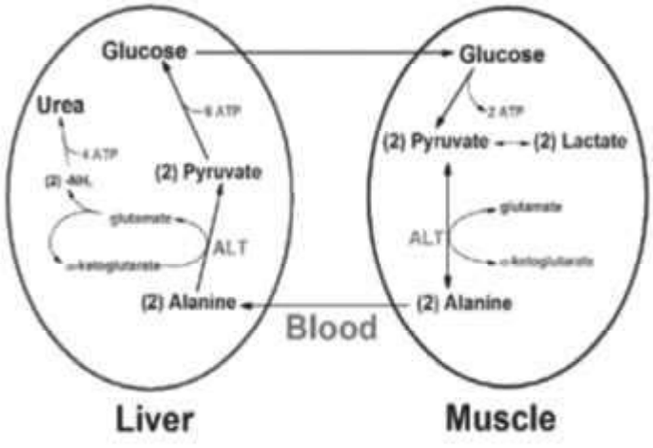


Fig. 4.31 *Glucose-Alanine Cycle*

Glutamate acts as a prominent intermediate compound in the elimination of nitrogen and anabolic pathways because AKG acts in several transaminations. Glutamate is produced in the process of nitrogen elimination. It is deaminated in the following reactions:

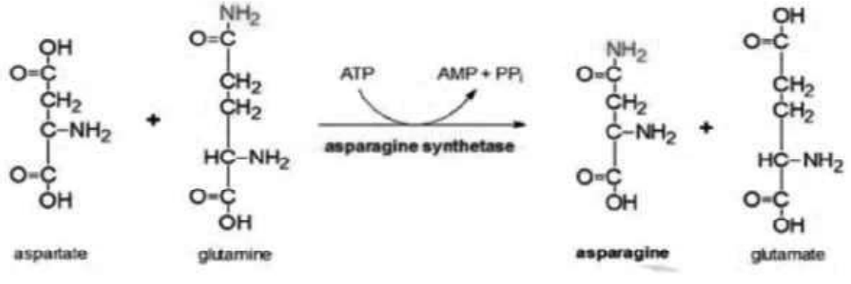
- Oxidative deamination by liver glutamate dehydrogenase forming ammonia
- Conversion to glutamine by glutamine synthase and transport to kidney tubule cells

In the kidneys, glutamine may undergo deamination by the following enzymes:

- Glutaminase
- Kidney glutamate dehydrogenase

Ammonia generated in the preceding reactions gets excreted as NH_4^+ in urine. NH_4^+ help maintain urine pH in the normal range of pH4 to pH8. Normal serum ammonium level ranges from 20 to 40 μmol/L. Increased level of ammonia in blood (approximately up to 400 μmol/L) causes neurotoxicity and alkalosis.

The amination of aspartic acid to generate asparagines is a therapeutically useful amino acid-related reaction. The enzyme asparagine synthase catalyses the transamination reaction using Adenosine Triphosphate (ATP) as per the following equation:



NOTES

A minor enzymatic pathway exists in the peroxisomes of mammalian tissues, particularly, liver for the removal of amino groups from amino acids. L-amino acid oxidase is FMN-linked and has a broad specificity for the L-amino acids. Several substances, including oxygen, behave as electron acceptors from the flavoproteins. Oxygen, as an electron acceptor, forms hydrogen peroxide, which is quickly dissociated by the catalysts found in liver and other tissues. Missing or abnormal biogenesis of peroxisomes or L-amino acid oxidase results in generalized hyperaminoacidemia and hyperaminoaciduria, which usually cause neurotoxicity and early death.

Pyridoxamine Phosphate

It has further been established that transamination reactions utilize pyridoxal phosphate or pyridoxamine phosphate as a coenzyme. Apparently, pyridoxal phosphate is a common enzyme for all the transminases known which undergoes a common reaction mechanism. Further analysis reveals that the covalent attachment of -amino group to the carbon atom of the aldehyde group of the pyridoxal phosphate (bound to transaminase) results in a structure called aldimine, which in its tautomeric form is converted to ketimine. This ketimine formed combines with a water molecule and, consequently, results in the formation of free -keto acid and pyridoxamine phosphate. It seems that pyridoxal phosphate (which is non-covalently bound to a specific transaminase) accepts an amino group from the amino acid and forms pyridoxamine phosphate, and keto acid is released.

Pyridoxamine phosphate so formed then transfers the amino group to another keto acid, forming an amino acid. Pyridoxal phosphate is regenerated during this transfer of the amino group. Pyridoxal phosphate is one of the most important coenzymes responsible for carrying on transamination reactions. Besides; it also catalyses other reactions as follows;

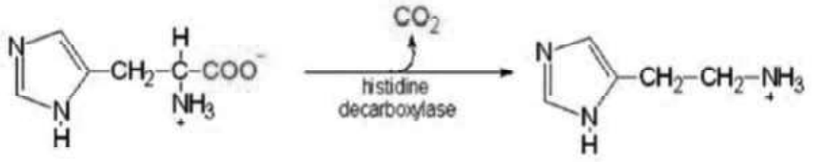
- Enzymatic decarboxylation of amino acid
- Racemization reaction (conversion of L-amino acid to D-amino acid and vice versa)
- Deamination reaction
- Serine dehydration
- Removal of sulphur from cysteine
- Aldol cleavages of amino acids

Decarboxylation of Amino Acids

Decarboxylation reactions are important in amino acid metabolism and they release primary amines, which have multiple functions in living cells. Some of the important decarboxylation reactions can be summarized as follows:

- The most important decarboxylase is one specific for L-glutamic acid in bacteria particularly *Clostridium* sp. The enzyme glutamic decarboxylase releases y-amino butyric acid by decarboxylating glutamic acid.

- Histamine is a strong vasodilator and is generally released during inflammations and allergic reactions. Histidine decarboxylase produces histamine, which is very important for gastric secretion in animal tissues, according to the following equation:



- Tyrosine decarboxylase decarboxylates 3, 4-dehydroxyphenylalanineand forms 3, 4-dihydroxyphenylethylamine (dopamine). Dopamine is an important intermediate in the formation of adrenalin.
- Tryptophan is also decarboxylated (tryptophan decarboxylase) to form tryptamine which is an important precursor for the synthesis of indoleacetic acid (plant growth hormone). All the amino acid decarboxylases require pyridoxal phosphate as a co enzyme.

Formation of Amides

Amides are formed from the amino acids with the further addition of an amino group. The amidation reactions are catalyzed by specific enzymes. The most important amide is glutamine, which is produced by the amination of glutamate. In the synthesis of glutamine, the hydroxyl group of one of the carboxyl groups of glutamic acid is replaced by a NH_3 group. The reaction is catalysed by glutamine synthetase in the presence of Mg; and ATP. The synthesis of glutamine takes place in two steps:

- Formation of g-glutamyl phosphate
- Formation of glutamine from g-glutamyl phosphate

However, asparagine synthetase has not yet been successfully obtained in the purified form.

Oxidative Deamination

Only liver mitochondria contain glutamate dehydrogenase, which deaminate glutamate. Thus all amino acids are first transaminated to glutamate, which is then finally deaminated. The coupling of deamination is called transdeamination. Amino acids are deaminated at the rate of about 70 gm per day during the transamination reaction the amino group of all other amino acids is funneled in to glutamate. Hence the glutamate dehydrogenase reaction is the final reaction; it needs NAD^+ as coenzyme. It is also an allosteric enzyme. Metabolic Pathway of deamination: L-Amino acids deoxidase can act on all amino acids except hydroxyl amino acids and decarboxylic amino acids. It uses Flavin Mononucleotide (FMN) as coenzyme. The peroxide formed in this reaction is decomposing by catalase in the peroxisome. D-amino acid oxidase can oxidize glycine and any D – amino acids that may be formed by bacterial metabolism.

NOTES

Pathways Leading to Acetyl CoA

Acetyl coenzyme A (acetyl CoA) is an important molecule in metabolism, used in many biochemical reactions. Its main function is to convey the carbon atoms within the acetyl group. In chemical structure, acetyl CoA is the thioester between coenzyme A and thiol. Acetyl CoA is produced during the pyruvate decarboxylation, which occurs in the matrix of the mitochondria. Acetyl CoA is also an important component in the biogenic synthesis of the neurotransmitter acetylcholine. Choline, in combination with acetyl-CoA, is catalysed by the enzyme choline acetyltransferase to produce acetylcholine.

The second pathway of threonine catabolism utilizes serine hydroxymethyltransferase. As indicated above this enzyme belongs to a family of one-carbon transferases and is alternatively named glycine hydroxymethyltransferase or threonine aldolase. The products of this reaction are acetyl CoA and glycine. The principal product from valine is propionyl CoA, the glucogenic precursor of succinyl-CoA. Isoleucine catabolism terminates with production of acetyl CoA and

Propionyl CoA; thus isoleucine is both glucogenic and ketogenic. Leucine gives rise to acetyl CoA and acetoacetyl CoA, and is thus classified as strictly ketogenic. Because this transamination reaction is not reversible, lysine is an essential amino acid. The ultimate end-product of lysine catabolism is acetoacetyl-CoA.

Control of Amino Acid Synthesis

In contrast to the simple routes for the formation of, glutamic acid and alanine, essential amino acids are synthesized by long pathways and the synthesis is subject to regulation by allosteric end-product inhibition. A very good example is that of valine synthesis. It is characteristic of such pathways in which allosteric inhibition by the product of the pathways occurs at the first reaction of the sequence.

A number of other amino acid synthesizing pathways (including histidine, leucine and threonine) are similar to those for valine in being unbranched and in exhibiting end product inhibition of the first reaction. Somewhat more complicated pathways and control mechanisms are noted in the case of other amino acids. The synthesis of tryptophan, phenylalanine and tyrosine, all partly include a common pathway. Consequently, simple feedback inhibition is adequately operational. Tryptophan serves as an allosteric inhibitor of the first reaction that leads to its own formation. Tyrosine and phenylalanine produce a similar effect and inhibit the first reaction in the pathway. This mechanism has been extensively investigated in *Bacillus subtilis* and is called sequential feedback control.

Another interesting control system is illustrated by threonine and lysine, both of which are synthesized from (3-aspartyl phosphate. The synthesis of these two amino acids has been very well studied in a bacterium *Rhodopseudomonas capsulatus*. The enzyme (3-1 asparatate phosphokinase is not inhibited by threonine or lysine when supplied individually. However, there is a strong inhibition of this enzyme when threonine and lysine are supplied together. Such a control mechanism is called concerted feedback inhibition.

In some cases, every final product formed during the reaction is able to inhibit the initial enzyme. If there are two different final products formed, then the inhibition is a cumulative effect of both the products. The best known example is that of the enzyme glutamine synthatase. The enzyme has been characterized in several bacterial and cyanobacterial systems. It has been isolated in the pure state from *E. coli* and is composed of 12 subunits each with a molecular weight of about 53,000 daltons. Electron micrographs indicate that the 12 subunits are arranged in two hexagonal layers and this aggregate can be dissociated into subunits by treatment with guanidine, urea or a detergent. Divalent cations, such as Mg_{2+}, appear to be involved in the binding of the subunits together. Glutamine is an important precursor of a number of compounds including tryptophan, adenylic acid, cytidylic acid, glucosamine, alanine, glycine, and histidine and carbamoyl phosphate. In the cells of *E. coli* as well as in those of a number of organisms studied, all these compounds exert allosteric inhibition of glutamine synthetase.

Analysis of the properties of the enzyme under different conditions reveals that there are more than eight separate sites to accommodate the several inhibitory molecules. Glutamine synthetase also exhibits another kind of control mechanism unlike repression or feedback inhibition. The mechanism is adenylylation or covalent modification (reversible) and there is covalent attachment of an AMP moiety to the hydroxyl group each subunit of glutamine synthetase. This is mediated by an enzyme called adenylyl transferase, and AMP moiety can be removed by a deadenylating enzyme. The glutamine synthetase is adenylylated in the presence of NH_4^+ ions in the medium. In other words, NH_4^+ ions repress the synthesis of glutamine by adenylating glutamine synthetase. This type of regulation is not of universal occurrence. Recently, it has been shown that the adenylylation control mechanism as observed in *E. coli* is in sharp contrast to the situation in gram positive bacteria, such as *B. subtilis*.

Formation of Nitrogenous Excretion Products

Nitrogen, nitrites and nitrates are acted upon by bacteria (nitrogen fixation) and plants and we assimilate these compounds as protein in our diets. Ammonia incorporation in animals occurs through the actions of glutamate dehydrogenase and glutamine synthase. Glutamate plays the central role in mammalian nitrogen flow, serving as both a nitrogen donor and nitrogen acceptor. Humans are totally dependent on other organisms for converting atmospheric nitrogen into forms available to the body. Nitrogen fixation is carried out by bacterial nitrogenases forming reduced nitrogen, NH_4^+ which can then be used by all organisms to form amino acids.

Reduced nitrogen enters the human body as dietary free amino acids, protein, and the ammonia produced by intestinal tract bacteria. A pair of principal enzymes, glutamate dehydrogenase and glutamine synthatase, are found in all organisms and effect the conversion of ammonia into the amino acids glutamate and glutamine, respectively. Amino and amide groups from these two substances are freely transferred to other carbon skeletons by transamination and transamidation reactions.

The glucose-alanine cycle is used primarily as a mechanism for skeletal muscle to eliminate nitrogen while replenishing its energy supply. Glucose oxidation produces pyruvate which can undergo transamination to alanine. This reaction is catalysed by alanine transaminase, ALT. Additionally, during periods of fasting, skeletal muscle protein is degraded for the energy value of the amino acid carbons and alanine is a major amino acid in protein. The alanine then enters the blood stream and is transported to the liver. Within the liver alanine is converted back to pyruvate which is then a source of carbon atoms for gluconeogenesis. The newly formed glucose can then enter the blood for delivery back to the muscle. The amino group transported from the muscle to the liver in the form of alanine is converted to urea in the urea cycle and excreted.

4.6.3 Urea Cycle and Ammonia Excretion

The formation of urea in the body is a complex process and involves different processes.

Assimilation of Ammonia (NH_3)

A major reaction in the assimilation of NH_3 is catalysed by glutamine synthetase, an enzyme that is ubiquitous in nature. The significance of glutamine in nitrogen metabolism results from the fact that the amide nitrogen atom serves as the precursor of nitrogenous compounds i.e. glutamic acid, asparagine, tryptophan, histidine glucosamine 6 phosphate, NAD^+, p-aminobenzoic acid and carbamoyl phosphate. The enzyme glutamate synthetase is widely distributed in bacterial species. The reactions catalysed by glutamate synthase can be coupled with glutamine synthetase and transamination to accomplish the synthesis of amino acids ($RCHNH_2COOH$) from keto acids (RCOCOOH) by a process that is unidirectional and driven by the hydrolysis of ATP. These reactions are as follows:

$$\text{L-Glutamic acid} + \text{ATP} + NH_3 \xrightarrow{\text{Glutamine synthatase}} \text{L-Glutamine} + \text{ADP} + H_3PO_4$$

$$\text{-}\alpha\text{-Ketoglutarate} + \text{L-glutamine} + \text{NADPH} + H^+ \rightarrow 2\,\text{L-Glutamic acid} + \text{NADP} + H_2O$$

$$\text{L-Glutamic acid} + \text{RCOCOOH} \rightarrow RCHNH_2COOH + \text{-}\alpha\,\text{Ketoglutarate (transamination)}$$

$$\text{RCOCOOH} + \text{ATP} + NH_3 + \text{NADPH} + H^+ \longrightarrow RCHNH_2COOH + \text{ADP} + H_3PO_4 + \text{NADP} + H_2O$$

Besides these reactions, the intestinal bacteria produce ammonia from dietary protein as well as from the urea present in fluids secreted into the gastrointestinal tract. This ammonia is absorbed from the intestine into the portal venous blood. Under normal conditions the liver promptly removes the ammonia from the portal blood. Minute quantities of ammonia are toxic to the central nervous system.

Ammonia is excreted as ammonium salts during metabolic acidosis but the majority is excreted as urea. Ammonia is present only in traces in blood because it is rapidly removed from the circulation by the liver and converted to glutamine or urea. In brain, the major mechanism for removal of ammonia is glutamine formation and in the liver, the most important pathway is urea formation.

Sir Hans Kreb and K. Henseleit were among the first to study the formation urea in animal tissues. In the initial step, carbamoyl phosphate reacts with ornithine to form citruline. In the next step of the cycle, argininosuccinic acid is formed by the combination of citruline and aspartic acid in presence of argininosuccinic acid synthetase and ATP. Argininosuccinic acid is cleaved to arginine and fumaric acid by arginosuccinase, which is present in mammalian liver and kidney. The fumarate formed is converted to oxaloacetate via fumarase and malate dehydrogenase reactions and then transaminated to regenerate asparate.

Arginase catalyses the irreversible hydrolysis of L-arginine to ornithine. Urease is the enzyme which converts the unidirectional sequence for biosynthesis of arginine into a cyclic process for making urea. Thus, all the above reactions accomplish the formation of arginine, a widely occurring amino acid, from ornithine, NH_3 and CO_2. The enzymes catalysing these reactions presumably occur in a wide number of tissues in animals, plants and microorganisms. Liver is the major site of urea formation in mammals although some urea synthesis can occur in brain and kidney.

The cycle accounts for the formation of urea from NH_3, CO_2 and the amino group of aspartic acid. The requirement for the oxidizable substrates reported by Kreb is explained by the participation of ATP in the formation of carbamoyl phosphate and argininosuccinic acid. By the eventual conversion of fumaric acid back to aspartic acid, another mole of amino nitrogen can be brought to the point of reaction in the cycle (Figure 4.32).

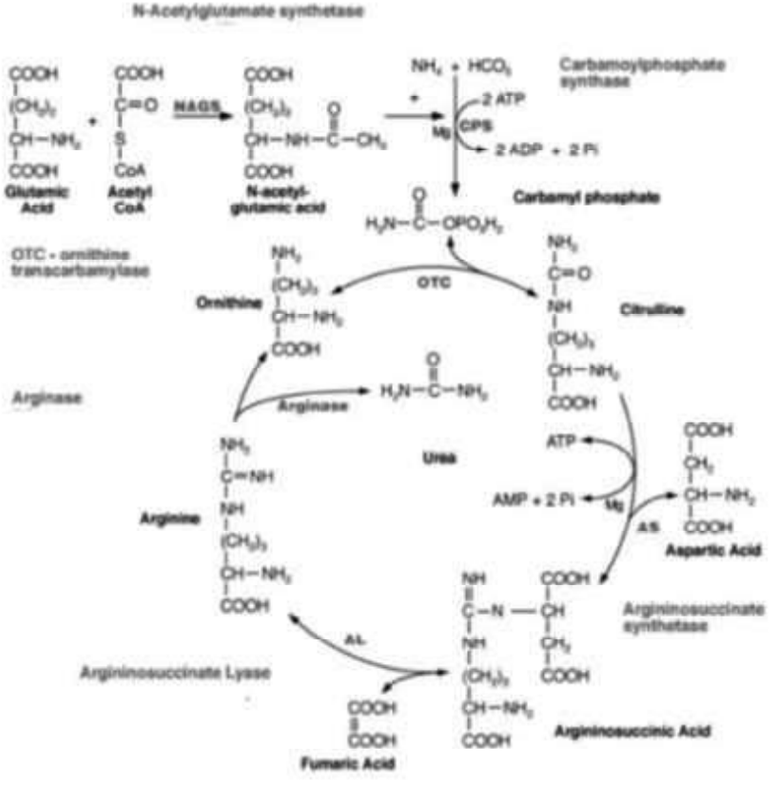


Fig. 4.32 *Kreb-Henseleit Cycle for Urea Formation*

NOTES

The enzyme ornithine transcarbamylase requires no factors and exhibits extreme substrate specificity.

Biomedical Significance

Urea constitutes about half of the total urinary solid and is the principal end product of protein metabolism. In human beings, it represents about 80-90% of the total urinary nitrogen about 25-30g of urea are excreted is proportional to the total protein metabolism including food protein and tissue protein undergoing catabolism. Urea formation is partly a cyclical process. An active man consuming about 300g of carbohydrates, 100 gm of fat, 100 g of protein daily excrete about 16.5 g of nitrogen. 95 percent is eliminated by the kidneys and 5 percent in the stool.

Urea Cycle Disorders (UCDs)

In general, the treatment of UCDs has as common elements the reduction of protein in the diet, removal of excess ammonia and replacement of intermediates missing from the urea cycle. Dietary supplementation with arginine or citruline can increase the rate of urea production in certain UCDs. Normal concentration of urea in the blood is 20 to 40 mg/100 ml. The daily output of urea through urine is 20 to 30 grams. A less quantity is excreted in the sweat. The quantity of urea excreted is proportional to the total protein metabolismdiseases. In severe acidosis, the output of urea is decreased. In nephritis, when the ability of the kidneys to excrete urea is severely impaired, the concentration of urea in the blood is increased (uremia).

Table 4.3 *UCDs: Enzyme Deficiency and Symptoms*

UCD	Enzyme Deficiency	Symptoms/Comments
Type I hyperammonemia, CPSD	Carbamoylphosphate synthetase I	With 24h–72h after birth infant becomes lethargic, needs stimulation to feed, vomiting, increasing lethargy, hypothermia and hyperventilation; without measurement of serum ammonia levels and appropriate intervention infant will die: treatment with arginine which activates N-acetylglutamate synthetase
N-acetyl glutamate synthatase deficiency	N-acetylglutamate synthetase	Severe hyperammonemia, mild hyperammonemia associated with deep coma, acidosis, recurrent diarrhea, ataxia, hypoglycemia, hyperornithinemia: treatment includes administration of carbamoyl glutamate to activate CPS I
Type II hyperamonia,OTCD	Ornithine transcarbamoylase	Most commonly occurring UCD, only X-linked UCD, ammonia and amino acids elevated in serum, increased serum orotic acid due to mitochondrial carbamoylphosphate entering cytosol and being incorporated into pyrimidine nucleotides which leads to excess production and consequently excess catabolic products: treat with high carbohydrate, low protein diet, ammonia detoxification with sodium phenylacetate or sodium benzoate

Classic citrulenemia, ASD	Argininosuccinate synthetase	Episodic hyperammonemia, vomiting, lethargy, ataxia, siezures, eventual coma: treat with arginine administration to enhance citrulline excretion, also with sodium benzoate for ammonia detoxification
Argininosuccinic aciduria, ALD	Argininosuccinate lyase (argininosuccinase)	Episodic symptoms similar to classic citrullinemia, elevated plasma and cerebral spinal fluid argininosuccinate: treat with arginine and sodium benzoate
Hyperargininemia, AD	Arginase	Rare UCD, progressive spastic quadriplegia and mental retardation, ammonia and arginine high in cerebral spinal fluid and serum, arginine, lysine and ornithine high in urine: treatment includes diet of essential amino acids excluding arginine, low protein diet

4.6.4 Fate of Carbon Skeleton of Amino Acids

Carbon skeletons are used for energy. The seven amino acids that are degraded entirely or in part to acetoacetyl-CoA and/or acetyl-CoA—phenylalanine, tyrosine, isoleucine, leucine, tryptophan, threonine, and lysine—can yield ketone bodies in the liver, where acetoacetyl-CoA is converted to acetoacetate and then to acetone and 3-hydroxybutyrate (see Figure 4.33). These are the ketogenic amino acids. Their ability to form ketone bodies is particularly evident in uncontrolled diabetes mellitus, in which the liver produces large amounts of ketone bodies from both fatty acids and the ketogenic amino acids. The amino acids that are degraded to pyruvate, á ketoglutarate, succinyl-CoA, fumarate, and/or oxaloacetate can be converted to glucose and glycogen by pathways described in previous unit. They are the glucogenic amino acids. The division between ketogenic and glucogenic amino acids is not sharp; five amino acids—tryptophan, phenylalanine, tyrosine, threonine, and isoleucine—are both ketogenic and glucogenic.

Catabolism of amino acids is particularly critical to the survival of animals with high-protein diets or during starvation. Leucine is an exclusively ketogenic amino acid that is very common in proteins. Its degradation makes a substantial contribution to ketosis under starvation conditions.

NOTES

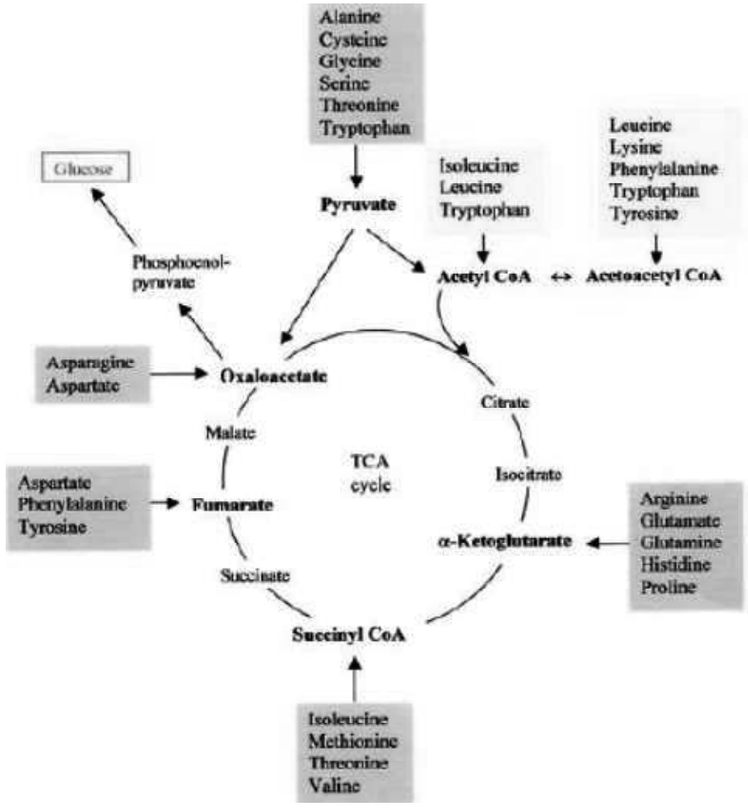


Fig. 4.33 *Amino Acid Catabolism*

Check Your Progress

14. What is alpha carbon?
15. What does SGPT stand for?
16. How is ammonia removed from the liver?
17. What is the treatment of UCD?

4.7 SEQUENCE DETERMINATION

In this section, we will discuss the different sequence determination methods.

4.7.1 Chemical

Chemical method of polypeptide sequence determination is based on using chemicals for hydrolyzing the peptide bond into its constituent amino acids. Various chemicals are used for this process; they are outlined below:

- **Edman Degradation Method:** This method of protein sequence estimation from their N- terminus was developed by Pehr Edman. This chemical method is based on purifying protein by serially removing one residue at a time from the amino end of a peptide. It is done to save the protein chain by the damage that occur on it due to hydrolyzing conditions.

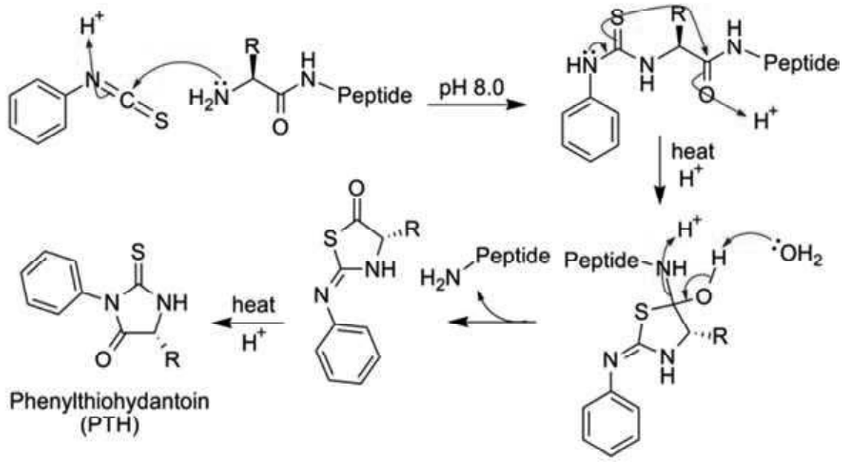


Fig. 4.34 *Edman Degradation*

Pehr Edman found a new way of labelling and cleaving the peptide with removal of only one residue at a time, which do not damage the overall sequencing. Phenyl isothiocyanate (Edman's reagent) when added the proteins, creates a Phenyl Thiocarbamoyl derivative (PTC- peptide) with the N-terminal derivative. The N-terminal is then cleaved under less harsh acidic conditions, creating a cyclic compound of phenylthiohydantoin PTH-amino acid. This does not damage the protein and leaves two constituents of the peptide. This method can be repeated for the rest of the residues, separating one residue at a time.

- **N-Terminal Analysis:** Reagent 1-fluro-2, 4-dinitrobenzene (FDNB) and Dansyl chloride are used for this method of detection of N-terminal residue of the protein chain. FDNB reacts with the free amino groups of the N-terminal amino acid residues of the peptide in the alkaline solution (pH 9.5) to form a distinctive yellow Dinitrophenyl (DNP) derivative. This derivative is released from the peptide chain by either acid or enzyme hydrolysis and are identified. This reaction was used by Sanger for the first time for the elucidation of the primary structure of insulin polypeptide protein. So, this reagent is also termed as Sanger's reagent.

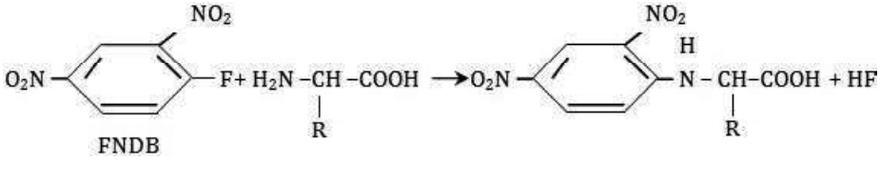


Fig. 4.35 *N-terminal Analysis using Sanger's Reagent*

- **Reduction and Alkylation:** In this method, firstly the protein is denatured to expose the higher order structure of protein and expose the many internal disulphide bonds. Denaturation of the protein chain is done using high temperature or chemical agent. After the denaturation, the exposed structures can be easily reduced using a reagent such as 1, 4-dithiothreitol (DTT),

NOTES

mercaptoethanol, or tris(2-carboxethyl) phosphine (TCEP). Usually, denaturation and reduction are conducted together (i.e., heating the sample in the presence of DTT) to avoids the problem of renaturation during the process. After this, alkylation of the cysteine is needed to avoid reformation of the disulphide bonds in the protein (renaturation). For this, protein chain is incubated with an alkylating agent such as 2-Iodoacetamide (IAA).

4.7.2 Enzymatic Method

The use of enzymatic method for sequence determination of protein is usually a very preferred method as other sequencing techniques in use cannot predict the sequence of polypeptide chain with more than 50 amino acid residues. So, in the enzymatic method, Proteolytic enzyme is used which cleaves the protein into the peptide fragments and at the specific position in the protein sequence to produce a 'fingerprint' of fragments, which are then chromatographically separated and identified by MS detection.

Chemical methods are also used for the protein sequence determination but the enzymatic method is the more preferred one. The reason for the same is that the fragments which are produced by enzymatic digestion are very predictable because the cleavage occurs at specific sites. The specificities are determined by the side chains of amino acid present at the either side of the peptide bond which are hydrolyzed. The proteolytic enzymes can be endopeptidases or exopeptidases. Endopeptidases cleave the internal peptide bonds, whereas, exopeptidases break the terminal points of the polypeptide chain.

Trypsin, chymotrypsin, elastase, thermolysin and pepsin are the examples of endopeptidases. Carboxypeptidases and aminopeptidases are the examples of exopeptidases. Trypsin is mostly used for peptide mapping due to its well-defined specificity. Trypsin hydrolyses the peptide bonds at the carbonyl group of either arginine (Arg) or a lysine (Lys) amino acid.

Table 4.4 *List of Enzymes*

Enzyme	**Site of Cleavage**
Trypsin	Lys, Arg (C)
Chymotrypsin	Phe, Trp, Tyr (C)
Asp-N-protease	Asp, Glu (C)
Pepsin	Leu, Phe, Trp, Tyr (N)
Elastase	Ala, Gly, Ser (C)
Endoproteinase Lys C	Lys (C)

The choice of enzyme for cleavage depends upon the sequence needed or for the purpose it is needed. Several enzymes are there which are comparable of cleaving the peptide bonds at the specific positions. All the enzymes are outlined in the above table.

NOTES

4.7.3 Mass Spectrometry (MS)

Mass Spectrometry (MS) is a widely used technique for determining the molecular mass and structure of a protein. The efficiency of the method for determination of the protein structure is due to the modern computational methods through comparisons of ion fragment data with computer databases of known protein structures. Initially, the protein sample is collected and broken into smaller peptides with an enzyme, for e.g., trypsin, which cleaves on carboxyl side of positively charge Lys and Arg side chains. As we know, the average size of proteins in the human proteome is about 50,000, and the average molecular mass of an amino acid in a protein is around 110 (18 subtracted since water is released on amide bond formation), so, the average number of amino acids in the protein would be about 454. If 10% of the amino acids are Arg and Lys, then on average there would be approximately 50 Lys and Arg, and hence 50 tryptic peptides of average molecular mass of 1000.

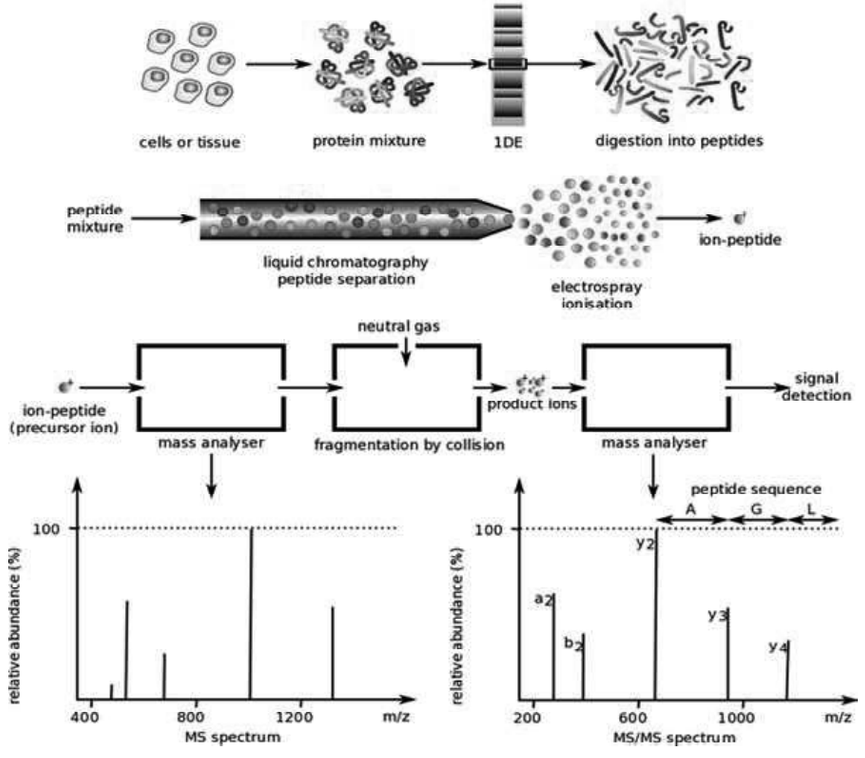


Fig 4.36 *Mass Spectrometry Protocol*

The fragments are introduced in the MS where a peptide fragment fingerprint analysis can be performed. The MWs of the fragments can be identified and compared to known peptide digestion fragments from known proteins to identify the analyte protein. After the fragmentation, the fragments are ionized in an ion source. Then, into a mass analyser, the charge particles are accelerated by electric field and are subjected to an external magnetic field. The interacting magnetic fields, one arising from particle movement and other externally applied, causes the particles to deflect. The deflection produced in the particles is proportional to the mass to charge ratio, m/z. Then, ions are encountered by a detector, usually a photomultiplier.

NOTES

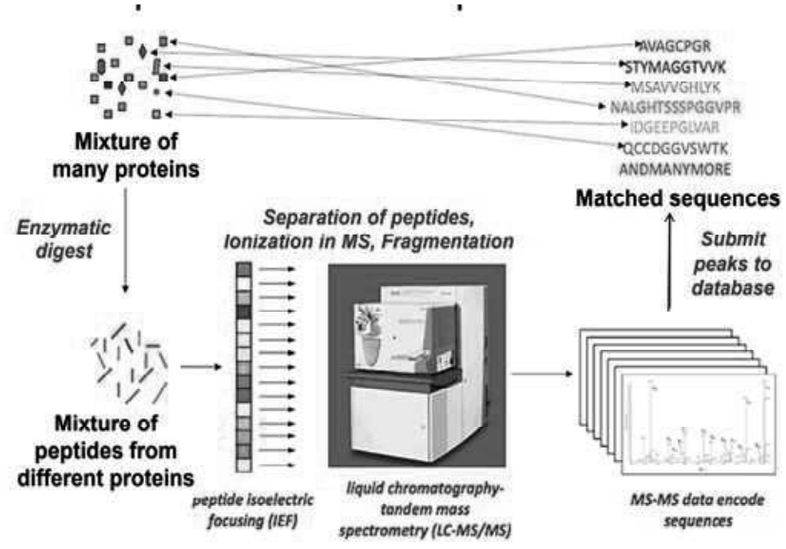


Fig 4.37 *Procedure of MS*

Various applications of using MS technique are:

- Protein identification
- *De novo* (peptide) sequencing
- Antigen presentation
- Protein quantitation
- Protein structure determination
- Proteogenomic

The components used in mass spectroscopy are outlined below.

Ion Source

Ions are ionized by Atmospheric Pressure Chemical Ionization (APCI), Chemical Ionization (CI), or Electron Impact (EI). The most common methods for protein/ peptide analyzes are Electrospray Ionization (ESI) and Matrix Assisted Laser Desorption Ionization (MALDI).

Electrospray Ionization (ESI)

Analyte id dissolved in a volatile solvent which may be methanol or acetonitrile and injected into the ion source through a fine stainless steel capillary at a slower rate. Positive charge is kept on the capillary by giving a high voltage of 3-4 kV with respect to the other oppositely charged electrode. The flowing liquid also carries the same polarity as of the positively charged capillary. The application of high field leads to the appearance of the sample like a charged aerosol spray of charged micro drops reducing the electrostatic repulsions in the liquid.

The principle of the method is based on using electrical energy for the production of the aerosol. As the volatile solvent evaporates, micro drops become smaller in size and increases the positive charge density on the drops. Ultimately electrostatic repulsions cause the drops to burst in a series of steps producing analyte

devoid of solvent. This method of ionization produces analytes that are not cleaved but ready for introduction into the mass analyzer. For gases, samples are introduced into the ion source by simple diffusion, volatile liquids from a reservoir, by injection of a liquid sample comprising the analyte by spraying a fine mist, or very large proteins by desorbing a protein from a matrix using a laser. Complex mixtures are then analysed by coupling HPLC with mass spectrometry in a LCMS.

Matrix Assisted Laser Desorption Ionization (MALDI)

This technique is used for larger biomolecules like proteins and polysaccharides. In it, the analyte is mixed with an absorbing matrix material. Then, with laser excitation, matrix is excited leading to energy transfer. Due to this, ionization and "launching" of the matrix occurs and analyte in the form of ion is produced from the solid mixture.

Mass Analyzer

The mass analyzer or quadrupole ion trap (used in ESI) is used for trapping the complex mixture of ions. Quadrupole ion trap analyzer can be of two types: linear and 3D quadrupoles. In it, the opposite electric or magnetic fields are contained on the opposite ends of the square or a cube. There sum of charges are not zero and this arrangement uses a combination of fixed and alternating electric fields. The trap comprises of He at 1 mTorr.

In a 3D trap, the ring electrode contain an oscillating RF voltage which keeps the ions trapped. The end caps also have an AC voltage. Determined by the frequency of the RF voltage and the m/z ratio, ions oscillate in the trap with a 'secular' frequency. By increasing the amplitude of the RF field across the ring electron, motion of ion get destabilized and leads to ion ejection into the detector. When the secular frequency of ion motion matches the applied AC voltage to the end cap electrodes, resonance occurs and the amplitude of motion of the ions increases, also allowing leakage out of the ion trap into the detector.

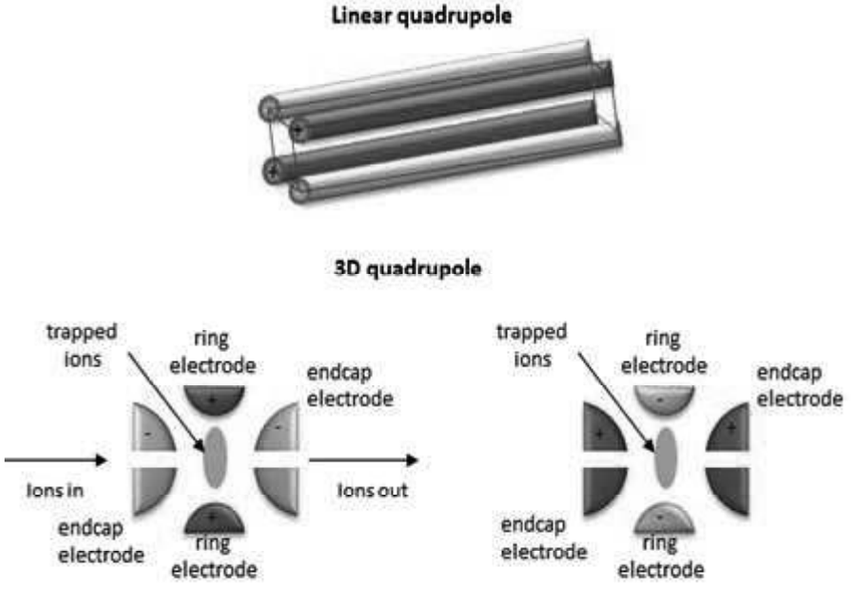


Fig 4.38 *Linear and 3D Quadrupole*

Time of Flight (TOF) tube (used in MALDI), which are long tubes, are used and the time required for ion detection is determined. The small molecular mass ions take the shortest time to reach the detector.

NOTES

Check Your Progress

18. Give some examples of endopeptidases.
19. Why is the mass spectrometry method of sequence determination efficient?
20. Name the types of quadrupole ion trap analyzer.
21. Mention any three applications of the MS technique.

4.8 RACEMIZATION/DETECTION

Racemization is a process of conversion of an optically active compound into a racemic (optically inactive) form by the application of heat or by chemical reaction. Due to this, half of the optically active substance becomes its mirror image (enantiomer) referred as racemic mixtures (i.e., contain equal amount of (+) and (“) forms). If the obtained racemization mixture contains D and L enantiomers in equal quantities, the resulting sample is described as a racemic mixture or a racemate. As different enantiomers may have different pharmaceutical effects, racemization plays an important role in pharmacology. Racemization can be completed by simply mixing equal quantities of two pure enantiomers or chemical interconversion. For example, when (*R*)-3-phenyl-2-butanone is dissolved in aqueous ethanol that contains NaOH or HCl, a racemate is formed. The racemization occurs by intermediate enol form in which the former stereocenter becomes planar and hence achiral. An incoming group can approach from either side of the plane, so there is an equal probability that protonation back to the chiral ketone will produce either an *R* or an *S* form, resulting in a racemate.

4.9 CHEMISTRY OF OXYTOCIN

The word ‘oxytocin’ is derived from a Greek word `κυτόκος” (*ōkutókos*), based on @ξύς (oxús), meaning ‘sharp’ or ‘swift’, and τόκος (tókos), meaning ‘quick birth’ after its uterine-contracting properties were discovered by Dale in 1906. Oxytocin is a pleiotropic polypeptide produced by the hypothalamus and secreted by the posterior pituitary gland of mammals. It has several hormonal functions in general health, adaptation, development, reproduction, and social behavior. The hormone is also known for its major role in lactation and parturition. The non-apeptide structure of oxytocin is composed of nine amino acids which are, Cys-Tyr-Ile-Gln-Asn-Cys-Pro-Leu-Gly-NH_2 or CYIQNCPLG-NH_2 having a sulfur bridge between the two cysteines. The c-terminus convert to primary amide. The structure of Oxt show resemblance with another nonapeptide, vasopressin (Avp) with the difference of two amino acids.

Oxytocin has a molecular mass of 1007 Da, and one international unit (IU) of oxytocin is the equivalent of 1.68 μg of pure peptide. Oxidized octapeptide oxytocin is the biologically active form of oxytocin usually measured by Radioimmunoassay and/or HPLC techniques. It is disulphide, but oxytocin can also exist as a dithiol nonapeptide called oxytoceine is a reduced straight-chain (non-cyclic). It has been hypothesized that oxytoceine act like a free radical scavenger, donating an electron to a free radical, which allows it to re-oxidize into oxytocin via the dehydroascorbate / ascorbate redox couple.

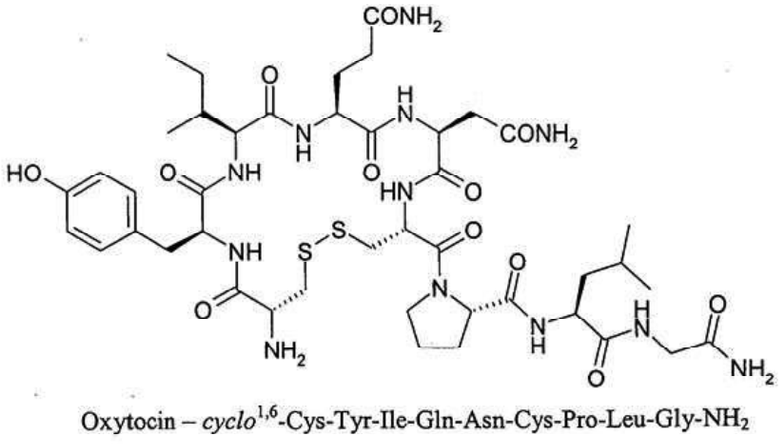


Fig 4.39 *Molecular Structure of Oxytocin*

Recent advances in analytical instrumental techniques highlighted the importance of liquid chromatography (LC) coupled with mass spectrometry (MS) for measuring oxytocin levels in various samples derived from biological sources.

Initially the oxytocin peptide is synthesized as inactive precursor protein from the *OXT* gene. Carrier protein neurophysin is also attached with the precursor protein. The inactive precursor protein is gradually hydrolyzed into smaller fragments (one of which is neurophysin I) by the series of enzymes. The last hydrolysis that releases the active oxytocin nonapeptide is catalyzed by Peptidyl Glycine Alpha-Amidating Monooxygenase (PAM). The activity of this hydrolyzing enzyme depends on vitamin C (ascorbate), which is a necessary vitamin co-factor.

The role of oxytocin on the behaviours and physiology depends very much on the steroid hormones and gender and that why their expression is typically higher in females. So, the numbers of Oxt-immunostained cells and the amounts of Oxt in females exceed the numbers and amounts that is found in males with greater numbers of oxytocin-immunostained axons too. The actions of Oxytocin are mediated by specific oxytocin receptors which is a G-protein-coupled receptor, OT-R, which requires magnesium and cholesterol and is expressed in myometrial cells (The myometrium is the middle layer of the uterine wall, which consist mainly of uterine smooth muscle cells).

Many parts of the brain and spinal cord have neurons carrying oxytocin receptors like amygdala, ventromedial hypothalamus, septum, nucleus accumbens, and brainstem. Milk ejection reflex/Letdown reflex: Oxytocin act on the mammary

NOTES

glands in the lactating mothers which causes the milk to come to lactiferous ducts and then it is excreated by the nipple. Suckling at the nipple is conveyed by spinal nerves to the hypothalamus. When stimulated, the neurons create action potentials in intermittent bursts which results in the secretion of pulses of oxytocin from the neurosecretory nerve terminals of the pituitary gland.

It is released in the bloodstream as a hormone in response to sexual activity and during labour. It can also administer in synthetic or pharmaceutical from, where in both the cases, its role is to stimulate uterine contractions to make the process of childbirth fast. In its natural form, it also plays a role in lactation. When the oxytocin is released during the initial stage of childbirth for the contractions of the uterus, it gives the positive feedback which controls the release of more hormone to increase in the intensity and frequency of contractions.

Oxytocin shows its very vital role in the uterine contraction during the second and third stages of birth and cervical dilation. After the birth, during first few weeks of lactation the released oxytocin causes mild but often painful contractions. This contraction also serves to support the uterus in clotting the placental attachment point postpartum. It has shown to induce erection in the males. When the burst of oxytocin is released during ejaculation in several species, including human males; it stimulates contractions of the reproductive tract, aiding sperm release. Oxytocin levels are found to increase in plasma during sexual stimulation and orgasm in both men and women.

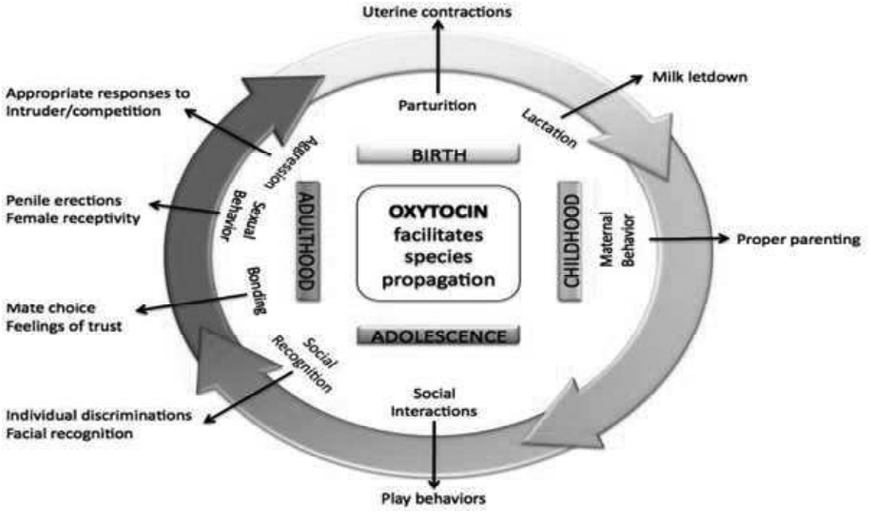


Fig 4.40 *Different Roles of Oxt in Life Affecting Behaviour and Physiology*

It is classified as antidiuretic as due to its similarity to vasopressin, it can reduce the excretion of urine slightly. In several species, it can also stimulate sodium excretion from the kidneys (natriuresis), and, in humans, high doses can result in low sodium levels (hyponatremia). It has been found to function in the embryonal development of the heart by promoting cardiomyocyte differentiation as seen in some rodents. However, there absence does not show any cardiac insufficiencies. Oxytocin has been concerned in the etiology of autism, as a mutation on Oxytocin Receptor Gene (*OXTR*) is a cause of it. Studies on Caucasian, Finnish and Chinese

Han families support the relationship of *OXTR* with autism. Autism may also be associated with an abnormal methylation of *OXTR*.

4.10 CHEMISTRY OF THYROTROPIN RELEASING HORMONE (TRH)

NOTES

Thyrotropin Releasing Hormone (TRH) is formed by hypothalamus gland which is present at the base of the brain just above the pituitary gland. This nerve cell cluster is known as the paraventricular nucleus. TRH is one of the smallest hormones found in the body. It plays vital role in the regulation of thyroid gland activity by stimulating the release of thyrotropin also known as thyroid-stimulating hormone (TSH). Thyroid-stimulating hormone controls the production of thyroid hormones in the thyroid gland. Thyroxine (T4) and triiodothyronine (T3) are the thyroid hormones which controls the body's metabolic rate, heat generation, neuromuscular function and heart rate, among other things.

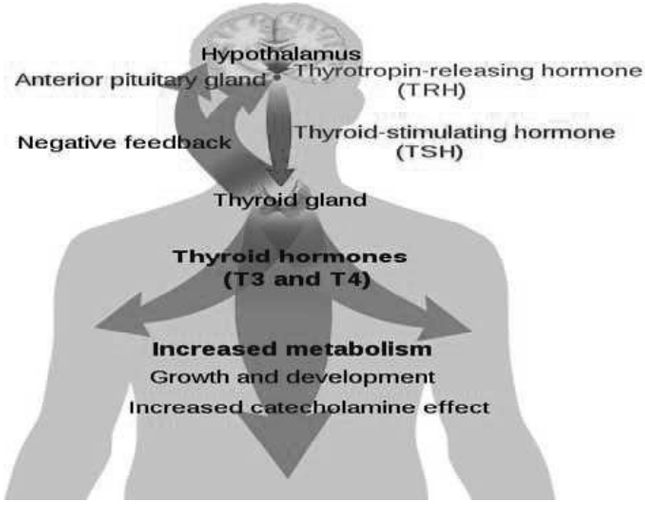


Fig 4.41 *Diagrammatic Representation of the Regulation of Production of Thyroid Hormones*

It also functions in stimulating prolactin, which is again a hormone secreted by pituitary gland and functioning as a neurotransmitter in the tissues of the nervous system. It has also shown its role in metabolism, cognition, mental health, energy balance or homeostasis, eating patterns, thermogenesis or heat production and autonomic regulation (the unconscious control of vital bodily functions) and more. As the quantity of thyroid hormones become low in the bloodstream, hypothalamus signals the pituitary gland by TRH to produce TSH for the thyroid to release more T4 or thyroid hormone. Very low TRH level leads to hypothalamic hypothyroidism, or central hypothyroidism. It is a very rare condition, which occurs due to any injury or tumour which destroys this area of the hypothalamus.

The synthesis of TRH occurs in hypothalamus. Its precursor translated form is the polypeptide of 242-amino acid with 6 copies of the sequence -Gln-His-Pro-Gly-, flanked by Lys-Arg or Arg-Arg sequences. It transforms into the mature

NOTES

tripeptide form with an amino acid sequence of pyro glutamyl-histidyl-proline amide by the series of enzymes. Following modifications are performed for the conversion which are:

- C-terminal side of the flanking Lys-Arg or Arg-Arg is cleaved by protease.
- Removal of Lys/Arg residues leaving Gly as the C-terminal residue by carboxypeptidase.
- Then, peptidyl glycine-alpha-amidating monooxygenase collectively converts Gly into an amide residue.
- Alongside with all these processing steps, the conversion of N-terminal Gln (glutamine) to pyroglutamate (a cyclic residue) occurs.

With this multistep process, a mature TRH molecule with six copies per precursor molecule for human TRH (5 for mouse TRH) is produced.

TRH travels from hypothalamus to anterior pituitary via the hypophyseal portal system, there it interacts and bounds to the TRH receptor, which leads to the stimulation of thyroid-stimulating hormone release from thyrotropes and prolactin from lactotropes. TRH has the half-life of approximately 6 minutes in the blood. The pharmaceutical form of Thyrotropin-Releasing Hormone is 'protirelin'. TRH test is usually done for testing the response of the anterior pituitary gland and for diagnosis of thyroid disorders such as secondary hypothyroidism and acromegaly by using the intravenous injection of TRH (brand name Relefact TRH).

TRH has anti-depressant and anti-suicidal properties. TRH has a central role in the regulation of metabolic and hormonal functions as been suggested in mice to be an anti-aging agent with a wide-ranging spectrum of activities. Intravenous TRH administration have minimal side effects. Like, Nausea, flushing, urinary urgency, and mild rise in blood pressure. Shaking, sweating, shivering, restlessness, and mild rise in blood pressure were observed after intrathecal administration.

Check Your Progress

22. What do you mean by oxytocin?
23. What is the role of T3 and T4 hormones?

4.11 ANSWERS TO 'CHECK YOUR PROGRESS'

1. In an amino acid, the amino and carboxylic groups are attached to the same carbon atom, which is called the a–carbon.
2. Peptide bond formation is an example of a condensation reaction, a common class of reactions in living cells.
3. The chemical properties of amino acids are mainly determined by the presence of two functional groups: carboxyl group and amino group.

4. Leucine is an amino acid that acts as persuasive stimulator of insulin. It helps promoting in bone and skin healing. It modulates the release of encephalin, which are natural pain-reducers. Its codons are CUC, CUU, CUA and CUG.
5. Tryptophan may be used as effective sleep supplement and reduces anxiety or nervousness. Tryptophan is very effective in treating some forms of depression, migraine headaches and insomnia. It also stimulates the production of growth hormone.
6. The term protein was primary suggested by a Berzelius, a Swedish chemist and Mulder, a Dutch chemist in 1838.
7. The proteins which are rich in proline amino acid include collagen presented in connective tissue, gliadin in wheat and zein in corn.
8. The four levels of protein structure were first defined by Linderstrom and Lang which includes primary structure, secondary structure, tertiary structure and quaternary structure.
9. Secondary structure of protein is formed by twisting or folding of polypeptide chain (primary structure of protein). These twisting and folding are because of interactions between atoms of polypeptide backbone which includes polypeptide chain (carboxyl group, amino group) but not R groups.
10. The word 'collagen' is derived from a Greek word which means glue as primarily it was used to define that constituent of connective tissue which gives gelatin on boiling.
11. The tertiary structure of protein refers to the arrangement of amino acids that are far apart in the chain. Each protein ultimately folds into a three dimensional shape with a distinct inside and outside. On the other hand, when a protein has two or more polypeptide subunits, their arrangement in space is referred to as quaternary structure.
12. Fibrous proteins are those proteins containing polypeptide chains arranged in the form of long strands or sheets. Fibrous proteins contribute to properties that provide strength or elasticity to the structures in which they occur.
13. The silk fibroin is composed mainly of glycine, alanine and serine amino acid residues linked together by the peptide bonds. Glycine constitutes approximately 45% of the total amino acid residues whereas alanine in addition to serine composes another 42% of the total amino acid residues.
14. All amino acids consist of an amino group and a carboxyl group bonded to a carbon atom, called alpha carbon.
15. SGPT stands for Serum Glutamate-Pyruvate Aminotransferase.
16. The major mechanism for removal of ammonia in the liver is urea formation.
17. The treatment of UCDs has as common elements the reduction of protein in the diet, removal of excess ammonia and replacement of intermediates missing from the urea cycle.

18. Trypsin, chymotrypsin, elastase, thermolysin and pepsin are the examples of endopeptidases.
19. The efficiency of the mass spectrometry method for determination of the protein structure is due to the modern computational methods through comparisons of ion fragment data with computer databases of known protein structures.
20. Quadrupole ion trap analyzer can be of two types: linear and 3D quadrupoles.
21. Various applications of using MS technique are:
    a. Protein identification
    b. De novo (peptide) sequencing
    c. Antigen presentation
22. Oxytocin is a pleiotropic polypeptide produced by the hypothalamus and secreted by the posterior pituitary gland of mammals. It has several hormonal functions in general health, adaptation, development, reproduction, and social behaviour.
23. Thyroxine (T4) and triiodothyronine (T3) are the thyroid hormones which control the body's metabolic rate, heat generation, neuromuscular function and heart rate, among other things.

4.12 SUMMARY

- Amino acids are colourless, crystalline substances. The crystal shape can differ from slender needles as found in tyrosine to thick hexagonal plates as presented in Cysteine.
- Peptide bond formation is an instance of a condensation reaction, a general class of reactions in living cells.
- In a similar manner, three amino acids can be combined by two peptide bonds for the formation of a tripeptide.
- Amino acids can be classified on the basis of their groups, namely R groups and on the basis of synthesis.
- Nonpolar, aliphatic R groups have nonpolar and hydrophobic R groups. Aromatic R groups are the aromatic side chains containing amino acids including Phenylalanine, tyrosine and tryptophan which are comparatively nonpolar (hydrophobic).
- The amino acids which get formed in our body and are not required in the course of any diet, are referred to as non-essential amino acids.
- An essential amino acid is an amino acid that cannot be synthesized by the human beings and other animals and must be supplied through diets.
- The essential functions of proteins are to build and repair of body and muscle tissues, form a part of enzymes and hormones, act as a source of energy, keep skin, hair and nails healthy and for normal bodily functions, such as respiration, require muscle contractions, which require proteins.

- Each protein, also known as polypeptide, is a polymer of L-amino acid. Proteins have four levels of structures: primary, secondary, tertiary and quaternary.
- The tasks of describing and understanding structure of large macromolecules such as proteins are approached at several levels of complexity which are therefore set in a category of theoretical chain of command.
- Secondary structure of protein is formed by twisting or folding of polypeptide chain. These twisting and folding are because of interactions between atoms of polypeptide backbone which includes polypeptide chain but not R groups.
- The linking of different amino acid residues in a polypeptide chain by covalent bonds mainly peptide bonds and disulfide bonds constitute the primary structure of protein**.**
- α-helices are regular right-hand turns of amino acids 3.6 residues long; 5.41 AÏ. Hydrogen bonding between the first backbone carbonyl oxygen atom and the fourth residue NH group stabilizes the structure, van der Walls interactions across the axis further stabilize the structure.
- β-pleated sheet secondary structure are generally fibrous, such as silk, but pleated sheet is observed as a significant part of secondary structure in other proteins.
- B-Sheets are not flat but have a pleated appearance due to the C atoms being successively above and below the plane of the sheet. The side groups are also successively above and below the plane of the sheet and them, therefore, cannot interact with each other.
- Many globular proteins contain combination of α-helix and β-pleated secondary structures. When this specific geometries of α-helix and β-pleated secondary structures are connected through loops are called super secondary structures, which are also called motifs.
- Collagen is the fibrous, structural protein found abundantly in the animals. It is the prime constituent of their various connective tissues like skin, tendons, ligaments, cartilage, bone, teeth, basement membranes, blood vessels, etc.
- Tertiary structure of protein refers to the arrangement of amino acids that are far apart in the chain. Each protein ultimately folds into a three dimensional shape with a distinct inside and outside.
- When a protein has two or more polypeptide subunits, their arrangement in space is referred to as quaternary structure**.**
- It is useful to classify proteins into two major groups for the considering these higher levels (quaternary) of structure which includes fibrous proteins and globular proteins.
- Fibrous proteins are those proteins containing polypeptide chains arranged in the form of long strands or sheets, whereas globular proteins are those proteins having polypeptide chains folded into a spherical or globular shape.

NOTES

- The crucial structural unit of a-keratin generally consists of three right-handed helical polypeptides that are stabilized by cross linking disulfide bonds in a left-handed orientation.
- Globular proteins frequently contain several types of secondary structure. Most enzymes and regulatory proteins are types of globular proteins which includes cytochrome c, lysozyme, ribonuclease and myoglobin.
- Cytochrome c is a major constituent of the respiratory chain presented in mitochondria. Cytochrome c is a heme protein and it contains a single polypeptide chain of about 100 amino acid residues with a single heme group.
- Hemoglobin is a heme-protein belonging to the myoglobin-hemoglobin family. It plays crucial role as the oxygen transporter in erythrocytes. It constitutes about 90% of the protein of red blood corpuscles or erythrocytes.
- Amino acids metabolism takes places in following two phases: anabolic and catabolic.
- The transfer of an amino (-NH2) group from an amino acid to a keto acid by enzyme catalysed reaction known as transamination.
- The removal of amino group from the amino acid as NH3 is known as deamination.
- All amino acids consist of an amino group and a carboxyl group bonded-to a carbon atom i.e., the so-called alpha carbon.
- Some of the ammonia generated is used in biosynthesis; the excess is excreted. Excess ammonia generated in extra hepatic tissues is transported to liver for conversion to the appropriate excreted form.
- Acetyl coenzyme A (acetyl CoA) is an important molecule in metabolism, used in many biochemical reactions. Its main function is to convey the carbon atoms within the acetyl group.
- Nitrogen, nitrites and nitrates are acted upon by bacteria (nitrogen fixation) and plants and we assimilate these compounds as protein in our diets.
- A major reaction in the assimilation of NH3 is catalysed by glutamine synthetase, an enzyme that is ubiquitous in nature.
- Ammonia is excreted as ammonium salts during metabolic acidosis but the majority is excreted as urea.
- Chemical method of polypeptide sequence determination is based on using chemicals for hydrolyzing the peptide bond into its constituent amino acids.
- The use of enzymatic method for sequence determination of protein is usually a very preferred method as other sequencing techniques in use cannot predict the sequence of polypeptide chain with more than 50 amino acid residues.
- Mass Spectrometry (MS) is a widely used technique for determining the molecular mass and structure of a protein.
- The mass analyzer or quadrupole ion trap (used in ESI) is used for trapping the complex mixture of ions. Quadrupole ion trap analyzer can be of two types: linear and 3D quadrupoles.

- Racemization is a process of conversion of an optically active compound into a racemic (optically inactive) form by the application of heat or by chemical reaction.
- The role of oxytocin on the behaviours and physiology depends very much on the steroid hormones and gender and that why their expression is typically higher in females.
- Thyrotropin Releasing Hormone (TRH) is formed by hypothalamus gland which is present at the base of the brain just above the pituitary gland. This nerve cell cluster is known as the paraventricular nucleus.

4.13 KEY TERMS

- **Amino Acids:** These are known as the building blocks of polypeptides and proteins. Their play important roles in metabolic pathway, gene expression, and cell signal transduction regulation. Its molecule contains two functional groups—amine and carboxyl – and a unique side chain.
- **Peptides:** These are short chains of amino acids that are linked by peptide bonds.
- **Urea:** It is a colourless crystalline nitrogenous compound which is obtained by the breakdown of the protein metabolism in mammals. It is excreted in urine.
- **Denaturation of Proteins:** It is the process of disruption and destruction of the secondary and tertiary structures of proteins.
- **Enzymes:** They are proteins that increase the rates of chemical reactions in enzymatic reactions.
- **Ammonia:** It is a colourless gas with a characteristic pungent smell, which dissolves in water to give a strongly alkaline solution.
- **mTorr:** Millitorr (mTorr) is a very small pressure unit used for high vacuum measurements and is a 1/1000x multiple of the Torr pressure unit. 1 mTorr equals 0.133322 pascals.

4.14 SELF ASSESSMENT QUESTIONS AND EXERCISES

Short-Answer Questions

1. Briefly explain the non-essential amino acids.
2. Write a short note on the formation of peptides.
3. Briefly discuss the triple helix structure of the collagen.
4. How does assimilation of ammonia (NH_3) take place?
5. What are the different types of chemical methods used in sequence determination?

NOTES

Long-Answer Questions

1. Discuss how are amino acids classified on the basis of their R groups.
2. Describe the bonding and types of secondary structure of proteins.
3. Explain the different types of globular proteins.
4. Describe the phases in amino acid metabolism.
5. Discuss the chemistry of oxytocin.

4.15 FURTHER READING

Blackstock, James C. 2014. *Guide to Biochemistry*. Oxford: Butterworth-Heinemann.

Park, K. H. 2008. *Carbohydrate-Active Enzymes: Structure, Function and Applications*. Amsterdam: Elsevier.

Bhardwaj, Uma. 2012. *Biochemistry for Nurses*. Noida: Pearson Education India.

Fromm, Herbert J. and Mark Hargrove. 2012. *Essentials of Biochemistry*. Berlin: Springer.

Fearon, William Robert. 2014. *An Introduction to Biochemistry*. Amsterdam: Elsevier.

UNIT 5 NUCLEIC ACIDS

NOTES

Structure

5.0 INTRODUCTION

Nucleic acids are the biopolymers, or small biomolecules, essential to all known forms of life. The term nucleic acid is the overall name for DNA and RNA. They are composed of nucleotides, which are the monomers made of three components: a 5-carbon sugar, a phosphate group and a nitrogenous base. If the sugar is a compound ribose, the polymer is RNA (ribonucleic acid); if the sugar is derived from ribose as deoxyribose, the polymer is DNA (deoxyribonucleic acid). Nucleic acids are the most important of all biomolecules. They are found in abundance in all living things, where they function to create and encode and then store information in the nucleus of every living cell of every life-form organism on Earth. In turn, they function to transmit and express that information inside and outside the cell nucleus—to the interior operations of the cell and ultimately to the next generation of each living organism. The encoded information is contained and conveyed via the nucleic acid sequence, which provides the 'ladder-step' ordering of nucleotides within the molecules of RNA and DNA. In this unit, we will discuss the structure and components of nucleic acid including RNA and DNA, along with the purine and pyrimidine bases of nucleic acid. It will also focus on the chemical basis of heredity and chemical synthesis of mono and trinucleoside.

5.1 OBJECTIVES

NOTES

After going through this unit, you will be able to:

- Describe the structure and components of nucleic acid including RNA and DNA
- Explain the purine and pyrimidine bases of nucleic acid
- Discuss the chemical basis of heredity and chemical synthesis of mono and trinucleoside

5.2 NUCLEIC ACIDS: STRUCTURE AND COMPONENTS

Nucleic acids constitute the most important biomolecules of the cell and are critical entities for all known forms of life. A nucleic acid is a macromolecule composed of chains of monomeric nucleotides. In biochemistry, these molecules carry genetic information or form structures within cells. The backbone of a nucleic acid is made of alternating sugar and phosphate molecules bonded together in a long chain. Each of the sugar groups in the backbone is attached to a third type of molecule called a nucleotide base.

A nitrogenous (nitrogen-containing) base is an organic compound that owes its property as a base to the lone pair of electrons of a nitrogen atom. Purines and pyrimidines make up the two groups of nitrogenous bases, including the two groups of nucleotide bases. The two most common nucleic acids are Deoxyribonucleic Acid (DNA) and Ribonucleic Acid (RNA).

5.2.1 Origin of Nucleic Acids

Nucleic acids were discovered by Friedrich Miescher in 1869. He reported that he had found a substance within the nuclei of human white blood cells, which was weakly acidic in nature and whose function was unknown. He had named this material as "nuclein". In 1874, Miescher was successfully able to separate nuclein into protein and nucleic acid components from salmon sperm nuclei. Nuclein was later named as nucleic acid in 1889 by Richard Altmann. They were so named because of their initial discovery from within the nucleus (~nucle), and due to the presence of phosphate groups in their molecules (phosphoric acid ~ ic acid) and it possesses acidic properties. In 1880s, Fischer identified purine and pyrimidine bases in nucleic acid. In 1881, Zacharis identified nucleic acid with chromatin. In 1882, Sachs discovered that nucleic acid of sperm and egg are different.

In 1894, Geheimrat Albrecht Kossel, recognized that histones ad protamins are associated with nucleic acid and the histones are the basic proteins. In 1910 he got Nobel Prize for the discovery of two purines and pyrimidine bases in nucleic acids. In 1914, Robert Feulgen discovered the method of colour test of DNA, which is called Feulgen test. In 1931, P.A. Levine defined that there are two types of nucleic acids; DNA and RNA. In 1941, Caspersson and Brachet independently stated that nucleic acid is directly related to protein synthesis.

In 1944, Oswald T. Avery, Colin M. Macleod and Maclyn McCarty stated that DNA is directly related to inheritance. In 1953, James D. Watson and Francis H.C. Crick constructed the double helical model for the DNA molecule. Before going on to studying the types of nucleic acids, you need to know what is gene and genome.

A gene is a unit of heredity in a living organism. It normally resides on a stretch of DNA that codes for a type of protein or for an RNA chain that has a function in the organism. All living things depend on genes. This is because they specify all proteins and functional RNA chains. Genes contain the information on which to build and maintain an organism's cells and pass genetic traits to offspring, although some organelles are self-replicating and not coded for by the organism's DNA. Most organisms have several genes corresponding to several different biological traits. Some of these traits are visible instantly, for example, the colour of the eye and the number of limbs. Some are not visible immediately, such as blood type or increased risk for specific diseases or the thousands of basic biochemical processes comprising life.

The word 'genome' is derived from the Greek word genome, which means 'I become, I am born, to come into being'. Some suggest it to be a mix of the gene and chromosome. In modern molecular biology and genetics, the genome is the hereditary information of an organism in totality. It is encoded either in DNA or for several types of viruses in RNA. The genome includes both the genes and the non-coding sequences of the DNA.

5.2.2 Types of Nucleic Acids

There are two types of nucleic acids; Deoxyribonucleic Acid (DNA) and Ribonucleic Acid (RNA). These two types of nucleic acids are present in all types of plants and animals. The viruses also contain DNA or RNA, but not both. DNA is present in the cell nucleus and nearly 90 per cent RNA is present in the cytoplasm and 10 per cent in the nucleolus.

5.2.3 Components of Nucleic Acid

The basic components of a nucleic acid include three different entities, namely a nitrogenous base, a sugar moiety and a phosphate group. These combine to give one unit of a nucleotide (discussed later), which are stacked in a nucleic acid molecule (Figure 5.1).

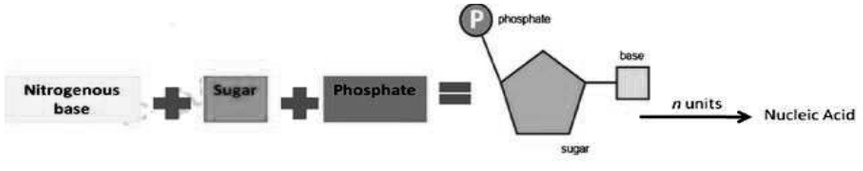


Fig. 5.1 *Components of Nucleic Acid*

The basic components of a nucleic acid are discussed in detail in sections below:

A. Nitrogenous Bases (Pyrimidine and Purine)

There are two types of nitrogenous bases present in all nucleic acids. These are pyrimidine and purine.

NOTES

Pyrimidine

These are heterocyclic aromatic compounds similar to pyridine and benzene. They possess six membered ring and two N atoms and three double bonds. Its melting point is 22^0C and boiling temperature of 123.5°C. Figure 5.2 shows the molecular structure of pyrimidine.

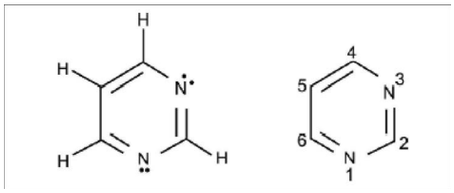


Fig. 5.2 *Molecular Structure of Pyrimidine.*

In nucleic acids, pyrimidines are of three types: (i) Cytosine (i) Thymine and (iii) Uracil.

(i) Cytosine ($C_5H_5ON_3$): It is found both in RNA and DNA. It is a white crystalline substance and its molecular weight is 111.12 daltons. Figure 5.3(i) shows the molecular structure of cytosine.

(ii) Thymine ($C_5H_6O_2N_2$): It is named thymine as it has being isolated from thymus. It is found in DNA and its molecular weight is 126.13 daltons. Figure 5.3(ii) shows the molecular structure of thymine.

(iii) Uracil ($C_4H_4O_2N_2$): It is found only in RNA. It is a white crystalline substance and its molecular weight is 111.10 daltons. Figure 5.3(iii) shows the molecular structure of uracil.

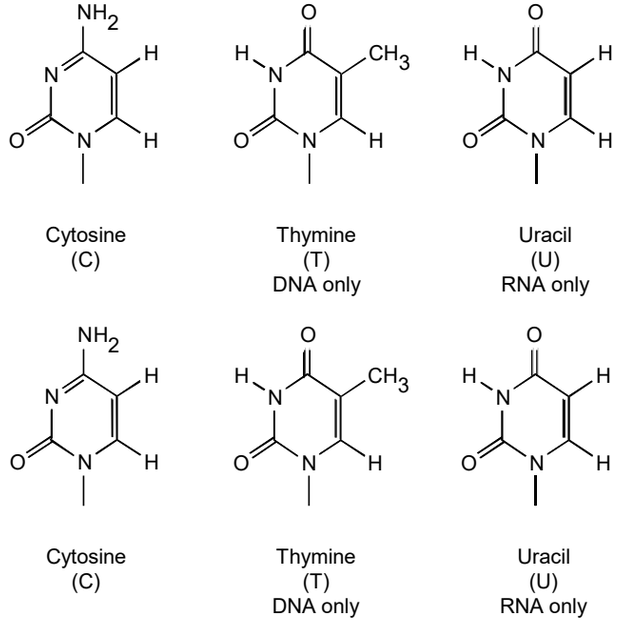


Fig. 5.3 *Pyrimidine: (i) Cytosine (ii) Thymine and (iii) Uracil*

Purine

In 1884, Purine name was coined by the German chemist Emil Fischer. Purine is a heterocyclic aromatic organic compound, in which a 6 membered pyrimidine ring is fused with an imidazole ring. Purine has a melting point of 216^{0}C. Figure 5.4 shows the molecular structure of purine. There are two different types of purines; adenine and guanine.

NOTES

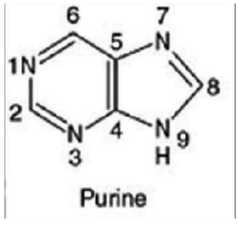


Fig. 5.4 *Structure of Purine*

(i) Adenine (C5H5N5): This is found in DNA as well as in RNA. It is a white crystalline substance and its molecular weight is 135.15 daltons. Its melting point is 360-365^{0}C. Figure 5.5 shows the molecular structure of adenine.

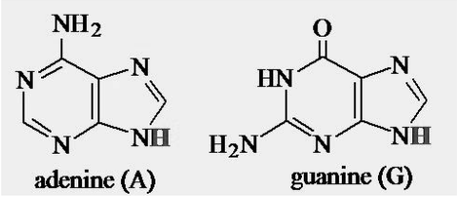


Fig. 5.5 *Molecular Structure of Adenine and Guanine*

(ii) Guanine (C5H5ON5): This is also present in DNA as well as in RNA. It is a white colourless, insoluble substance and its molecular weight is 151.15 daltons. Figure 5.5 shows the molecular structure of guanine.

Properties of Purines and Pyrimidines

The properties of purines and pyrimidines are:

- **Shape:** Purines and pyrimidines differ in their shape. The shape of the pyrimidine ring is planar, whereas the shape of the purine rings is nearly planar but exhibits some amount of puckering.
- **Solubility:** Purine and pyrimidine molecules are hydrophobic in nature and have a relatively low solubility in water near neutral pH. However, at acidic or alkaline pH, the purines and pyrimidines become charged, and their solubility therefore increases.
- **Chemical Properties:** They are conjugated molecules and weakly basic in nature.
- **Tautomerism:** Both purines and pyrimidines exhibit keto-enol tautomerism. The keto tautomer is known as a lactam ring, whereas the enol tautomer is known as a lactim ring (Figure 5.6). At neutral pH, the keto-tautomer remains the more predominanting form. On interaction with other molecules, ring nitrogens in the lactam serve as donors of hydrogen bond (H-bond), and the keto oxygens behave as H-bond acceptors.

NOTES

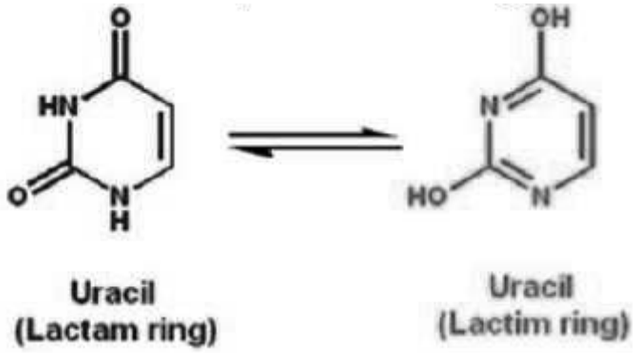


Fig. 5.6 Keto-enol Tautomerism in Uracil

- **Absorption**: As a consequence of aromatic ring structure and associated resonance, pyrimidine and purine bases absorb ultraviolet light (UV light), with an absorption maxima at a wavelength 260 nm (Figure 5.7). The measurement of the concentration of DNA or RNA in a given sample is therefore performed by measuring the UV absorbance at this wavelength.

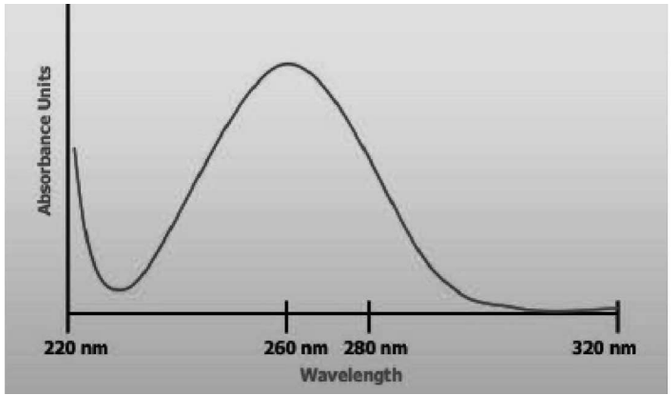


Fig. 5.7 An Absorption Spectra of Purified DNA Sample

B. Phosphate

Phosphate is another important component of the nucleic acid molecule. It gets attached to C-5' OH group of the sugar and gets incorporated into nucleic acid (both DNA and RNA).The molecular formula of phosphoric acid is H_3PO_5. It attaches three monovalent hydrogen atoms and three divalent oxygen atoms to the pentavalent phosphorus atom. Figure 5.8 shows the molecular structure of phosphoric acid.

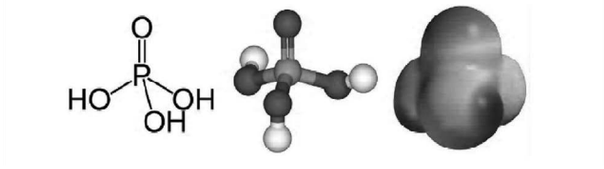


Fig. 5.8 Phosphoric Acid

C. Pentose Sugar

Two types of pentose sugars are found in nucleic acids, namely ribose and 2-deoxy ribose. The carbons in the ribose sugar are numbered according to convention. Ribose differs from deoxyribose in the presence of a hydroxyl group at the 2'C. The structures of both ribose and deoxyribose are shown in Figure 5.9. The D-ribose and D-deoxyribose are found in RNA and DNA respectively, in their furanose (closed five-membered ring) forms.

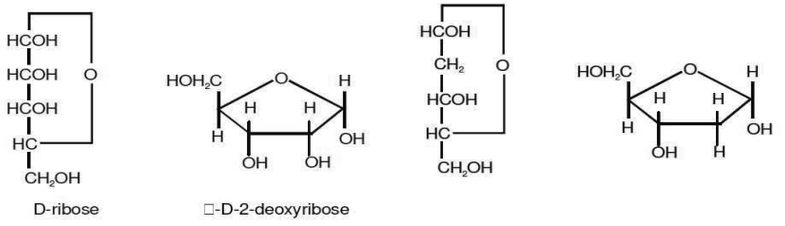


Fig. 5.9 *Molecular structure of (a) D-ribose and (b) D-2-deoxyribose*

The main difference between these two structures is that only D-2-deoxyribose sugar replaces OH atom at position 2 by an H atom. These two sugars are also differentiated by means of colour reaction.

5.2.4 Chemical and Enzymatic Hydrolysis of Nucleic Acids

Hydrolysis of nucleic acids by either chemical or enzymatic methods leads to the cleavage of phosphodiester backbone producing smaller oligonucleotide containing up to 20 residues, nucleosides, free purine or pyrimidine bases, ribose or deoxyribose and phosphates.

In enzymatic hydrolysis, cleavage can occur on either side of the phosphate. Enzyme hydrolyses the bond specifically. Enzymes that hydrolyse nucleic acids are called nucleases. They are needed for normal 'housekeeping' functions in biosynthesis and turnover of nucleic acids. The nucleases are also required for digestion of nucleic acids in the diet. Nucleases can of two types: Exonucleases and endonuclease. Exonuclease cleave the ends of the DNA and endonuclease act on regions in the middle of target molecules. They are further subcategorized as deoxyribonucleases and ribonucleases which act on DNA and RNA respectively. Restriction endonucleases are the ones that cleave DNA at the specific sequences recognized by the restriction enzymes.

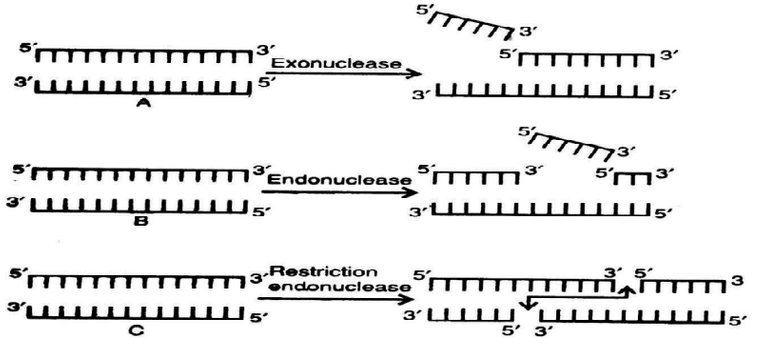


Fig. 5.10 *Enzymatic Actions of Exonuclease, Endonuclease and Restriction Nuclease on DNA Strand*

NOTES

The chemical hydrolysis consists of:

- **Acid Hydrolysis:** RNA is resistant to the effects of dilute acid, but gentle treatment of DNA with 1Mm HCl leads to hydrolysis of purine n-glycosidic bonds and causes depurination of DNA without disturbing the pyrimidine deoxyribose bonds or the phosphodiester bonds of the backbone. At other harsh chemical environments, selective removal of pyrimidine bases occurs. Both Nucleic acids can be hydrolysed to their constituent bases by the treatment with 72 per cent perchloric acid ($HClO_4$) for 1hour in many cases. The resulting nucleic acid derivative which is devoid of purine bases is called apurinic acid; while that devoid of pyrimidine bases is called apyrimidinic acid.
- **Alkali Hydrolysis:** DNA is not susceptible to alkaline hydrolysis. On the other hand, RNA is alkali labile and is voluntarily hydrolysed by dilute sodium hydroxide. Hydroxyl ion attacks on the 2'-OH group of sugar and yields cyclic-2', 3-phosphonucleoside intermediate that further hydrolyses to nucleoside-2'-phosphates and nucleoside-3'-phosphates.

5.2.5 Base Pairing of Purines and Pyrimidines

Purines and pyrimidines, being complementary bases, can participate in base pairing, based on the specific shapes and hydrogen bond properties. Guanidine, being a complement of cytosine, pairs with cytosine through three hydrogen bonds. Adenine (A) is the complement of thymine (T) in DNA and uracil (U) in RNA. Adenine base pairs with thymine and uracil through two hydrogen bonds. The pairings of the bases are as follows (Figure 5.11):

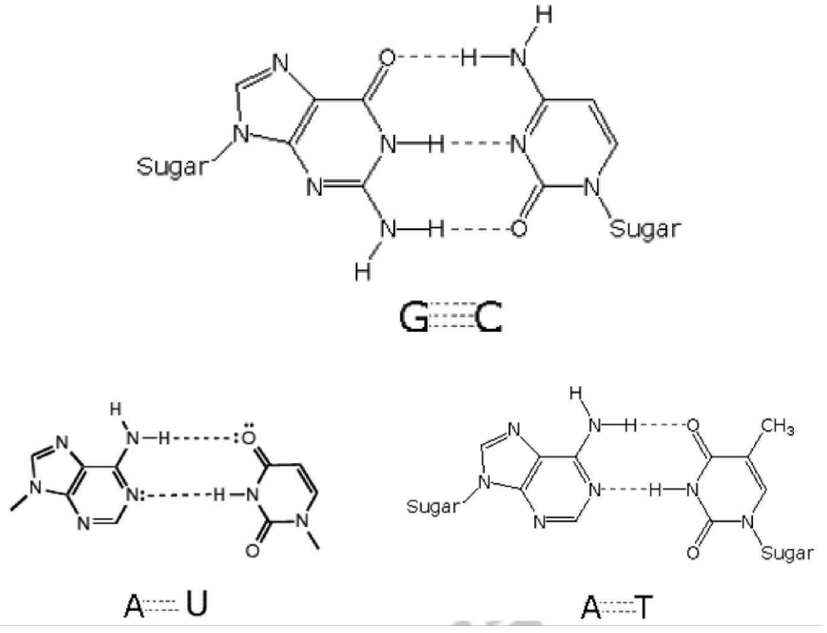


Fig. 5.11 *Base Pairing in Purines and Pyrimidines*

NOTES

Chargaff's Rule

Erwin Chargaff (1905-2002), an Austrian-American biochemist gave the Chargaff's rule, according to which DNA always contains equal amounts of certain base pairs. He observed that the amount of adenine (A) always equaled with the amount of thymine (T), and the amount of guanine (G) always equaled the amount of cytosine (C), regardless of the DNA source. The ratio of (A+T) to (C+G) varied from 2.70 to 0.35 in various organisms.

$$A + G = T + C$$

Nucleosides and Nucleotides

A nucleoside consists of a combination of a nitrogenous base and a sugar (ribose or deoxyribose). The bond between them is called the beta-glycosidic linkage. The position of attachment is shown below.

Nucleotides = nitrogenous base + sugar + phosphate
Nucleotides = Nucleosides + phosphate

Nucleotides comprises of a nitrogenous base linked to a 5-carbon sugar and one or more phosphate group. The phosphate is attached to 5' CH_2OH group of sugar part of nucleoside. They function as the building blocks of nucleic acids. The position of attachment is shown below in Figure 5.12.

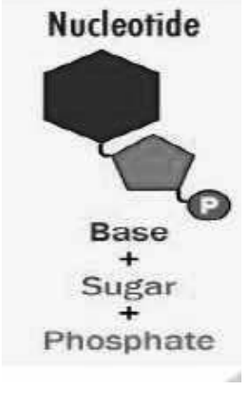


Fig. 5.12 *Nucleotide*

The base of a nucleotide (position N-1 of pyrimidines or N-9 of purines) is forms a covalent N—glycosyl bond with the 1' carbon of the pentose, by removal of a water molecule. The phosphate is esterified to the 5' carbon (Figure 5.13).

NOTES

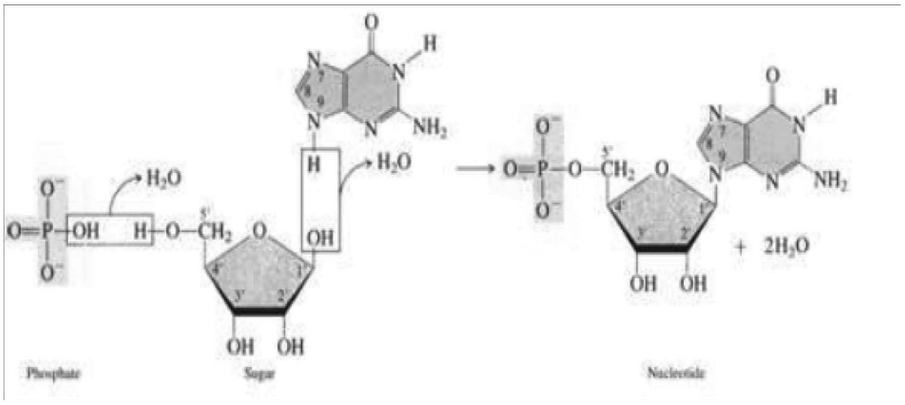


Fig 5.13 *Formation of Nucleotides*

Examples of nucleotides include deoxyadenosine monophosphate, deoxycytidine monophosphate, deoxyguanoside monophosphate, deoxythymidine monophosphate (Figure 5.14).

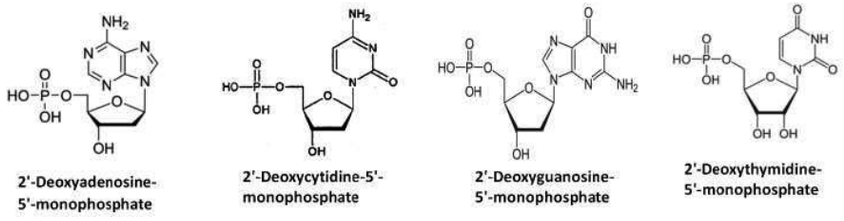


Fig. 5.14 *Examples of Nucleotides*

Nucleotide Di- and Tri-Phosphates

The term 'nucleotide' generally refers to a nucleoside monophosphate, But in case additional phosphoric acid groups are present, they can link to the existing phosphate (in nucleotide monophosphates) to produce nucleotide diphosphates and nucleotide tri-phosphates (Figure 5.15).

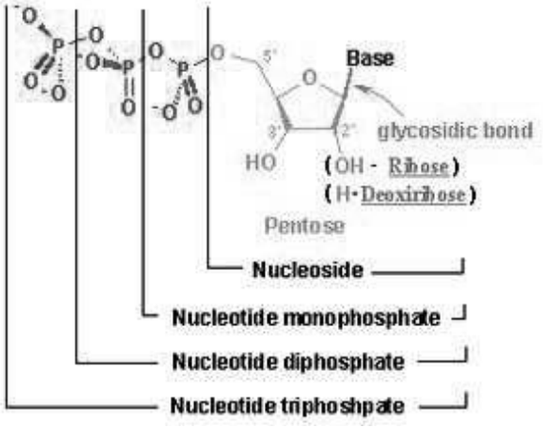


Fig. 5.15 *Nucleoside Mono, Di and Triphosphates*

NOTES

Cyclic Nucleotides

Nucleoside monophosphates can sometimes form two ester bonds with the phosphoric acid, at the 5' and 3' hydroxyl groups of the ribose sugar. This therefore results in the formation of cyclic nucleotides. These are designated as cNMP where 'c' stands for cyclic and 'N' stands for the respective nucleoside.

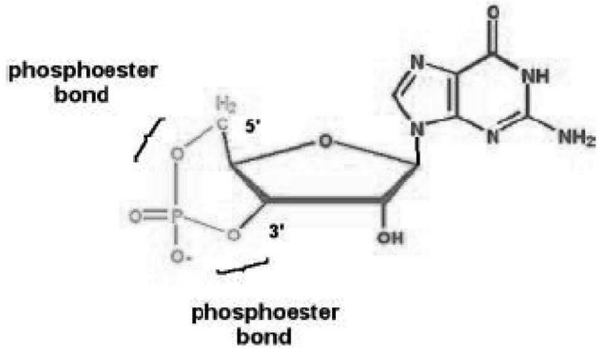


Fig. 5.16 *3', 5'-Cyclic Guanosine Monophosphate (3', 5'-GMP)*

Figure 5.16 above shows a cyclic guanosine monophosphate (cGMP). Cyclic AMP (cAMP) and cyclic GMP (cGMP) are the two of the well-studied cyclic nucleotides. These are found in all cells and play important role in the regulation of cell metabolism.

Polynucleotides

A polynucleotide formation is initiated when many nucleotides continue to join together by phosphodiester linkages. The formation of polynucleotide is catalyzed by polymerase enzymes (DNA polymerase in case of DNA or RNA polymerase in case of RNA). The -OH group on the 3'-carbon of sugar in one nucleotide reacts with the phosphate attached to the 5'-carbon of another to form phosphate ester bonds and a dinucleotide. Repeated formation of such bonds leads to further elongation of the polynucleotide chain.

DNA and RNA are examples of polynucleotides (Figure 5.17), where the nucleotides are arranged in linear way and proceeds in the 5' ——> 3' direction. A common representation of polynucleotide is given below:

5'pApTpGpC OH_3'

While RNA is single stranded, DNA is double stranded and contains two such polynucleotide chains spiraling round each other to form a double helical structure. The two chains in the double helix are held together by hydrogen bonds by complementary bases on different chains.

NOTES

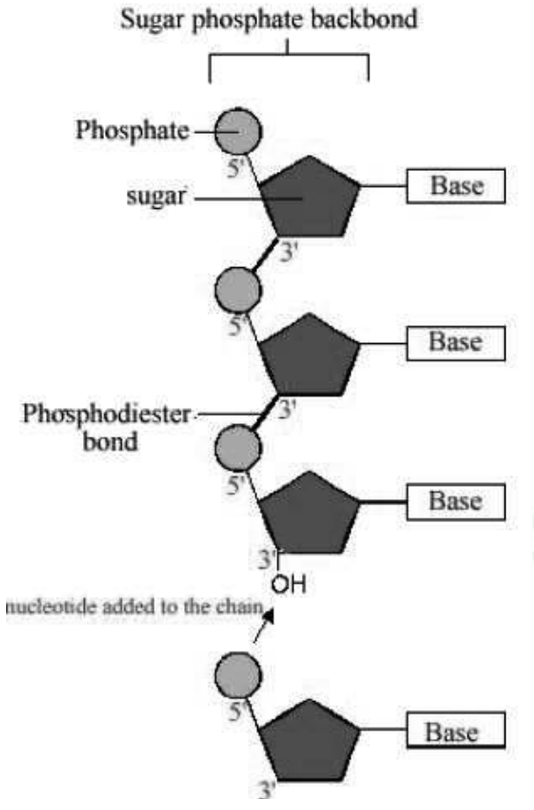


Fig. 5.17 Formation of Polynucleotides

5.3 STRUCTURE OF DNA

In the early 1950s, four scientists, James Watson and Francis Crick at Cambridge University and Maurice Wilkins and Rosalind Franklin at King's College, determined the true structure of DNA from data and X-ray crystallography of a molecule taken by Franklin. In 1953, Watson and Crick published a paper in the scientific journal Nature describing their research. They showed that DNA molecule is not only double-stranded, but also has two strands attached around each other forming a coil, or helix. A double helix represents the exact structure of DNA.

Except viruses, in most living organisms, genetic information is stored in DNA. DNA resides in the nucleus of living cells. Four different nucleotide bases occur in DNA, namely adenine (A), cytosine (C), guanine (G) and thymine (T).The nucleotide bases of the DNA molecule form complementary pairs with the nucleotides hydrogen bond to another nucleotide base in a strand of DNA, opposite to the original. This bonding is specific, with adenine always pairing with thymine (and vice versa) and guanine always pairing with cytosine (and vice versa).

A unique feature of DNA molecule is that it is able to make exact copies of itself, or is self-replicate. When an organism requires more DNA, such as during reproduction or cell growth, a break in the hydrogen bonds between the nucleotide bases takes place, leading to the separation of two single strands of DNA. New complementary bases are brought in by the cell and paired up with each of the

two separate strands, thus forming two new, identical, double-stranded DNA molecules.

A DNA molecule consists of two long polymers, which face in the opposite direction and are known as 'anti-parallel'. This means that in a double helix the direction of the nucleotides in one strand is opposite to the direction of the other strand.

NOTES

Every sugar moiety is connected to one of the four types of molecules known as bases. Each stand has polarity, that is, a top and a bottom that imparts 3' (three prime) and 5' (five prime) ends. Information is encoded within the sequence of these bases along the backbone. The 5' end has a terminal phosphate group while whereas the 3' end has a terminal hydroxyl group. Hydrogen bonds that bind the bases attached to the two strands help in stabilizing the DNA double helix. Figure 5.18 shows the DNA molecular topography.

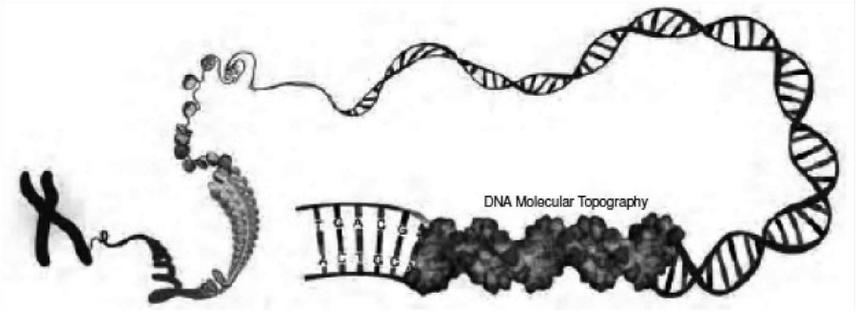


Fig. 5.18 *DNA Molecular Topography*

Chemically, DNA is a long polymer that is comprised of nucleotide units. The DNA chain is 22–26 Å (Ångströms) wide and each nucleotide unit is 3.3 Å long (Figure 5.19). DNA polymers are large molecules that may have millions of nucleotides.

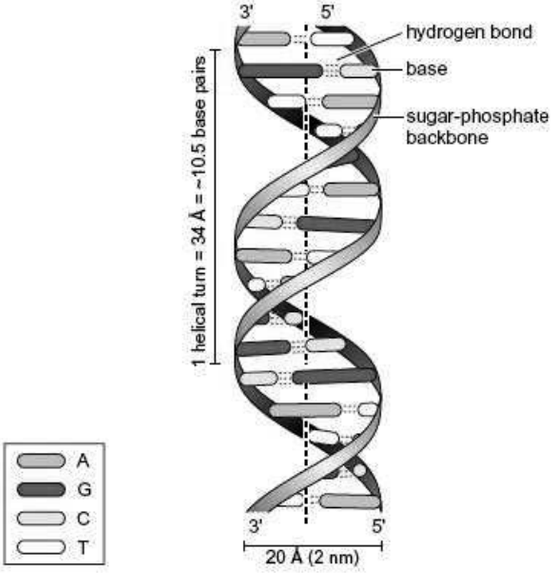


Fig. 5.19 *Double Helix Structure of DNA*

NOTES

Grooves

A twisted DNA molecule shows two types of grooves that are different based on size. The major groove is 22 Å wide while the minor groove is 12 Å wide. Because it is narrow, the edges of the bases are less accessible in the minor groove in comparison to major groove that leads to lesser interaction. Proteins such as transcription factors that can bind to specific sequences in double-stranded DNA therefore make contacts to the sides of the bases exposed in the major groove normally.

Base Pairing

Complementary base pairing is when each type of base on one strand establishes a link with exactly one type of base on the other strand. Purines form hydrogen bonds with pyrimidines. The arrangement of two nucleotides binding together across the double helix is known as a base pair. The double helical DNA strands can be pulled apart like a zipper either by mechanical force or by applying high temperature. The concept of 'complementarity' causes duplication of all information in the double-stranded sequence of a DNA helix, which is essential in DNA replication. This reversible, specific interaction between complementary base pairs is crucial for all DNA functions in living organisms

The two types of base pairs are able to form various numbers of hydrogen bonds. The AT base pair forms two hydrogen bonds while the GC base pair forms three hydrogen bonds. DNA that has high GC content is more stable than DNA that has low GC content. The DNA double helix that separates easily tends to have a high AT content. The strength of the interaction can be calculated by finding the temperature that is needed to break hydrogen bonds. This is known as the melting temperature or Tm value. When all the base pairs in a DNA double helix melt, the strands separate and are there in solution as two completely independent molecules.

Sense and Anti-Sense

DNA sequence is referred to as 'sense' if its sequence is the same as that of a messenger RNA copy that is translated into protein. Its opposite is known as the 'antisense' when the sequence is on the opposite strand. Some DNA sequences in prokaryotes and eukaryotes, and more in plasmids and viruses, distort the difference between sense and anti-sense strands by having overlapping genes.

Supercoiling

The process of DNA supercoiling involves twisting DNA into a rope-like structure. If the DNA is twisted in the direction of the helix, it is known as positive supercoiling. The bases are held more tightly together but if they are twisted in the opposite direction, it is known as negative supercoiling. This way the bases come apart more easily. Most DNA have slightly negative supercoiling introduced by enzymes called topoisomerases.

5.3.1 Different Forms of DNA Structure

The many conformations in which DNA exists include A-DNA, B-DNA and Z-DNA (Figure 5.20). The DNA molecule that Watson and Crick described was in

NOTES

the B form. The conformation adopted by DNA is dependent on the hydration level, DNA sequence, amount and direction of supercoiling, chemical modifications of bases, type and concentration of metal ions and the presence of polyamines in solution. As compared with B-DNA, A-DNA forms a wider, right-handed spiral with a shallow, wide minor groove and a narrower, deeper major groove.

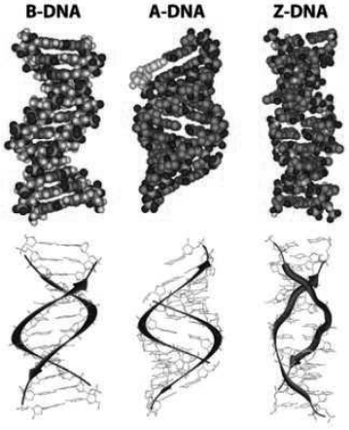


Fig. 5.20 Structures of B, A, and Z DNA (from Left to Right)

The A form occurs under non-physiological conditions in partially dehydrated samples of DNA. DNA segments, where the bases have been chemically modified by methylation, may undergo a larger change in conformation and adopt the Z form. Here, the strands turn about the helical axis in a left-handed spiral, the opposite of the more common B form (Table 5.1).

Table 5.1 *Features of Different Forms of DNA*

Feature	B-DNA	A-DNA	Z-DNA
Type of helix	Right-handed	Right-handed	Left-handed
Helical diameter (nm)	2.37	2.55	1.84
Rise per base pair (nm)	0.34	0.29	0.37
Distance per complete turn (pitch) (nm)	3.4	3.2	4.5
Number of base pairs per complete turn	10	11	12
Topology of major groove	Wide, deep	Narrow, deep	Flat
Topology of minor groove	Narrow, shallow	Broad, shallow	Narrow, deep

5.3.2 Structure and Properties of RNA

Ribonucleic Acid (RNA) has derived its name from the sugar group present in its backbone, that is, ribose. Some of the features of DNA and RNA are common. However, there are also features that differentiate DNA from RNA. Both have a sugar-phosphate backbone with nucleotide bases attached to it. Adenine, cytosine and guanine are common in both. However, RNA lacks thymine. It contains uracil in place of thymine. Further, DNA is a double-stranded molecule, whereas RNA is a single-stranded molecule.

RNA is the key genetic material used in microorganisms, such as viruses. It also plays an important role in the production of proteins in other living organisms.

NOTES

Due to its ability to move around the cells of living organisms, it functions as a genetic messenger, transmitting the information contained in the cell's DNA to other parts of the cell. This transmission helps in protein synthesis.

RNA is a biologically important molecule that consists of a long chain of nucleotide units. Each nucleotide has a nitrogenous base, a ribose sugar and a phosphate. RNA is transcribed from DNA by enzymes known as RNA polymerases and is usually processed further by other enzymes. RNA is considered very important in the synthesis of proteins. Messenger RNA (mRNA), a type of RNA, carries information from DNA to ribosomes, made of proteins and ribosomal RNAs that bind together to form a molecular machine that can read mRNAs and translate the information they carry into proteins. Other RNAs with diverse roles exist – in particular, these regulate which genes are expressed and are also the genomes of most viruses.

Each nucleotide in RNA comprises ribose sugar, with carbons numbered from 1' to 5'. A base is attached to the 1' position, generally adenine (A), cytosine (C), guanine (G) or uracil (U). Adenine and guanine are purines while cytosine and uracil are pyrimidines. A phosphate group is attached to the 3' position of one ribose and the 5' position of the next ribose. Phosphate groups have a negative charge at physiological pH, making RNA a charged molecule. The bases may form hydrogen bonds between cytosine and guanine, between adenine and uracil and between guanine and uracil (Figure 5.21).

The functional form of single-stranded RNA molecules frequently asks for a certain tertiary structure. The basis for this structure is given by secondary structural elements that are hydrogen bonds within the molecule. This causes many recognizable secondary structure 'domains', such as hairpin loops, bulges and internal loops. Since RNA is charged, metal ions such as Mg^{2+} are required to stabilize many secondary and tertiary structures.

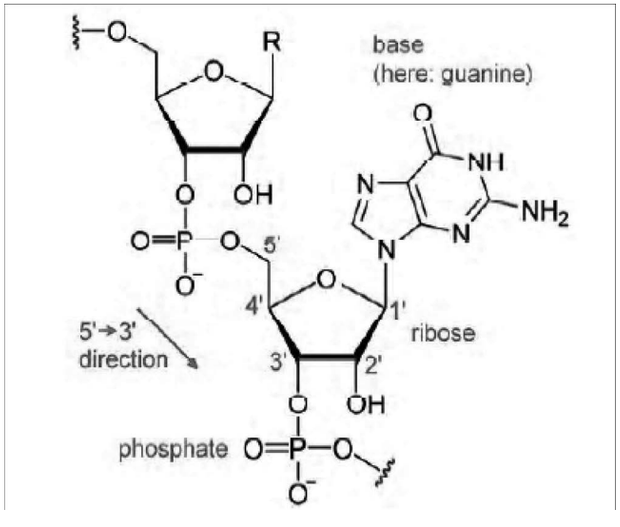


Fig. 5.21 *RNA Nucleotide with Ribose Sugar, Phosphate and Base*

Most biologically active RNAs, such as mRNA, tRNA, rRNA, snRNAs and other non-coding RNAs, have self-complementary sequences that enable

parts of the RNA to fold and pair with itself. This causes double helices to form. Structural analysis of these RNAs shows that they are highly structured and their structures, unlike DNA, do not consist of long double helices but rather collections of short helices packed together into structures similar to that of proteins. Thus, RNAs can attain chemical catalysis like enzymes do.

Types of RNA

There are three types of RNA, which are as follows:

1. **Ribosomal RNA (rRNA) (Insoluble RNA):** This form of RNA has the highest molecular weight. It is the most abundant RNA of all types of RNAs. There are two different types of RNA in the prokaryotic and in eukaryotic cells. In prokaryotes the ribosome possesses two different subunits, large and small. Large subunits (50 s) contains two types of rRNA: 23 s and 5 s named upon the sedimentation behavior. The smaller subunit (30 s) contains 16 s types. The ribosome of eukaryotes also possesses three different types of rRNAs named on the sedimentation behavior. Larger subunit contains 5 s and 28 s types. The smaller subunit contains 18 s types. The mammalian rRNA possesses 5 s, 5.8 s and 28 s types in larger subunit and in smaller subunit it contains only one 18 s types. Figure 5.22 shows the structure of rRNA.

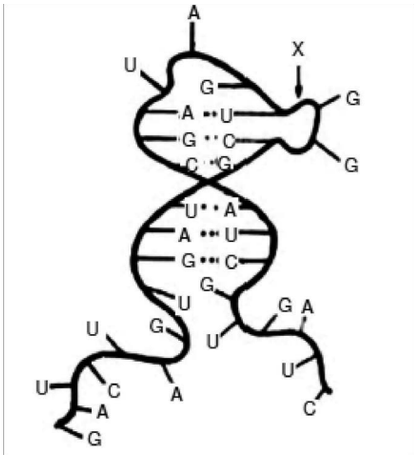


Fig. 5.22 *Structure of rRNA*

2. **Messenger RNA (mRNA) (Template RNA):** Messenger RNA is the most heterogeneous in size and stability. Its amount is 5 per cent of total RNA of the cell. It is synthesized on the surface of the DNA template. This carries the genetic information to assemble the amino acids from DNA to ribosomes; hence it is called messenger RNA. In prokaryotes the mRNA is unstable rather than in eukaryotes. The mRNA is complementary to the DNA template strand. In eukaryotes there are specific 5' cap, which is the recognition site for ribosome subunit, whereas in prokaryotes there is no specific 5' recognition site. As a result, there can be many ribosomal binding sites in prokaryotes, each one giving rise to a different type of protein.

NOTES

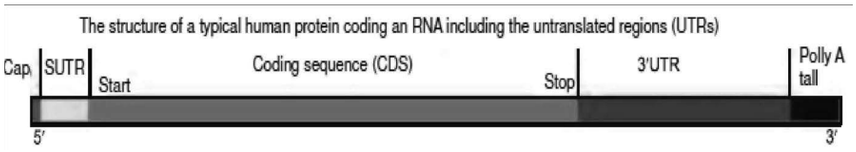


Fig. 5.23 *Structure of mRNA*

In mammals the 5' cap is methylated that means a cap of 7-methyle guanosine triphosphate is linked to the 5' end. The 3' end of mRNA possesses 20 - 250 nucleotides of adenylate residue, which is called poly A tail. This tail probably works as a stabilizer in the cell. Figure 5.23 shows the structure of mRNA.

3. **Transfer RNA (tRNA) (Soluble RNA):** It is the smallest polymeric form of RNA. It is nearly 15 per cent of total RNA of the cellular RNA. The most important function of tRNA is to act as a specific carrier of activated amino acids to specific site on the protein synthesis. Now, nearly 50 sequences are known. Robert W. Holley presented the clover leaf model for the tRNA. Figure 5.24 shows the structure of tRNA.

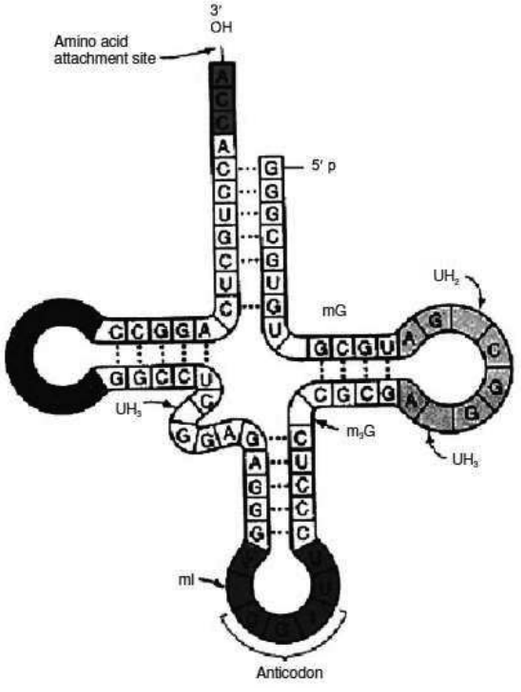


Fig. 5.24 *Structure of tRNA*

The common structural features of tRNA are as follows:

(i) All tRNAs have a common design, consisting of three folds. This gives them a shape of a clover leaf.

(ii) The molecular weight of all tRNA molecules range from 24,000 to 31,000.

(iii) tRNAs contain 7-15 unusual bases. Many of them are methylate of unmethylated derivatives of A, U, G, C.

(iv) The 5' end is phosphorylated and terminal residue is usually guanylate.

(v) The base sequence of 3' end of all tRNAs is CCA.

(vi) Nearly 60 per cent of the nucleotides in tRNAs bases are paired to form double helices.

(vii) There are 5 groups of bases which are not base paired. These forms loops which are as follows:

- 3' CCA terminal region
- The ribothymine-pseudouracil-cytosine (TOC) loop
- The extra arm or little loop
- The diyhdrouracil loop (DHU)
- Anticodon loop

(viii) The unique feature of all the tRNAs is that distance between CCA to anticodon is constant.

5.3.3 Organization of DNA in Cell

DNA is present in cells in the long structures in the nucleus, also known as chromosomes. These chromosomes duplicate before the cells divide and this process is known as DNA replication. Chromosomes are condensed thread-like structures of DNA. DNA is present in the eukaryotes inside the cell nucleus. A small amount of DNA is also available in cell organelles like mitochondria or chloroplasts.

On the other hand, in prokaryotes like bacteria and archaea, DNA is present in the cytoplasm. Chromatin proteins, such as, histones, compact and organize DNA into chromosomes.

Organization of DNA in Prokaryotes

In prokaryotes, the genome is composed of a single, double-stranded DNA molecule in the form of a loop or circle (Figure 5.25). The region in the cell containing this genetic material is called a nucleoid (remember that prokaryotes do not have a separate membrane-bound nucleus). Some prokaryotes also have smaller loops of DNA called plasmids that are not essential for normal growth.

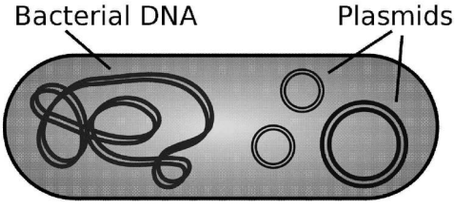


Fig. 5.25 *Bacterial DNA and Plasmids*

Bacteria can exchange these plasmids with other bacteria, sometimes receiving beneficial new genes that the recipient can add to their chromosomal DNA. Antibiotic resistance is one trait that often spreads through a bacterial colony through plasmid exchange. Most bacteria (including E. coli) lead existences as individual cells, but in some bacterial species cells tend to associate in clusters or filaments, and a few (the mycobacterium, for example) demonstrate simple social behaviour.

NOTES

Organization of DNA in Eukaryotes

The DNA in a diploid human cell is nearly 2 m long. To fit into cells, the DNA is tightly packaged into chromatin. Chromatin are consists of DNA and proteins. The lowest level of packaging is the nucleosome, which consists of DNA wrapped around histone proteins which are small and basic proteins rich in amino acids such as lysine and/or arginine.

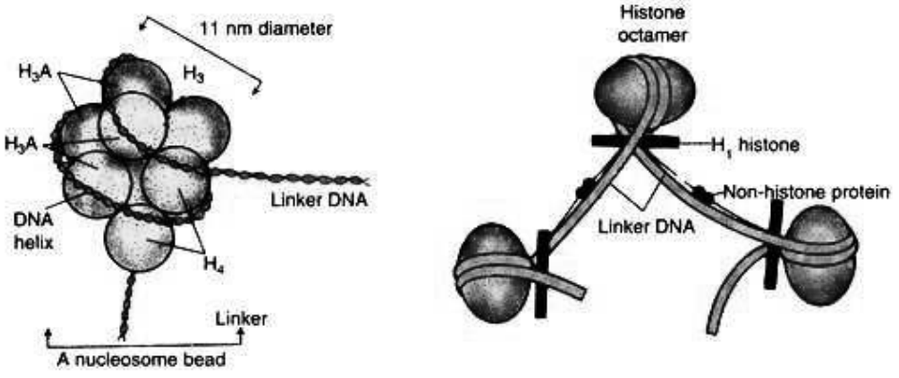


Fig. 5.26 *Nucleosome*

Almost in all eukaryotic cells, there are five types of histones e.g. H_1, H_2A, H_2B, H_3 and H_4. Eight histone molecules (two each of H_2A, H_2B, H_3 and H_4) form an octamer ellipsoidal structure of about 11 nm long and 6.5-7 nm in diameter. DNA coils around the surface of ellipsoidal structure of histones 166 base pairs (about 7/4 turns) before proceeding onto the next and form a complex structure, the nucleosome (Figure 5.26). Thus a nucleosome is an octamer of four histone proteins complexed with DNA.

The histones play an important role in determining of eukaryotic chromosomes by determining the conformation known as chromatin. The nucleosomes are the repeating units of DNA organization which are often termed as beads. The DNA isolated from chromatin looks like string or beads.

The 146 base pairs of DNA lie in the helical path and the histone-DNA assembly is known as the nucleosome core particle. The stretch of DNA between the nucleosomes is known as 'linker' which varies in length from 14 to over 100 base pairs.

The H1 is associated with the linker region and helps the folding of DNA into complex structure called chromatin fibres which in turn get coiled to form chromatin. As a result of maximum folding of DNA, chromatin becomes visible as chromosomes during cell division.

5.4 CHEMICAL BASIS FOR HEREDITY AND OVERVIEW OF REPLICATION OF DNA

The formation of new DNA chains from raw material including one original DNA molecule is known as DNA replication. It acts as the basis for biological inheritance.

Besides DNA chain synthesis, DNA replication also includes initiation and termination of the synthesis. Watson and Crick proposed a mechanism for the self-duplication of DNA, during which the two strands of double helix separated gradually by progressive breakdown of hydrogen bonds between the nitrogenous bases, with each strand containing information for the duplication of its complementary strand.

5.4.1 Features of DNA Replication

The feature of DNA replication include:

- **Semiconservative Nature of DNA Replication:** DNA is made up of a double helix of two complementary strands and as suggested by Watson and Crick, the two strand of DNA molecule separate and each strand serves as a template alongside which a new DAN strand is synthesized following the base pairing rules (A≡T and G=C). As a result of semi-conservative replication, the new helix will be composed of an original DNA strand and a newly synthesized DNA strand. Since only one parent strand is conserved in each newly formed daughter strand, this mode of DNA replication is known as semiconservative. Cellular proofreading and error-checking mechanisms ensure near perfect fidelity for DNA replication.
- **DNA Replication is Bidirectional and forms Replication Bubble:** In a cell, DNA replication begins at specific locations or Origins of Replication (ORI) in the genome. Unwinding of DNA at the Origin and Synthesis of new strands, accommodated by an enzyme known as helicase, results in replication forks growing bi-directionally from the origin. A number of Proteins are associated with the replication fork to help in the initiation and continuation of DNA synthesis. Most prominently, DNA polymerase synthesizes the new strands by adding nucleotides that complement each (template) strand. DNA replication occurs during the S-stage of interphase (Refer Figure 5.27).

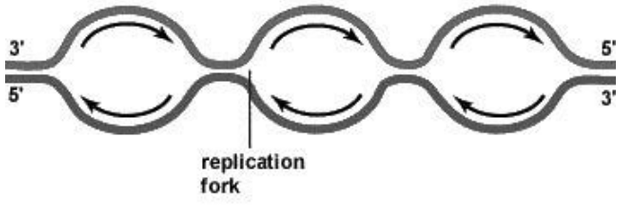


Fig. 5.27 *Bidirectional Replication from Many Origins in Linear Eukaryotic Chromosome*

- **DNA Replication Occurs in 5' → 3' Direction:** DNA strands have a directionality, the two strands are antiparallel as they have two opposite strands differ in chemical polarity. One strand of the double helical DNA has 5' → 3' polarity while other has 3' → 5'. That is, if the base sequence of a single strand of DNA is given, the left end of the sequence is the 5' end, while the right end of the sequence is the 3' end. DNA polymerase can add nucleotide molecules in 3' end of a growing DNA strand, thus the replication

NOTES

of new daughter DNA strands along with the old parent strand must occur in the 5' → 3' direction (Refer Figure 5.28).

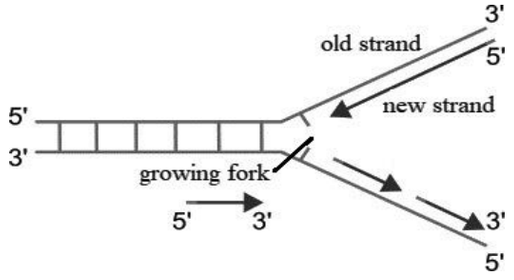


Fig. 5.28 *Unidirectional Replication of Single DNA Strand*

- **DNA Replication Takes Place as Continuous Leading Strand and Discontinuous Lagging Strand:** Since the strands of DNA duplex molecule are tightly winded with each other and are antiparallel, as a result cannot be pulled apart through their entire length at once. For both the strands to be copied at a growing fork, one of the two new strands must be synthesized as a discontinuous strand (grows in segments towards 5' → 3' direction). This strand is also known as Lagging Strand and segments of lagging strand are later joined together by DAN ligase. The other new DNA strand is synthesized continuously at the growing fork in 5' → 3' direction is called Leading Strand (Refer Figure 5.29).

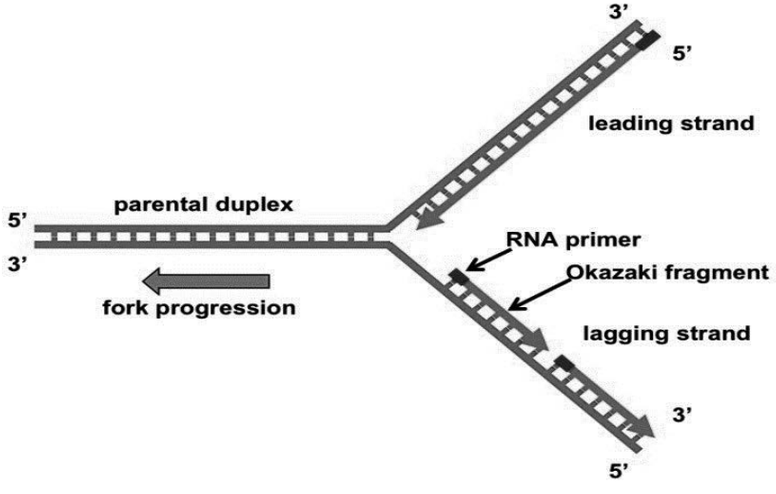


Fig. 5.29 *Formation of Leading and Lagging Strands in DNA Replication*

- **RNA Primer is Required to Initiate DNA Synthesis**: DNA polymerase cannot initiate the synthesis of new DNA strand on its own, it requires RNA primer (with 3' OH end) to initiate DNA synthesis. The RNA primers are formed by the enzymes called RNA primase. Once RNA primers are located on the newly exposed strand, new DNA strands then grow from the RNA primer in the 5' → 3' direction by adding nucleotides by DNA polymerase.

NOTES

The resulting hybrid short nucleotide (up to 1000) segments of RNA and DNA are known as Okazaki fragments. As an Okazaki fragment reaches the 5' end of the adjacent fragment synthesized earlier, the RNase removes the RNA primer by its additional 5' → 3' exonuclease property. The gaps are then filled by adding deoxyribonucleotides by DNA polymerase. Lastly, DNA ligase joins the adjacent completed fragments into a new daughter DNA strand (Refer Figure 5.30).

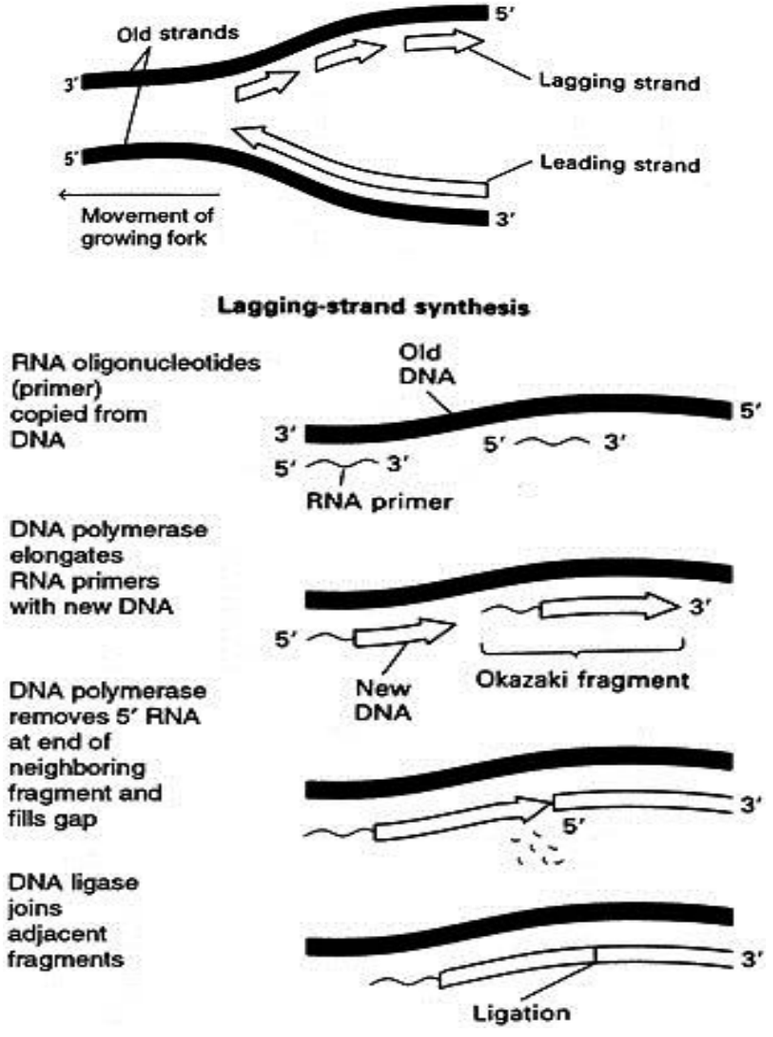


Fig. 5.30 *Synthesis of Lagging Strand*

5.4.2 Process of DNA Replication

DNA replication is coordinated process that is assisted and catalysed by many enzymes and proteins. It occurs in three steps: Initiation, Elongation and Termination.

NOTES

Initiation

For a cell to divide, it must first replicate its DNA. In Eukaryotes it occurs in nucleus during S phase of the cell cycle when chromosomes are in their extended form and are not readily visible. In the initial step of replication, activation of deoxyribonucleoside monophosphate occurs which are found free floating in nucleus and serve as raw material. Deoxyribonucleoside Monophosphate (dAMP, dGMP, dCMP, dTMP) combines with phosphate of ATP by an enzyme, phosphorylase and produce deoxyribonucleoside triphosphates (dATP, dGTP, dCTP, dTTP). This reaction is called Phosphorylation.

In the next step there occur exposure of parent DNA bases. The double helix parent DNA uncoils into two single strands by breakdown of hydrogen bonds. The unwinding of tightly coiled DNA is done by enzyme helicase, using ATP hydrolysis as energy source. Further unwinding is assisted and stabilized by Single Stranded DNA Binding proteins (SSB). Which binds strongly to the site. Unwinding creates a fork in the DNA, with two single strands trailing behind the fork and an intact double stranded helix in front of it. Supercoils produced by unwinding are removed by another group of enzymes, known as topoisomerase. Topoisomerase II or gyrase may cut and re-join one strand of DNA to facilitate uncoiling by enabling free ends to rotate around other chain. The DNA strands start uncoiling and separating at particular points in the DNA known as origin of replication which are targeted by initiator Proteins. Once the origin has been located, the initiator Proteins recruit other Proteins and form the pre-replication complex, which unwinds the double-stranded DNA.

Elongation

During this step, the two single parent strands are acted by DNA polymerase and RNA primer for addition of new polynucleotides. DNA polymerase has 52 –32 activity. DNA replication systems require a free 3' hydroxyl group before synthesis can be initiated (DNA template is read in 32 to 52 direction whereas a new strand is synthesized in the 52 to 32 direction).

As helicase unwinds DNA at the replication fork, the DNA ahead is forced to rotate. This process results in a build-up of twists in the DNA ahead. This build-up forms a torsional resistance that would eventually halt the progress of the replication fork. Topoisomerases are enzymes that temporarily break the strands of DNA, relieving the tension caused by unwinding the two strands of the DNA helix; topoisomerases (including DNA gyrase) achieve this by adding negative supercoils to the DNA helix. Bare single-stranded DNA tends to fold back on itself forming secondary structures; these structures can interfere with the movement of DNA polymerase. To prevent this, single-strand binding proteins bind to the DNA until a second strand is synthesized, preventing secondary structure formation (Refer Figure 5.31).

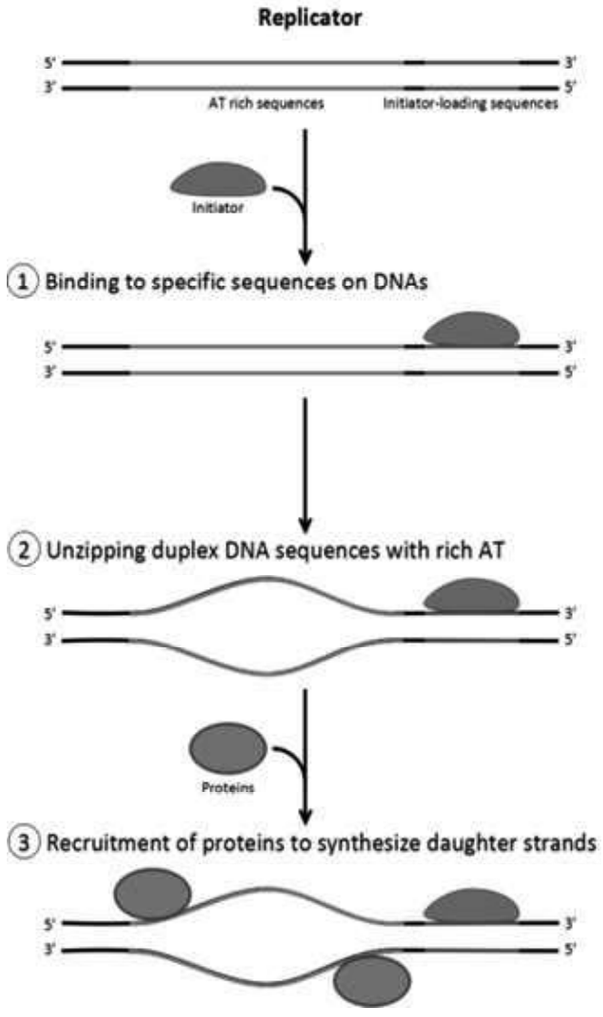


Fig. 5.31 *Role of Initiators for Initiation of DNA Replication*

Clamp Proteins form a sliding clamp around DNA, helping the DNA polymerase maintain contact with its template, thereby assisting with processivity. The inner face of the clamp enables DNA to be threaded through it. Once the polymerase reaches the end of the template or detects double-stranded DNA, the sliding clamp undergoes a conformational change that releases the DNA polymerase. Clamp-loading proteins are used to initially load the clamp, recognizing the junction between template and RNA primers.

Once the two strands are separated, RNA primase adds RNA primers to the template strands. On leading strand (continuous strand) one RNA primer is added while on lagging strand (discontinuous strand) many RNA primers are added. Once RNA primers are loaded on the newly exposed strand, new DNA strands then grow from the RNA primer in the 5' $\rightarrow$ 3' direction by adding nucleotides by DNA polymerase. The resulting hybrid short 1000 nucleotide segments of RNA and DNA are known as Okazaki fragments. As an Okazaki fragment reaches the 5' end of the adjacent fragment synthesized earlier, the RNase removes the RNA primer by exonuclease property. Thus, the leading strand is continuously extended from the primer by a DNA polymerase, while the lagging strand is extended

NOTES

discontinuously from each primer forming Okazaki fragments. After completion, a single nick on the leading strand and several nicks on the lagging strand are filled by ligase thus completing the newly replicated DNA molecule.

The primase used in this process differs significantly between Bacteria and Eukaryotes. Bacteria use a primase belonging to the DNAG protein superfamily which contains a catalytic domain of the TOPRIM fold type. The TOPRIM fold contains a α/β core with four conserved strands in a Rossmann-like topology. This structure is also found in the catalytic domains of topoisomerase IA, topoisomerase II, the OLD-family nucleases and DNA repair Proteins related to the RecR protein. In Eukaryotes, the primase contains the RNA Recognition Motif (RRM). This primase is structurally similar to many viral RNA-dependent RNA polymerases, reverse transcriptases, cyclic nucleotide generating cyclases and DNA polymerases of the A/B/Y families that are involved in DNA replication and repair. In eukaryotic replication, the primase forms a complex with Pol α (DNA polymerase).

DNA Polymerase

The primary function of DNA polymerase is to add nucleotides to the 3' end of a DNA strand. Multiple DNA polymerases take on different roles in the DNA replication process. In *E. coli*, DNA Pol III is the polymerase enzyme primarily responsible for DNA replication. It is a large, asymmetrical dimer containing two copies of most subunits. It has two catalytic sites which add nucleotides to the leading and lagging strand simultaneously. It assembles into a replication complex at the replication fork that exhibits extremely high processivity, remaining intact for the entire replication cycle. In contrast, DNA Pol I is the enzyme responsible for replacing RNA primers with DNA. DNA Pol I has a 5′ to 3′ exonuclease activity in addition to its polymerase activity, and uses its exonuclease activity to degrade the RNA primers ahead of it as it extends the DNA strand behind it, in a process called Nick Translation. Pol I is much less processive than Pol III because its primary function in DNA replication is to create many short DNA regions rather than a few very long regions.

In Eukaryotes, the low-processivity enzyme, Pol α, helps to initiate replication because it forms a complex with primase. In Eukaryotes, leading strand synthesis is thought to be conducted by Pol ε and Pol δ. Primer removal is completed Pol δ while repair of DNA during replication is completed by Pol ε (Refer Figure 5.32).

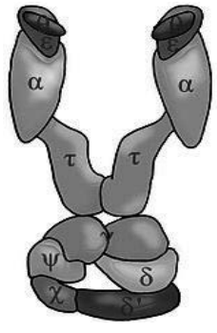


Fig. 5.32 *DNA Polymerase III in Prokaryotes*

DNA Replication Proteins

At the replication fork, accessory enzymes assemble on the DNA into a complex molecular machine. These enzymes and the DNA polymerase are together referred as replisome. Following is a list of enzymes that participate in DNA replication and form the replisome.

Table 5.2 *Role of Different Enzymes in DNA Replication Process*

Enzyme	Function in DNA Replication
DNA Helicase	Helicase separates the two strands of coiled DNA duplex at the replication fork. It is also known as helix destabilizing enzyme.
DNA Polymerase In Prokaryotes	DNA polymerase catalyses the addition of polynucleotides in the new daughter DNA strand in 5′→ 3′ direction. Besides polymerization, it also helps in proof-reading of newly synthesized DNA and its error correction during replication process.
	DNA polymerase III: Helps in synthesis of both leading and lagging DNA strands.
In Eukaryotes	DNA polymerase I: Fills in gaps left by removal of RNA primers.
	DNA polymerase α: Helps in synthesis of lagging strand.
	DNA polymerase β: Helps in synthesis of leading strand.
	DNA polymerase δ: Fills in gaps left by removal of RNA primers.
	DNA telomerase: Completes lagging strand and lengthens telomeric DNA by adding repetitive nucleotide sequences to the ends of Eukaryotic chromosomes.
Single-Strand DNA Binding (SSB) Proteins	These proteins bind themselves to single stranded DNA template and prevent the reannealing of the unwinded DNA double helix, thus stabilize the strand separation, and facilitate the synthesis of the new DNA strand.
Topoisomerase (DNA Gyrase)	Removes the super coiling in the growing fork induced by unwinding.
DNA Ligase	Seal nicks left by filling gaps between Okazaki Fragments of the lagging strand.
Primase	Primase helps in the synthesis of short RNA primer which initiates DNA replication.

Replication Machinery

Replication machineries consist of factors involved in DNA replication and appearing on template ssDNAs. Replication machineries include primosotors are replication enzymes; DNA polymerase, DNA helicases, DNA clamps and DNA Topoisomerases, and Replication Proteins; for example Single-Stranded DNA Binding Proteins (SSB). In the replication machineries these components coordinate. In most of the Bacteria, all the factors involved in DNA replication are located on replication forks and the complexes stay on the forks during DNA replication. These replication machineries are called Replisomes or DNA Replicase Systems.

The replication factories perform disentanglement of sister chromatids. The disentanglement is essential for distributing the chromatids into daughter cells after DNA replication. Because sister chromatids after DNA replication hold each other by Cohesin rings, there is the only chance for the disentanglement in DNA replication. Fixing of replication machineries as replication factories can improve

the success rate of DNA replication. If replication forks move freely in chromosomes, catenation of nuclei is aggravated and impedes mitotic segregation.

Termination

NOTES

Eukaryotes initiate DNA replication at multiple points in the chromosome, so replication forks meet and terminate at many points in the chromosome. Because Eukaryotes have linear chromosomes, DNA replication is unable to reach the very end of the chromosomes. Due to this problem, DNA is lost each replication cycle from the end of the chromosome. Telomeres are regions of repetitive DNA close to the ends and help prevent loss of genes due to this shortening. Shortening of the telomeres is a normal process in somatic cells. This shortens the telomeres of the daughter DNA chromosome. As a result, cells can only divide a certain number of times before the DNA loss prevents further division. Within the germ cell line, which passes DNA to the next generation, telomerase extends the repetitive sequences of the telomere region to prevent degradation. Telomerase can become mistakenly active in somatic cells, sometimes leading to cancer formation. Increased telomerase activity is one of the hallmarks of cancer.

Termination requires that the progress of the DNA replication fork must stop or be blocked. Termination at a specific locus, when it occurs, involves the interaction between two components:

- A termination site sequence in the DNA
- A protein which binds to this sequence to physically stop DNA replication

In various bacterial species, this is named the DNA replication terminus site-binding Protein, or Ter Protein.

Because Bacteria have circular chromosomes, termination of replication occurs when the two replication forks meet each other on the opposite end of the parental chromosome. *E. coli* regulates this process through the use of termination sequences that, when bound by the Tus Protein, enable only one direction of replication fork to pass through. As a result, the replication forks are constrained to always meet within the termination region of the chromosome.

5.4.3 DNA Polymerase

DNA polymerase is responsible for polymerization of deoxyribonucleotides the building blocks of DNA, during the replication of DNA. During this process, DNA polymerase 'Reads' the existing DNA strands to create two new strands that match the existing ones. DNA polymerase adds nucleotides to the 3' end of a DNA strand, one nucleotide at a time.

Every time when a cell divides, DNA polymerase is required to duplicate the DNA, so that a copy of the original DNA molecule can be passed to each daughter cell. In this way, genetic information is passed down from generation to generation. As stated in earlier unit (replication of DNA), during initiation of DNA replication, an enzyme called *Helicase* unwinds the highly coiled double helix DNA molecule by breaking down the Hydrogen bonds between the nucleotide base pairs. Unwinding leads to the opening of the double-stranded helical DNA to two single strands of DNA that can be used as templates for replication.

DNA polymerases have highly conserved structure, as their different catalytic subunits vary among species and are independent of their domain structures. Its shape resembles a right hand with thumb, finger and the palm domains. The major role of finger domain is to bind the nucleoside triphosphates (dATP, dGTP, dCTP, dTTP) with the DNA template bases. The thumb domain helps in translocation and positioning of the DNA whereas the palm domain helps in catalysing the transfer of phosphoryl groups in the phosphoryl transfer reaction. DNA is bound to the palm domain when the enzyme is active. Prokaryotic and Eukaryotic DNA polymerase differs in structure and function of various domains during replication process and is described as follows:

Prokaryotic DNA Polymerase

Prokaryotic DNA polymerase exists in two forms as core polymerase and Poloenzyme. Core polymerase synthesizes DNA from the DNA template but it cannot initiate the synthesis alone but holoenzyme can accurately initiate DNA synthesis. Following are the types of Prokaryotic DNA polymerases:

- **DNA Polymerase I (Pol I)*:*** DNA polymerase I enzyme (Pol I) is a part of prokaryotic family A polymerases, encoded by the polA gene. This enzyme is involved in excision repair with both 3'–5' and 5'–3' exonuclease activity and processing of Okazaki fragments generated during lagging strand synthesis. Pol I is the most abundant polymerase, accounting for more than 95% of polymerase activity in *E. coli*. Pol I adds about 15-20 nucleotides per second, which shows its poor processivity.
- **DNA Polymerase II (Pol II):** DNA polymerase II belongs to family B polymerase and is a polB gene product also known as DNAA. Polymerase II has 3'–5' exonuclease activity and participates in DNA repair. Pol II is also thought to be a backup to Pol III as it can interact with holoenzyme Proteins and assume a high level of processivity. The main role of Pol II is to direct polymerase activity at the replication fork and helped stalled Pol III bypass terminal mismatches.
- **DNA Polymerase III (Pol III):** DNA polymerase III holoenzyme is the primary enzyme involved in DNA replication in *E. coli* and belongs to family C polymerases. It consists of three parts:
  - Pol III Core
  - Beta Sliding Clamp
  - Clamp-Loading Complex

  The core consists of three subunits: α, the polymerase activity hub, ε, exonucleolytic proof-reader and θ, which may act as a stabilizer for. The holoenzyme contains two cores, one for the lagging and leading strand. The beta sliding clamp is present in duplicate, one for each core, to create a clamp that encloses DNA. The third assembly is a seven-subunit (τ2γδδ2 χψ) clamp loader complex (Refer Figure 5.33).

NOTES

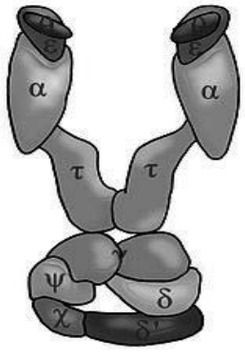


Fig. 5.33 *Structure of DNA Polymerase III in Prokaryotes with Different Subunits*

Table 5.3 *Subunits of the E. coli DNA Polymerase III Holoenzyme and Their Functions*

Subunit	Gene	Properties and Function
Core Subunits α	DNA E	5′→ 3′ polymerase activity required for DNA synthesis.
ε	DNA Q	3′→5′ exonuclease activity required for proofreading.
ϕ	Unassigned	May help to assemble other subunits.
Accessory Sub Units τ	DNA.X[a]	DNA dependent ATPase required for initiation. Promotes dimerization.
γ	DNA.X[a]	Associates with four peptides to form a DNA dependent ATPase required for initiation and also facilitates binding of β-subunit.

Eukaryotic DNA Polymerase

The Eukaryotic DNA Polymerase consists of:

Polymerases β, λ, σ and μ (Beta, Lambda, Sigma, and Mu)

The eukaryotic polymerase pol β, Pol σ, Pol λ, Pol μ, and terminal deoxynucleotidyl transferase belongs to family X polymerases. These polymerases have highly conserved regions that include two helix-hairpin-helix motifs which play an important role in the DNA-polymerase interactions. One motif interacts with downstream DNA and another motif interacts with the primer strand. Pol β, encoded by POLB gene, is required for short-patch base excision repair. Pol λ and Pol μ, encoded by the POLL and POLM genes respectively, are involved in non-homologous end-joining, a mechanism for rejoining DNA double-strand breaks due to hydrogen peroxide and ionizing radiation, respectively. TdT is expressed only in lymphoid tissue, and adds n nucleotides to double-strand breaks formed during V (D)J recombination to promote immunological diversity.

Polymerases α, δ and ε (Alpha, Delta, and Epsilon)

Family B Polymerases contain Polymerases Alpha, Delta, and Epsilon. These polymerases play a major role in nuclear DNA replication. Pol α complex (pol α and DNA primase complex) consists of four subunits: the Catalytic Subunit POLA1,

the regulatory subunit POLA2, the Small Primase Subunit PRIM1 and the Large Primase Subunits PRIM2. Once RNA primase has synthesized the RNA primer during DNA replication, Pol α initiates addition of polynucleotides with RNA primer in lagging strand whereas leading strand is taken by Pol δ.

Gene POLD1 encodes pol δ and creates the catalytic subunits POLD2, POLD3 and POLD4, which interacts with Proliferating Cell Nuclear Antigen (PCNA)- A DNA clamp that allows Pol δ to possess processivity. Whereas pol ε is encoded by the POLE1, the catalytic subunit POLE2 and POLE3 gene. The function of Pol ε is to extend the leading strand during replication.

Polymerases η, ι and κ (eta, iota, and kappa)

Polymerase eta, iota and kappa are Family Y DNA polymerases involved in the DNA repair by translesion synthesis and encoded by genes POLH, POLI and POLK respectively. Members of Family Y have five common motifs to aid in binding the substrate and primer terminus and they all include the typical Right-Hand Thumb, Palm and Finger domains with added domains like Little Finger (LF), Polymerase-Associated Domain (PAD) or Wrist.

Telomerase

Telomerase is a ribonucleoprotein recruited to replicate ends of linear chromosomes. The single-strand 3' overhang of the double-strand chromosome with the sequence 5'-TTAGGG-3' recruits telomerase. Telomerase acts like other DNA polymerases by extending the 3' end.

5.4.4 RNA Polymerase

RNA polymerase is DNA dependent RNA polymerase. During RNA transcription, RNA polymerase opens the double-stranded DNA to utilize one strand of the exposed nucleotides as a template for RNA synthesis. Before RNA polymerase can initiate the DNA unwinding at specific location (a promoter region), a transcription factor and associated transcription mediator complex must be attached to the promoter region. Since RNA polymerase has intrinsic helicase activity, therefore no separate enzyme is needed to unwind the DNA (in contrast to DNA polymerase). It also plays several functions during transcription which includes: Initiation of RNA Transcription, Guiding the Nucleotides into Position, Attachment and Elongation of Nucleotides, Proofreading and Replacement Capabilities, and Termination Recognition Capability.

Prokaryotic RNA Polymerase

In most of the prokaryotes, a single type of RNA polymerase specifies transcription of all the three RNA types. In *E. coli* RNA polymerase core consists of five subunits:

- Two alpha (α) subunits of 36 kDa
- One beta (β) subunit of 150 kDa
- One beta prime subunit (β2) of 155 kDa
- A small omega (ω) subunit

NOTES

A sigma (σ) factor binds to the core and forms the holoenzyme. After transcription begins, this sigma (σ) factor can unbind and let the core enzyme proceed with its function. The whole core RNA polymerase complex forms a clamp-jaw structure with an internal full-length channel. Eukaryotic and archaeal RNA polymerases have a similar core structure and work in a similar manner, although they have many extra subunits.

All RNA polymerases contain metal cofactors, specifically zinc and magnesium cations which aid in the transcription process (Refer Figure 5.34).

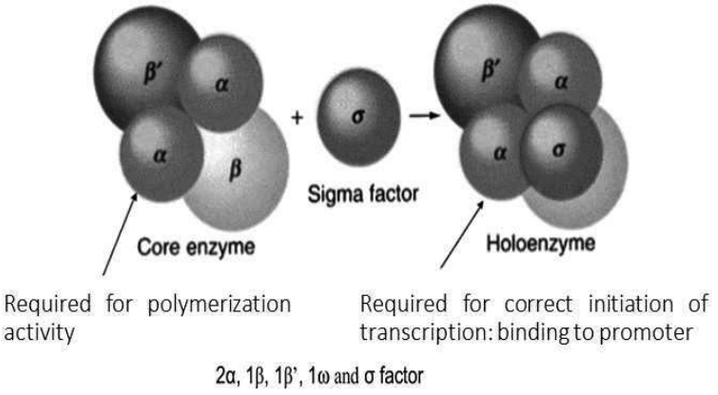


Fig. 5.34 *Structure of E. coli RNA Polymerase with Different Subunits*

Eukaryotic RNA Polymerase

In eukaryotes, RNA polymerase can build long chains of about 2.4 million nucleotides. Eukaryotes have multiple types of nuclear RNA polymerases, each responsible for synthesis of a distinct type of RNA. All are structurally and mechanistically related to each other and to prokaryotic RNA polymerase:

- **RNA Polymerase I:** It synthesizes a pre-rRNA 45S (35S in Yeast), which matures into 28S, 18S and 5.8S rRNAs which will form the major RNA sections of the ribosome.
- **RNA Polymerase II:** It synthesizes precursors of mRNAs and most snRNA and microRNAs. This is the most studied type, and, due to the high level of control required over transcription, a range of transcription factors are required for its binding to promoters.
- **RNA Polymerase III:** It synthesizes tRNAs, rRNA (5S) and other small RNAs found in the nucleus and cytosol.
- **RNA Polymerase IV:** It synthesizes siRNA in plants.
- **RNA Polymerase V:** It synthesizes RNAs involved in siRNA-directed heterochromatin formation in plants.

Table 5.4 Different Types of RNA Polymerases, their Products and Location

Form	Product	Location
I	rRNA	Nucleolus
II	mRNA, snRNA	Nucleoplasm
III	tRNA, 5sRNA	Nucleoplasm
IV	siRNA	Nucleus and Cytoplasm
V	siRNA-Directed Heterochromatin Formation	Nucleus and Cytoplasm

NOTES

5.4.5 Mitotic/Spindle Apparatus

The spindle apparatus is the cytoskeletal structure formed during cell division to separate sister chromatids between daughter cells and found in eukaryotic cell. It is also referred as mitotic spindle during mitosis, a process that produces genetically identical daughter cells, or the meiotic spindle during meiosis, a process that produces gametes with half the number of chromosomes of the parent cell (Refer Figure 5.35).

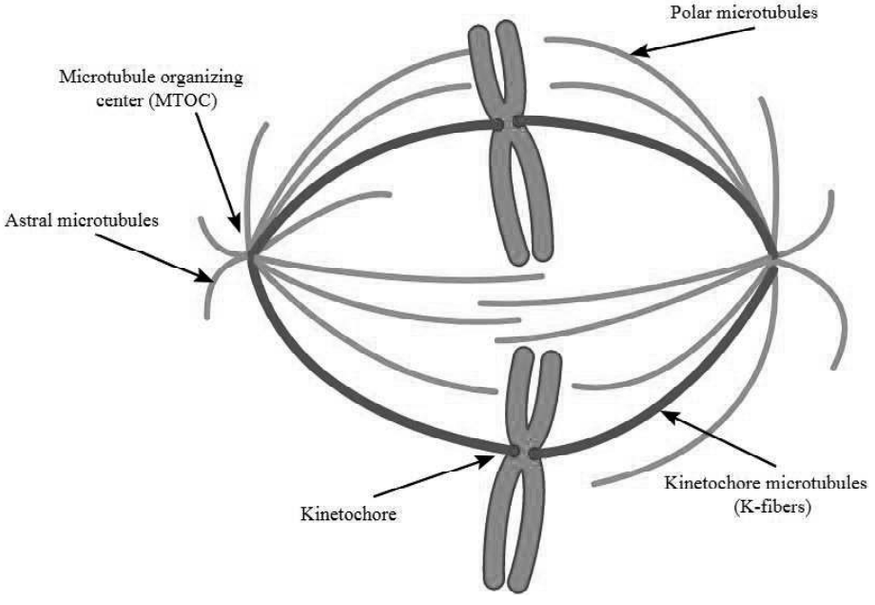


Fig. 5.35 Organization and Different Types of Mitotic Spindle in Animal Cells

Types of Microtubules

Besides chromosomes, the spindle apparatus is composed of hundreds of proteins. The most abundant components of the machinery comprises microtubules. Spindle apparatus includes the microtubules, associated proteins (kinesin and dynein), condensed chromosomes and any centrosomes or aster.

During mitotic metaphase stage of cell cycle, the spindle apparatus is comprised of kinetochore fibres, astral microtubules and interpolar microtubules. The kinetochore microtubules are connected to the kinetochore of the chromosomes, the astral microtubules hold the pole ends in position while the

NOTES

polar microtubules form the structure of the spindle apparatus. The spindle apparatus is somewhat ellipsoid with wide middle portion and tapers at the ends in cross sectional view. In the middle portion (spindle midzone) antiparallel microtubules are bundled by kinesins. At the pointed ends (spindle poles) microtubules are nucleated by the centrosomes in most animal cells. Acentrosomal or anastral spindles lack centrosomes or asters at the spindle poles.

- **Midzone and Midbody**: The kinetochore of sister chromatid forms bivalent attachments with the kinetochore fibres (spindle microtubules) and are positioned with respect to the division plane of the cell during anaphase. Bivalent attachment of the chromosome to the spindle is achieved when the plus ends of the microtubules emerge from each pole interacts with the kinetochores of each sister pair and then becomes embedded. At the beginning of anaphase, the kinetochore fibres shorten, delivering sister chromatids to the poles and astral microtubules elongate. The region between the two poles is called the Spindle Midzone and the microtubules in this region are called Midzone Microtubules. The term central spindle refers to the structure at the centre of the midzone, where the plus ends of the microtubules interdigitate. The microtubules of the central spindle eventually lose their interaction with the spindle poles. As the formation of the cleavage furrow progresses, the central spindle becomes compacted dense structure known as the Midbody.
- **Astral Microtubules**: Astral microtubules play a large role in determining the plane of division due to the interaction of astral microtubules with cortical actin and non-actin cortical factors. They are part of the spindle-dependent mechanism that induces formation of the cleavage apparatus as a cortical ring encircling the mitotic spindle. As a result, cytokinesis always cleaves the cell between the separated chromosomes independent of spindle orientation within the cell body. The interactions of the astral microtubule with cortical actin generate forces that align the spindle fibres parallel to the polarity axis of the cell so that the cleavage apparatus bisects the both separated chromosomes and polarized material. Therefore, astral microtubules ensure accurate chromosome segregation and determine the amount of symmetry by establishing and coordinating positions of the cleavage plane and spindle.

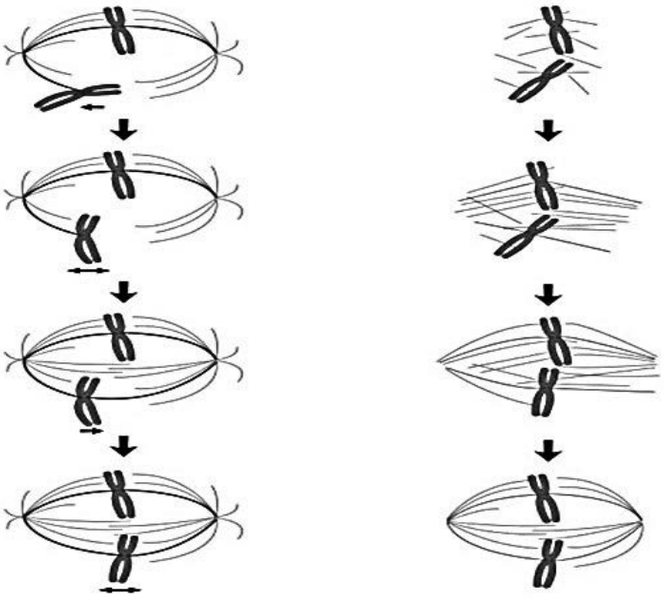
Organization of Spindle Apparatus

Bioriented chromosomes are aligned along the equator of the cell with spindle microtubules oriented roughly perpendicular to the chromosomes, their plus-ends (+) inserted in the kinetochores while their minus-ends (-) anchored at the cell poles. The precise orientation of this complex is required to ensure accurate chromosome segregation and to specify the cell division plane. Two models: search-and-capture model and self-assembly model differently explains the organization of spindle apparatus

- **Search and Capture Model:** In this model, microtubules are nucleated at microtubule organizing centres and undergo rapid growth and catastrophe to search the cytoplasm for kinetochores. Once they bind a kinetochore,

they are stabilized and their dynamics are reduced. The newly mono-oriented chromosome oscillates in space near the pole to which it is attached until a microtubule from the opposite pole binds the sister kinetochore. This second attachment further stabilizes kinetochore attachment to the mitotic spindle. Gradually, the bi-oriented chromosome is pulled towards the centre of the cell until microtubule tension is balanced on both sides of the centromere; the congressed chromosome then oscillates at the metaphase plate until anaphase onset releases cohesion of the sister chromatids. In this model, microtubule organizing centres are localized to the cell poles, their separation driven by microtubule polymerization and 'Sliding' of antiparallel spindle microtubules with respect to one another at the spindle midzone mediated by bipolar, plus-end-directed kinesins. Such sliding forces may account not only for spindle pole separation early in mitosis, but also spindle elongation during late anaphase.

- **Self-Assembly Model:** According to self-assembly model microtubules are nucleated acentrosomally near chromosomes and spontaneously assemble into anti-parallel bundles and adopt a spindle-like structure. The microtubules undergo acentrosomal nucleation among the condensed chromosomes. Further, the shape and size of the mitotic spindle are a function of the biophysical properties of the Cross-Linking Motor Proteins (Refer Figure 5.36).



(A) *(B)*

Fig. 5.36 *Models of Spindle Apparatus Organization A) Chromatin-Mediated Search and Capture Model, B) Chromatin-Mediated Self-Assembly Model*

Microtubule Associated Proteins and Their Role in Spindle Structure

The lengthening and shortening of spindle microtubules, through a process known as dynamic instability determines to a large extent the shape of the mitotic spindle and promotes the proper alignment of chromosomes at the spindle midzone.

NOTES

Microtubule-Associated Proteins (MAPs) associate with microtubules at the midzone and the spindle poles to regulate their dynamics. γ-tubulin is a specialized tubulin variant that assembles into a ring complex called γ-TuRC which nucleates polymerization of α/β tubulin heterodimers into microtubules. Recruitment of γ-TuRC to the pericentrosomal region stabilizes microtubule minus-ends and anchors them near the microtubule-organizing center. The Microtubule-Associated Protein Augmin acts in conjunction with γ-TURC to nucleate new microtubules from existing microtubules. The growing ends of microtubules are protected against catastrophe by the action of Plus-End Microtubule Tracking Proteins (+TIPs) to promote their association with kinetochores at the midzone.

Microtubule Associated Proteins are Majorly of Two Types: Type I MAP's and type II MAP's Type I includes MAP1 proteins and type II includes MAP2, MAP4 and tau proteins. MAPs bind directly to microtubules to stabilize or destabilize them and link them to various cellular components including other microtubules.

Type I: MAP 1

MAP1 family includes two major members: MAP1A and MAP1B. These Proteins bind to microtubules through charge interactions. The Proteins of this family are involved in microtubule assembly, which is an essential step in neurogenesis. The carboxyl terminus (-C) of these Proteins bind the microtubules and the N terminal bind other parts of the cytoskeleton or the plasma membrane to control spacing of the microtubule within the cell. The product of this gene is a precursor polypeptide that presumably undergoes proteolytic processing to generate the final MAP1A heavy chain and LC2 light chain. These Proteins are found in the axons and dendrites of nerve cells.

Type II: MAP2, MAP4 and Tau Proteins

Type II MAPs are found particularly in nerve cells in mammals. MAP2 (found mostly in dendrites) and Tau Proteins (found in in the axon) participate in determining the structure of different parts of nerve cells. Both the Proteins have microtubule-binding domain with conserved C-terminal and outwardly projecting variable N-terminal domains which interacts with other Proteins. MAP2 and Tau Proteins stabilize microtubules by binding to the outer surface of the microtubule protofilaments and favour the addition of new subunits thus accelerating the microtubule growth. In addition, MAP2 protein binds in a cooperative manner, with many MAP2 Proteins binding a single microtubule to promote stabilization and Tau has the additional function of facilitating bundling of microtubules within the nerve cell.

The function of Tau protein is linked to Alzheimer's disease. In the nervous tissue of Alzheimer's patients, Tau Protein forms abnormal aggregates. The aggregated Tau is often severely modified, through hyperphosphorylation which causes them to detach from microtubules. Thus, the hyperphosphorylation of Tau leads to massive detachment, which in turn greatly reduces the stability of microtubules in nerve cells. This increase in microtubule instability may be one of the main causes of the symptoms of Alzheimer's disease.

NOTES

MAP4 is found in almost all types of cells. MAP4 is responsible for stabilization of microtubules and has also been linked to the process of cell division like MAP2 and Tau Proteins.

Besides the above mentioned microtubule associated Protein category, some other MAPs have also been identified, which bind to the length of the microtubules, for example MAP6, MAP7 and EB1, 2, 3, CLIP170, CLIP115, CLASP1 and CLASP2 (plus end tracking Proteins bind to the tip of growing microtubules).

There are many other proteins which affect microtubule behaviour, such as catastrophin which Destabilizes Microtubules and Katanin and serves a number of motor Proteins that transport vesicles.

Functions of Spindle Apparatus

Spindle apparatus perform its three major functions: Chromosome alignment, chromosome segregation and bipolarity.

- **Chromosome Alignment:** During the onset of metaphase, the spindle apparatus is responsible for the alignment of chromosomes. The chromosomes attached to the mitotic spindle eventually align halfway between the two spindle poles and form metaphasic plate. The continuous lengthening and shrinkage of the microtubules and the actions of the microtubule associated Proteins are involved in keeping the chromosomes at the plate (Refer Figure 5.37).

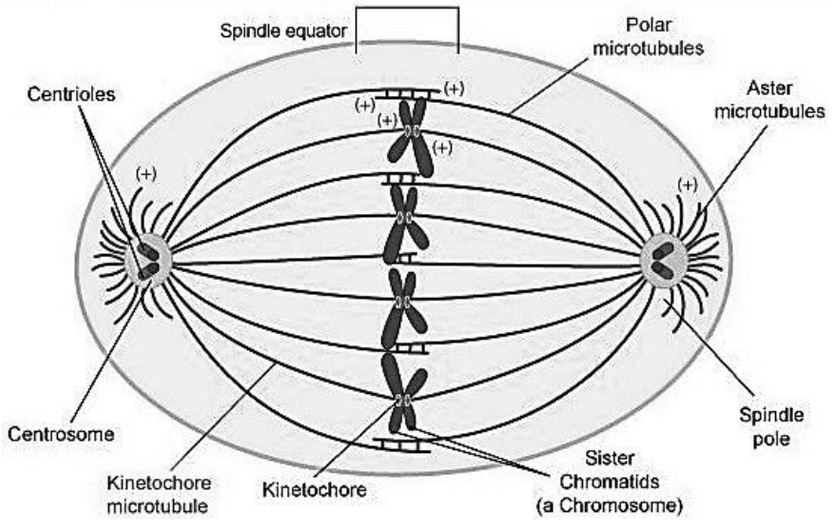


Fig. 5.37 *Schematic Representation of Chromosomes Alignment in the Mitotic Spindle during Metaphase of Cell Cycle*

- **Chromosome Segregation:** During anaphase the spindle poles segregate the separated sister chromatids. All the chromatids move apart at the same speed due to two processes involving the mitotic spindle:
  - o Microtubule associated proteins operate at the kinetochore to shorten the kinetochore microtubules through depolymerisation. This causes the chromatids to move poleward.

NOTES

- The spindle poles themselves move apart therefore contributing to chromosome segregation. Kinesin and dynein families operate on different types of spindle fibres to provide the driving force. For example, the kinesin-5 Motor Proteins require contact with another microtubule for it to cross-link with and push against for it to activate and become an ATP hydrolysing directional motor that drives spindle elongation. It acts on the interpolar microtubules and causes them to slide from opposite poles past one another at the spindle equator, therefore pushing the spindle poles apart.

Spindle migration and chromosome movements depend on the tiny mechanical forces that are generated by many different numbers of Motor Proteins associating with microtubule assembly and disassembly (Refer Figure 5.38).

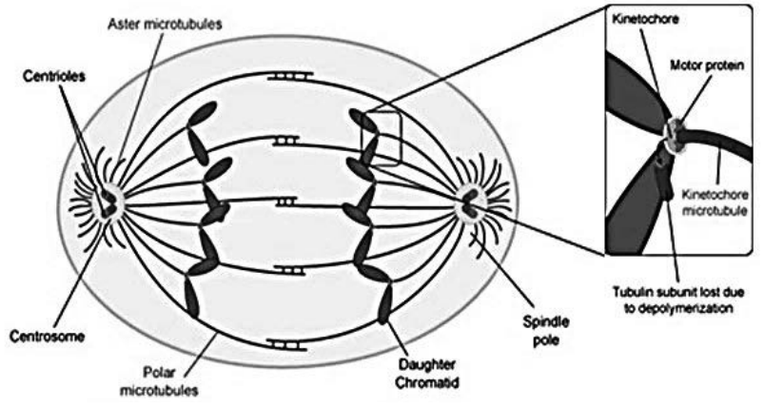


Fig. 5.38 *Chromosome Segregation during Anaphase of Cell Cycle*

- **Bipolarity:** In the spindle, kinetochore microtubules have their plus ends embedded in the kinetochores of the sister chromatids and their minus ends at the spindle pole ends. Kinesins are important to maintain spindle bipolarity. The simultaneous Kinesin Inhibitor (KinI) induced disassembly at both the plus and minus ends may result in the poleward driving forces. Bipolarity is critically important for correct segregation of the chromosomes. kinesin Eg5 (a plus ended kinesins 5 molecules) is responsible for allowing spindle function. Motor domains favour the second bound microtubule to be in an anti-parallel orientation and this may contribute to the formation of the bipolar spindle. Bipolarity of the spindles ensure the chromosomes separate with the highest possible fidelity.

5.4.6 Transcription

Transcription is a process of synthesis of RNA or transcript. It is based on the complementary base pairing from DNA templates catalyzed by DNA dependent RNA polymerases. The chain is synthesised in 5' to 3' direction unidirectionally. Each transcribed segment from DNA is known as transcription unit. Eukaryotes carry monocistronic transcription in which the coding sequence is present for only one polypeptide. In other words, it carries the information for just one gene.

NOTES

Whereas, prokaryotes carry polycistronic transcription unit in which coding sequence for more than one polypeptide is present and so codes for more than one gene products. The initial product of transcription is called primary transcript.

- **Start Point**: Start site is the first base pair from where transcription starts. The process of synthesising RNA is done by RNA polymerase that moves from start point along with template up to terminator sequence.
- **Upstream:** The sequence before the start point in the 5' end or minus direction is called upstream. It is the non-template nucleotide.
- **Downstream:** The sequence in the 3' end or plus direction present after the start point. This is the part of transcription unit.

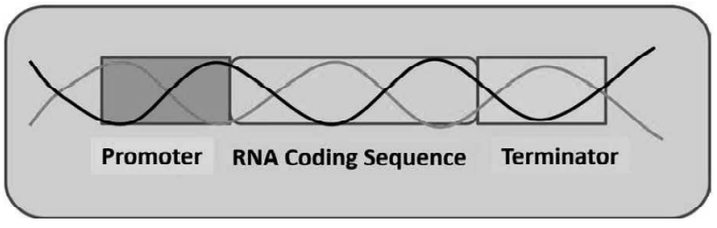


Fig. 5.39 *Structure of RNA Coding Gene*

During transcription, the complementary bases are added only to one strand of the DNA. That DNA strand which acts like a template is called antisense strand. The other stand which is not copies but is identical to the transcribed sequence is termed s coding or sense strand.

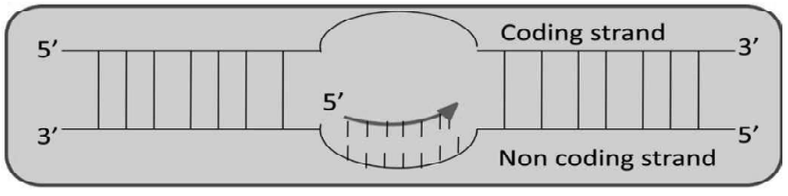


Fig. 5.40 *Coding and Non-Coding Strands*

Although the basic steps of transcription are same between prokaryotic and eukaryotic organism, but there are some in the process in both like:

Table 5.5 *Difference between Prokaryotic and Eukaryotic Termination*

Prokaryotic Transcription	Eukaryotic Transcription
Transcription is of polycistronic type.	Transcription is of monocistronic type.
Occurs in cytoplasm	Occurs in nucleus
Coupled transcription –translation process occurs.	Coupled transcription –translation process not occurs.
RNA Polymerase are made up by 5 subunits	RNA Polymerase are made by 10-15 subunits
No need of any transcription factor for initiation	Its require transcription factor for initiation
Single type of RNA Polymerase required for synthesis of all type of RNA	Three different type of RNA Polymerase required for synthesis of all type of RNA

NOTES

Transcription in Prokaryotes

One type of RNA polymerase is present that is responsible for synthesis of all type of RNA namely, rRNA, mRNA, tRNA. This DNA dependent RNA polymerase or RNA polymerase was discovered by Samuel B. Weiss and Jerard Hurwitz I 1960. Eubacterial RNA pol is a multisubunit protein composed of five subunits ααββ'ωí. This holoenzyme can be separated into two components: The core enzyme (ααββ'ω) and the sigma factor (í polypeptide).

Holoenzyme = Core Enzyme + Sigma Factor

The complete holoenzyme has a molecular mass of approximately 466 kDa. All the subunits are responsible for their particular roles. Like, α subunit is responsible for the assembly of core enzyme, promoter recognition and also in interaction of RNA polymerase with regulatory factors. β β'- performs all enzymatic and catalytic function together and β is also involved in chai elongation. í is specifically plays role in promoter sequence recognition. ω facilitates assembly of RNA polymerase and stables assembled RNA polymerase. ααββ' forms core enzyme.

The promoter sequence consists of 40 bp region located near to 5' end side of transcription start site. It is a cis acting, position dependent DNA sequence with is responsible for the initiating the transcription efficiently. It contains two 6 bp consensus sequences elements- Pribnow box or TATA box and -35 region. Pribnow box is a consensus sequence of TATAAT located 10 bp upstream of start point and -35 region has consensus sequence of TTGACA.

Fig. 5.41 *Structure of Prokaryotic Promoter*

Transcription occurs in four stages:

- **Binding of RNA Polymerases to Template DNA and Chain Initiation:** To approach to the single stranded DNA template, DNA duplex should be opened. With the í factor, RNA polymerase recognizes the promoter sequence at the template DNA and at this site the RNA polymerase holoenzyme binds to form a closed binary complex. Initially on binding to the DNA sequence, the RNA polymerase covers the DNA sequence from -55 to +20. Then melting of short region of DNA (upto-10bp) occurs where the enzyme has been bound, this local unwinding leads to the formation of transcription bubble. -10 region of template is essential for recognition, in the closed complex, the promoter region is double stranded and is single stranded in open complex. After the bubble has been formed, the next step is the incorporation of nucleotides and formation of phosphodiester bond between them. As the chain reaches nearly up to 10 bases, sigma factor is released leaving core enzyme for further elongation.

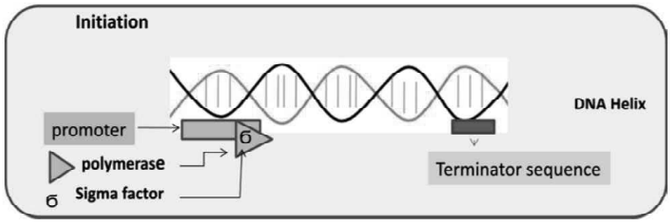


Fig. 5.42 *Initiation of Transcription*

RNA polymerase contains two binding sites for nucleotides: a. initiation sites: Within the open promoter site, the RNA polymerase binds to the first nucleotide at +1 position which is usually a purine-A or G. So, usually the first nucleotide is ATP or GTP. b. Elongation site: To the +2 site, the second incoming nucleotide binds. As these two nucleotides are joint together, the first base is released from the initiation site, and the initiation is complete

- **Chain Elongation:** Chain is elongated in 5'-3' direction with the addition of nucleotides at 3' end complementary to the template DNA (3'-5'). As the RNA polymerase moves along with DNA and RNA chain grows gradually. The 3' OH group of last nucleotide bond with the incoming 5' Ò phosphate nucleotide; α and β phosphate groups are removed and only Ò phosphate is used in the formation of phosphodiester bond, hydroxyl group is also removed from 3' carbon nucleotide presents at end of chain. In the similar manner, the chain is elongated and then the translocation of RNA chain occurs. Transcription rate of bacteria is nearly 40 to 50 nucleotide per second at 37. As the transcription proceeds, transcription bubble moves continuously by disrupting the DNA structure, the duplex reforms again and transcribed RNA hangs as free polynucleotide chain.

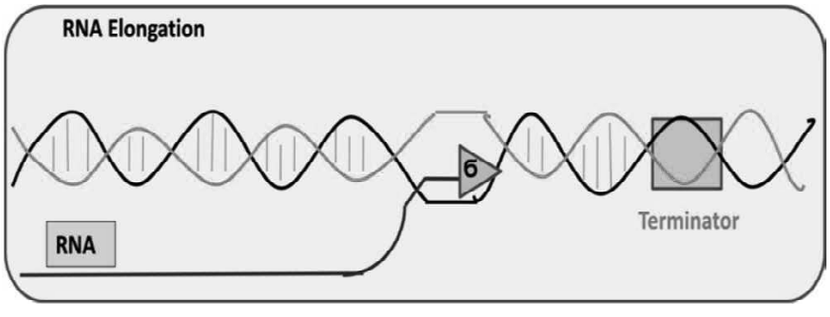


Fig. 5.43 *Elongation of Transcription*

- **Chain Termination:** Upon completion of the transcription, RNA polymerase stops adding nucleotide at RNA chain and it releases the completed RNA chain from the terminator sequence. All the hydrogen bonds are broken between the RNA-DNA hybrid and RNA chain is separated DNA again reform duplex. Transcription is halted at chain termination site. Generaly in bacteria there are two mechanisms for termination:

NOTES

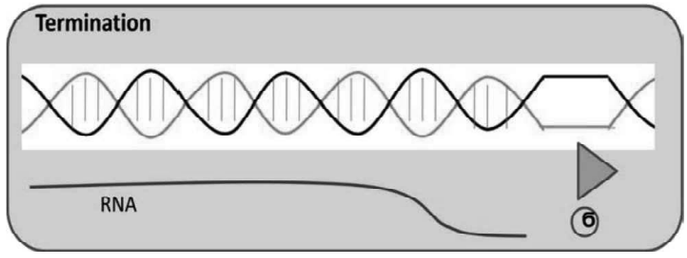


Fig. 5.44 Termination of Transcription

- **Intrinsic Termination:** This mechanism depends on the recognition of the sequence at which no bases can be added further. As the transcription goes on, the RNA- DNA hybrid requires forces for holding the elongation complex together.

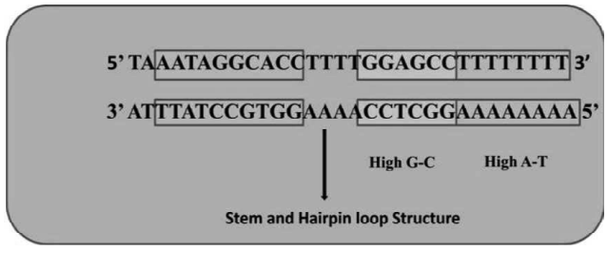


Fig. 5.45 (a) Stem Hairpin Loop Structure

- On the transcript two G-C rich GCCCGC inverted repeats followed by 7U residue sequence are present. This led to the formation of hairpin structure due to the complementary base pairing between the inverted repeat sequence. This stem loop structure causes the hybrid to detach as -U bonds get break down and leads termination and RNA molecules get separated.

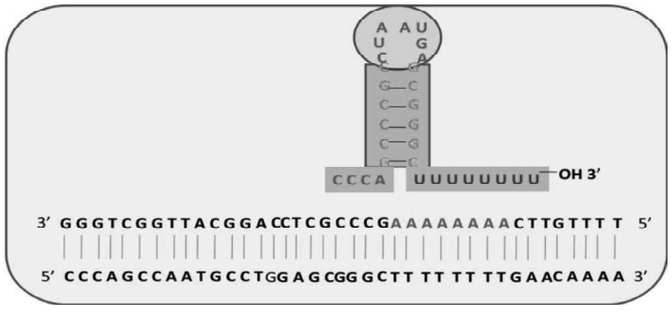


Fig. 5.45 (b) Chain Termination

- **Rho Dependent Termination:** Rho dependent termination requires Rho protein for disturbing RNA –DNA hybrids and terminating transcription. Rho is a hexameric ATP dependent helicase with subunit containing RNA binding and ATP hydrolysis domain. It binds to the sequence upstream of termination site called rut site. These sites are rich in C residues, so when RNA polymerase reaches this site, the movement of polymerase is slowed down and it catches the rho factor.

NOTES

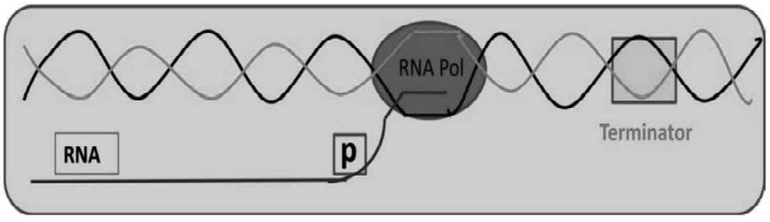


Fig. 5.46 (a) *Rho Factor is following RNA Polymerase*

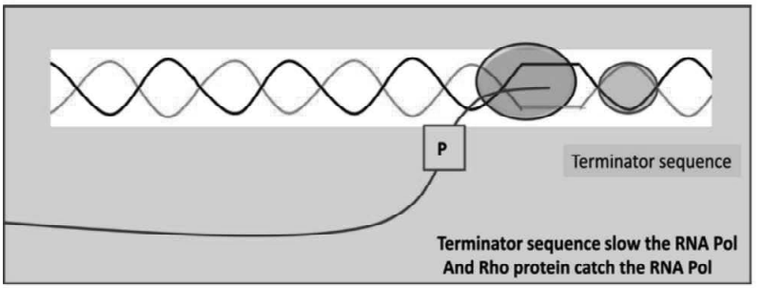


Fig. 5.46 (b) *Terminator Sequence Slow the RNA Pol and Rho Protein Catch the RNA Pol*

ATP hydrolysis controls the helicase activity of the Rho protein by which it follows the RNA polymerase enzyme. When RNA polymerase reaches terminator site, the rho protein freeze the structure of polymerase and collapse it. This causes termination and pre synthesises RNA chain get released.

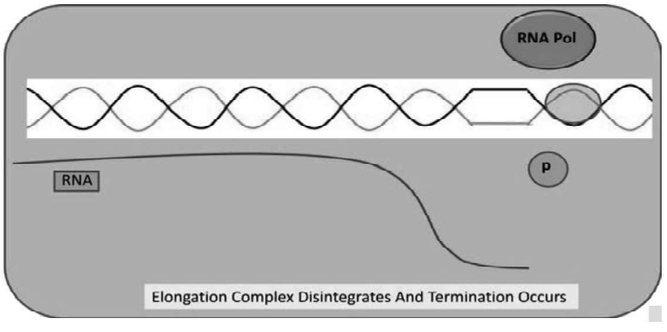


Fig. 5.46 (c) *Elongation Disrupt the RNA–DNA Hybrid and Leads Termination*

Eukaryotic Transcription

Transcription process is different in prokaryotes and eukaryotes. In eukaryotes, transcription binding factors, enhancer along with RNA polymerase are also required. Transcription factors are the proteins responsible for transcription. These factors bind sequentially to the DNA template and make an initiation complex with DNA polymerase. These facilitates the recognition of the promoter sequence by RNA polymerase. Mitochondria and chloroplast have RNA polymerase similar to bacteria. Eukaryotic RNA polymerases are multi subunit proteins containing distinct RNA polymerases for different RNA's.

NOTES

Table 5.6 *Different RNA Polymerases of Eukaryotic System*

Type of polymerase	Synthesis of RNA	Occurrence
RNA Pol I	synthesis of r RNA	Nucleoli
RNA pol II	synthesis of m-RNA	Nucleoplasm
RNA pol III	Synthesis of t-RNA	Nucleoplasm

Promoters are the specific sequence which are recognised by the transcription factor for initiating transcription. Promoter of RNA POL l and POL ll are mostly located at upstream site, but promoter of rRNA lll are located on downstream site of the start point. Eukaryotic promoter includes core promoter element at which RNA polymerase get attached and form initiation complex and also for efficient transcription it requires an upstream promoter element, which are basically G+C rich region and at which transcription factors are bind

Initiation

Eukaryotic mRNA transcription requires initiation complex which consist General Transcription Factors (GTFs) and mediator. Various transcription factor and their functions are as follows:

Table 5.7 *Eukaryotic Transcriptio Factors*

General Transcription Factors (GTFs)	Function
TFIID: I) TBP (Tata binding protein) ii) TAFs (TBP Associated Factors)	Recognize core promoter (TATA box) Recognize core promoter (non-TATA box)
TFIIA	Stabilizes TBP and TAFs binding
TFIIB	It helps in RNA polymerase II and TFIIF recruitment and in start site selection
TFIIF	Help RNA Pol in promoter binding
TFIIE	Help in TFIIH recruitment, modulation of TFIIH, Helicase, ATPase, Kinase activities
TFIIH	Help in promoter melting with helicase activities. Prompter clearance by phosphorylation

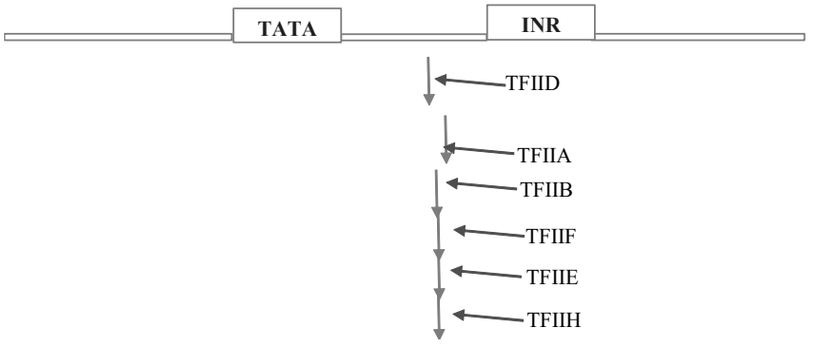


Fig. 5.47 *Sequential Addition of Transcription Factors*

These transcription factors bind sequentially bind to TATA box DNA and form a pre initiation complex. TFIIH binds in the last and phosphorylates Pol II to initiate transcription in the presence of ATP. Mediator is a protein complex that mediates between co activator and enhancer, and between promoter and RNA polymerase. For transcription activation mediator is essential.

Termination

The termination stage involves:

Table 5.8 *Transcription Termination*

For Polymerase l gene	rho dependent termination
For polymerase lll gene	rho independent termination
For polymerase ll gene	It more complex, pol II termination generally coupled with the RNA processing event, in which 3'end of transcript undergoes cleavage and polyadenylation.

RNA pol II transcription genes may continue up to hundreds or even thousands of nucleotides beyond the end of a coding sequence. Then the cleavage of RNA strand occurs by a complex which appears to associate with the polymerase. Cleavage of RNA is couple with termination process and it occurs at same consensus sequences. The polyadenylation of mature pol ll m RNA occurs at the 32 , which result in a poly (A) tail; this process is followed by cleavage and termination. Both processes polyadenylation and termination occur at same consensus sequence, and both these processes are interdependent.

Poly A dependent termination is basically coupled with the RNA maturation process in which 3' end of nascent RNA undergoes polyadenylation and cleavage, and these 3'end processing reaction are carried out in two steps: Transcription of poly A followed by cleavage of nascent transcript, and then the upstream product is polyadenylated and downstream product is degraded.

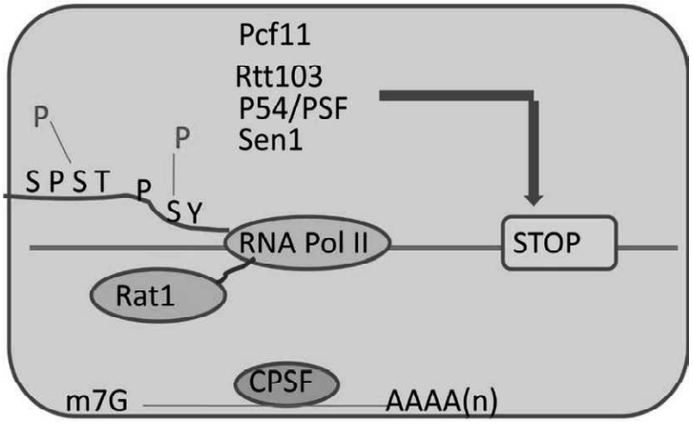


Fig. 5.48 *Termination of RNA Pol II Transcript*

Basically 3' end processing starts when the cis acting element in the Poly A site of nascent RNA transcript is recognised by the binding factors. When these factors bind at 3'end, it forms a very complex structure which result in high shear

forces consequently processing slow down which causes disruption of polymerase ll and DNA –RNA hybrid complex; and ultimately termination occurs.

Post-Transcriptional Modification

NOTES

There is no post transcriptional modification in prokaryotes as transcription and translation both processes are carried out instantaneously in them. The RNA copy of a gene called m-RNA, is ready for translation into protein just after transcription. Even, translation starts before transcription is finished. In eukaryotes, the primary RNA transcript of a gene needs to be processed further before going for translation into protein. This step is titled as "RNA processing". Also, after the processing, it needs to be transported out of the nucleus into the cytoplasm through the nuclear pores. Following modification happens in Pre-m-RNA. A eukaryotic gene comprises of several non-coding region (introns) and coding region (exons). So, upon transcription, the primary transcript has both introns and exons. Therefore, by the process of splicing, introns are removed from primary transcript and exons are joined together in a continuous sequence forming mature m-RNA transcript ready to be transported out of the nucleus and translated. Also, eukaryotic m-RNA undergoes 5' capping and polyadenylation at 3' end. All three processes of splicing, capping and adenylation occurs inside nucleus and then the mature m-RNA is transported to cytoplasm for further translation.

5.4.7 Translation

mRNA templates obtained from the transcription of DNA, are converted into the proteins by the process known as translation. The translation of mRNA occurs in 5′ to 3′ directions converting the nucleotide sequence into the amino acid sequence based on the universal genetic code. The process occurs on ribosomes and tRNAs act as adaptors among the mRNA template and the amino acids that polymerases to form polypeptide chain of a protein. The mechanism of translation is same in all the cells.

Post-Translational Modifications

Post-translational modifications are the critical chemical modifications which plays an important role in functional proteomics as they regulate position, activity and interaction of proteins with other cellular molecules like proteins, lipids, nucleic acids, cofactors, etc. Post-translational modifications take place at different amino acids side chains or at peptide linkages by different enzymes like phosphatases, kinases, transferases and ligases whose main function is to add or remove different functional groups or sugars to or from amino acids side chain. It can also involve protease which cleave peptide bonds or remove specific sequences or regulatory subunits from a large polypeptide.

Post-translational modification is like a step of a protein life cycle as some proteins are modified gradually after translation because it requires proper protein folding of nascent protein for stability. Some other modifications occur after folding and localization to stimulate or deactivate catalytic activity or to stimulate the biological activity of the protein. Proteins are also linked covalently with the tags

NOTES

that target a protein. It can also be reversible or irreversible which directly depend upon the nature of the protein modification. Example, kinase phosphorylate side chains of precise amino acid of any protein, which is a common method of protein activation or inactivation. Phosphatases hydrolyse the phosphate group to remove it from the side chain of specific amino acids of a protein and reverse the biological activity against the kinases.

5.4.8 Genetic Code

The genetic code is the set of rules used by living cells to translate information encoded within genetic material into proteins. In this, the four bases of DNA—the A, C, G, and Ts—are strung together in such a way that the cellular machinery, the ribosome, can read them and transform them into a protein. Each three nucleotides in a row counts as a triplet in the genetic code and codes for a single amino acid.

mRNA is a random sequence comprised of four types of nucleotides bases, Uracil (U), Adenine (A), Cytosine (C), and Guanine (G). Three nucleotides together form a codon which code for one amino acid. The reason behind the codon being a triplet of nucleotide is that, as the whole genome is composed of the sequence comprising of four nucleotides, so if the codon would be of two nucleotides, it would give the possible combination $4^2=16$ for amino acid residues. But, as the total amino acid in the body are 20, therefore the codon must code for all amino acid or must have the combination equal or greater than twenty. So, with the triplet codon, the number of combinations produced are $4^3=64$, which is much sufficient for coding 20 amino acid. So, this mathematics led Francis Crick, Sydney Brenner, and their colleagues to give triplet nature of codon. Coding sequence is composed of start and stop codon to initiate and terminate the translation process respectively. Start codon is AUG which code for methionine and stop codons are UAA, UAG, and UGA. In some cases, starting codons can be GUG or UUG.

Each codon can code for one amino acid which suggests that code is unambiguous. But, one amino acid can be coded by two or more codon, so, the code is degenerate. For example, a codon UUU can code for phenylalanie only, but phenylalanine can be coded by UUU and UUC. Tryptophan and methionine are the only amino acids coded by one codon. A codon carries no commas or gaps. The sequence also not overlapping. So, the sequence for example, AUGCUGGGUGAUUUUGUA would be having codons, AUG, CUG, GGU and so on and not AUG, UGC, GCU and so on.

Genetic code is common for all life forms and so it is universal. But there are some exceptions to this rule. For example, as UGA is a stop codon but it codes for tryptophan in mycoplasma, spiroplasma, and mitochondria of several species. Similarly, CUG codes for Leucine in general but in yeast mitochondria its codes for threonine. The different codon triplets coding respective amino acids is represented in figure 5.49.

NOTES

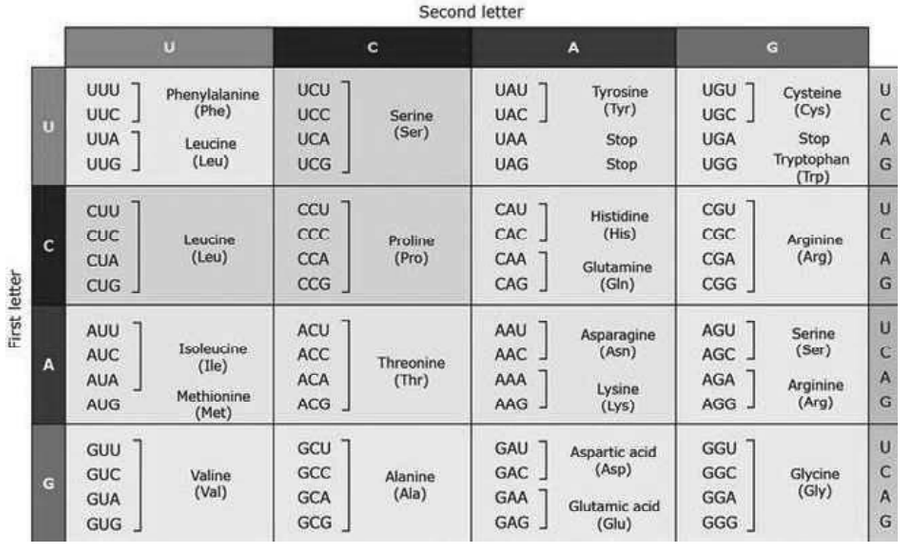


Fig. 5.49 *Coding Dictionary*

Messenger RNA (mRNA)

mRNA carries the genetic information from DNA in the form of three nucleotides or a triplet representing one amino acid to a polypeptide chain. mRNA has a 5' end, 5' UTR, ribosomal binding site, coding sequence, 3' UTR. In eukaryotes there are additional structures as 5' Guanine cap and poly (A) tail. Messenger RNA (mRNA) has 3 reading frames out of which only one codes for desired protein. If in the sequence of bases there is no stop codon to interrupt the translation then that synthesis entire polypeptide chain and is that is called open reading frames (ORF).

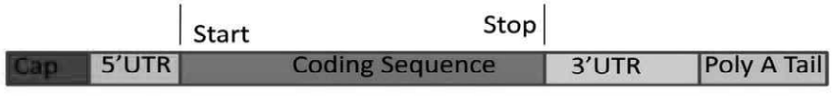


Fig. 5.50 *Structure of m-RNA*

Transfer RNA (t-RNA)

Transfer RNA (tRNA) is clover leaf structured molecule in two dimension and L-shaped structure in 3 dimensions with 73 to 94 ribo-nucleotides in length.

A tRNA molecule has a 5' phosphate terminal, an acceptor arm that ends in CCA terminal at 3', D loop which frequently comprises of dihydrouridine, anticodon loop, and T arm which has TΨC where Ψ is pseudouridine. CCA sequence is the site of attachment of amino acid important and plays important role in recognition of tRNA. Each t-RNA is specific to amino acid that it carries it in CCA arm. There are more then one tRNA for one amino acid as there are 30-45 different tRNA in prokaryotes and 50 types in eukaryotes. For example, glycine has two tRNA: tRNAGly1 and tRNAGly2.

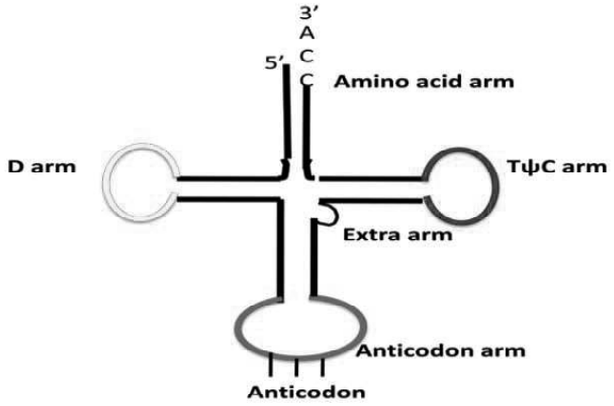


Fig. 5.51 *Structure of t-RNA*

Ribosomes

Ribosomes are ribonucleoprotein particles that comprises of r-RNA and proteins. Each ribosome has two subunits. Mitochondria, chloroplast of eukaryotes and prokaryotes has 70S ribosome which is composed of 50s and 30s subunits. In *e. coli*, 30s subunit contain 16s rRNA (1541 nucleotides) and 21 r-proteins and 50s subunit cover 23s rRNA (2904 ntds), 5s rRNA (120 ntds) and 31 proteins. In eukaryotes there is 80S ribosome which contain 60s and 40s ribosomal subunit. 60s subunit consists of 28s rRNA (4718 nucleotides), the small 5s rRNA (120 nucleotides), 5.8s rRNA (160 nucleotides) and about 50 proteins. The 40s subunit has 18s rRNA (1874 nucleotides) and 33 r-proteins. Here, S= Svedberg's unit of sedimentation coefficient. The 70s ribosome contain three tRNA binding sites- a. P-site (or peptidyl-tRNA binding site), b. A-site (aminoacyl-tRNA-binding site), and c. E-site (deacylated tRNA, also called the exit site).

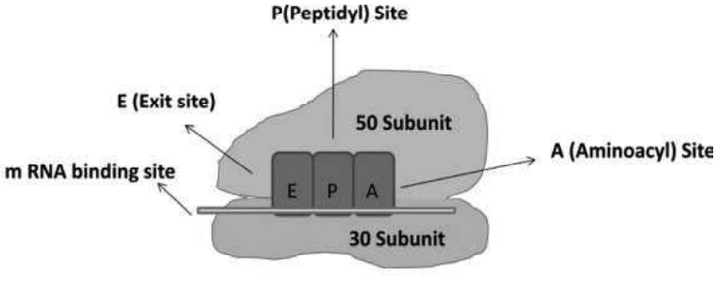


Fig. 5.52 *Structure of 70S Ribosome*

Activation of Amino Acid

By enzyme Amino acyl-t-RNA synthetase amino acids are attached covalently to the t-RNA covalently with a high energy bond, so they are called activated amino acids.

$$\text{Amino acids} + \text{tRNA} + \text{ATP} \xrightarrow{\text{Amino acyl t-RNA synthase}} \text{Amino acyl-t-RNA} + \text{AMP} + \text{PPi}$$

NOTES

Initiation

N formyl-methionine is the first amino acid in Eubacteria which is specific to three codes as AUG, GUG, and UUG. So, this codes at the initiation point codes for N-formyl-methionine but if they are present in between the coding sequence then they code for methionine and valine respectively. Because initiator tRNA and the one used in between the process of translation are different. Initiator tRNA has unique features that distinguish from elongating tRNA in eubacteria. Several initiation factors are required in the process, which include IF-1, IF-2, and IF-3**.** The steps include:

- Binding of small subunit of ribosome to mRNA such that initiation codon lies in partial P site. It is controlled by of IF-3. It principally prevents untimely re-association of large and small subunit of ribosome. Ribosomal binding site of mRNA consists of Shine -Dalgarno sequence and initiation codon. This Shine-Dalgarno sequence 5'-AGGAGGU-3' is located 10bp upstream of initiation codon is complementary to region near 3' end of 16s rRNA, a constituent of small subunit of RNA.

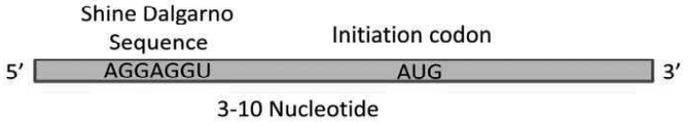


Fig. 5.53 *Ribosome Binding Site*

- Initiator tRNA with N-formyl methionine enters partial P site and binds to mRNA via its anti-codon loop controlled by IF-2. It guides initiator tRNA to its correct position in the initiation complex. It also displays ribosome dependant GTPase activity. Once GTP is hydrolysed then 50S subunit joins to form complete ribosome.
- Upon completion of complex, it forms complete P site and A-site. Then, Second charged tRNA (amino acid laden) enters A site. This tRNA contain an anticodon corresponding to codon in mRNA.

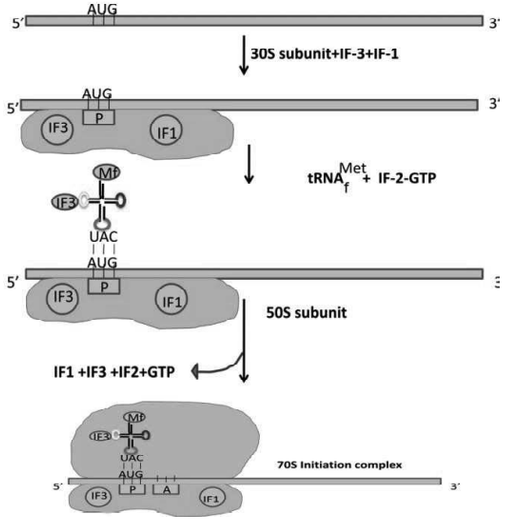


Fig. 5.54 *Initiation of Protein Synthesis*

NOTES

The differences between initiation of prokaryotes and eukaryotes are:

- Eukaryotes contain only one start codon for AUG and it codes for methionine and not N-formyl-methionine. Eukaryotic cells need more initiation factors than prokaryotes. Eukaryotic cells require 12 initiation factors.
- Process of association of mRNA with smaller subunit (40s) is much more complex in eukaryotes, than prokaryotes. 40s first subunit recognizes 5' methylated cap of m-RNA which then leads to scanning process wherein initiation codon is recognized which is aided by ATP dependant helicases that hydrolyse ATP. This recognition of initiation codon is also assisted by Kozak sequences 5'-ACCAUGG-3' similar to Shine-Dalgarno sequence in prokaryotes. But, in prokaryotes there is no scanning process, and directly, 16s subunit binds to region so that initiation codons are in P-site whereas in eukaryotes there is proper scanning process.

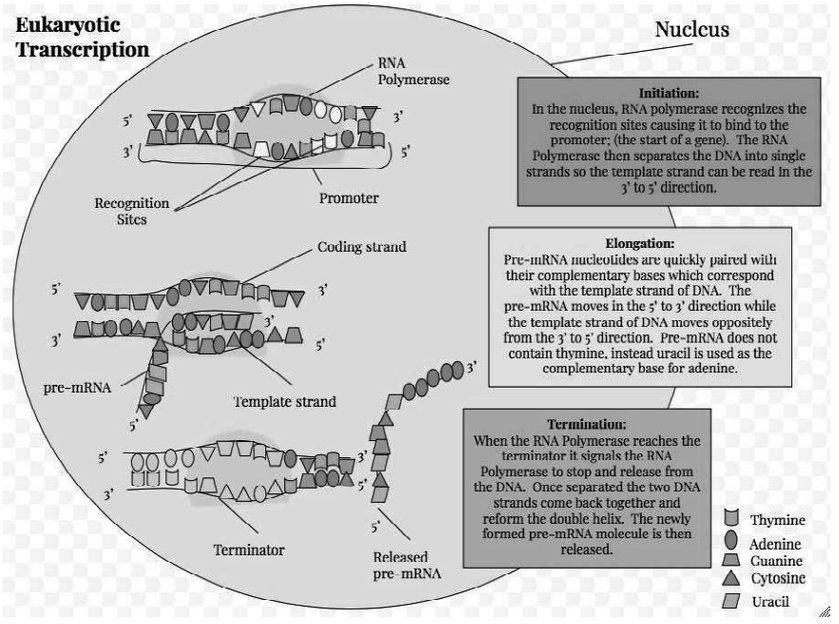


Fig. 5.55 *Translation in Eukaryotes*

5.4.9 Chemical Synthesis of Mono and Tri Nucleosides

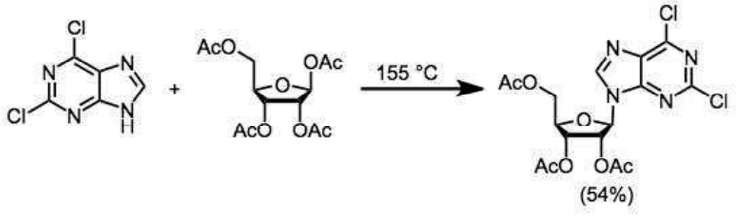
Nucleosides are the glycosylamines obtained by chemical or enzymatic decomposition of nucleic acids and contain two components: a five-carbon sugar (ribose or 2' deoxyribose) and a nitrogen base. The nitrogenous bases are planar, aromatic, heterocyclic molecules. Mostly, they are the derivatives of purine or pyrimidine. The major purine components of nucleic acids are adenine (A) and guanine (G) residues and the major pyrimidine residues are cytosine (C), Uracil (U) (which mainly occurs in RNA), and thymine.

Generally, the nucleosides are synthesized by the coupling of nucleophilic purines and pyrimidine's along with other basic electrophilic heterocyclic compound with a ribose or deoxyribose derivative at the anomeric carbon.

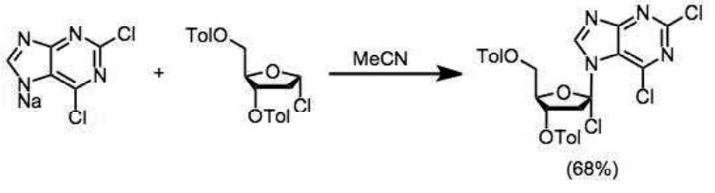
NOTES

The following three methods have been used to synthesize the nucleosides from nucleophilic bases and electrophilic sugars:

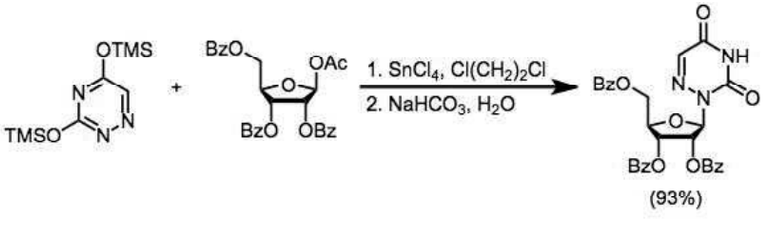
1. **The Fusion Method:** This method involves heating the base and acetyl-protected 1-acetoxyribose to 155°C and results in the formation of the nucleoside with a maximum yield of 70%.



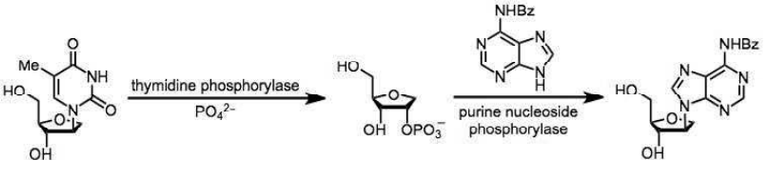
2. **The Metal Salt Method:** This method involves the combination of a metal salt of a heterocyclic compound with a protected sugar halide. Silver and mercury salts were used originally; however, the recently developed methods prefer using the sodium salts.



3. **The Silyl-Hilbert-Johnson (SHJ) Reaction (Vorbrüggen Reaction):** It is the most accepted method for the formation of nucleosides. It is the combination of a silylated heterocyclic compound and a protected sugar acetate (such as 1-O-acetyl-2,3,5-tri-O-benzoyl-beta-D-ribofuranose) in the presence of a Lewis acid.



A second alternative to the above mentioned conventional methods is the enzymatic trans-glycosylation method, which is a kinetically controlled (surpassing the issues of chemical trans-glycosylation associated with thermodynamic control). However, operational complications associated with the use of enzymes are a disadvantage of this method.



Check Your Progress

9. What is DNA replication?
10. What happens in elongation?
11. What is DNA polymerase responsible for?
12. Define Prokaryotic DNA.
13. What is DNA polymerase III? Name its types.
14. What is spindle apparatus?
15. In what direction does the translation of mRNA occur?

NOTES

5.5 ANSWERS TO 'CHECK YOUR PROGRESS'

1. Nucleic acids constitute the most important biomolecules of the cell and are critical entities for all known forms of life. A nucleic acid is a macromolecule composed of chains of monomeric nucleotides.
2. There are two types of nucleic acids; deoxyribonucleic (DNA) acid and ribonucleic (RNA) acid.
3. The basic components of a nucleic acid include three different entities, namely a nitrogenous base, a sugar moiety and a phosphate group.
4. A nucleoside consists of a combination of a nitrogenous base and a sugar (ribose or deoxyribose).
5. Nucleotides comprises of a nitrogenous base linked to a 5-carbon sugar and one or more phosphate group. The phosphate is attached to 5' CH2OH group of sugar part of nucleoside.
6. Complementary base pairing is when each type of base on one strand establishes a link with exactly one type of base on the other strand. Purines form hydrogen bonds with pyrimidines. The arrangement of two nucleotides binding together across the double helix is known as a base pair.
7. A DNA sequence is referred to as 'sense' if its sequence is the same as that of a messenger RNA copy that is translated into protein.
8. DNA is present in cells in the long structures in the nucleus, also known as chromosomes. These chromosomes duplicate before the cells divide and this process is known as DNA replication. Chromosomes are condensed thread-like structures of DNA.
9. The formation of new DNA chains from raw material including one original DNA molecule is known as DNA replication.
10. In elongation, the two single parent strands are acted by DNA polymerase and RNA primer for addition of new polynucleotides. DNA polymerase has 52 –32 activity. DNA replication systems require a free 3' hydroxyl group before synthesis can be initiated (DNA template is read in 32 to 52 direction whereas a new strand is synthesized in the 52 to 32 direction).

NOTES

11. DNA polymerase is responsible for polymerization of deoxyribonucleotides the building blocks of DNA, during the replication of DNA. During this process, DNA polymerase 'Reads' the existing DNA strands to create two new strands that match the existing ones. DNA polymerase adds nucleotides to the 3' end of a DNA strand, one nucleotide at a time.
12. Prokaryotic DNA polymerase exists in two forms as core polymerase and holoenzyme. Core polymerase synthesizes DNA from the DNA template but it cannot initiate the synthesis alone but holoenzyme can accurately initiate DNA synthesis.
13. DNA polymerase III holoenzyme is the primary enzyme involved in DNA replication in E. coli and belongs to family C polymerases. It consists of three parts:
    a. Pol III Core
    b. Beta Sliding Clamp
    c. Clamp-Loading Complex
14. The spindle apparatus is the cytoskeletal structure formed during cell division to separate sister chromatids between daughter cells and found in Eukaryotic cell.
15. The translation of mRNA occurs in 5′ to 3′ directions converting the nucleotide sequence into the amino acid sequence based on the universal genetic code.

5.6 SUMMARY

- Nucleic acids constitute the most important biomolecules of the cell and are critical entities for all known forms of life. A nucleic acid is a macromolecule composed of chains of monomeric nucleotides.
- Nucleic acids were discovered by Friedrich Miescher in 1869. He reported that he had found a substance within the nuclei of human white blood cells, which was weakly acidic in nature and whose function was unknown.
- The basic components of a nucleic acid include three different entities, namely a nitrogenous base, a sugar moiety and a phosphate group.
- There are two types of nitrogenous bases present in all nucleic acids. These are pyrimidine and purine.
- Hydrolysis of nucleic acids by either chemical or enzymatic methods leads to the cleavage of phosphodiester backbone producing smaller oligonucleotide containing up to 20 residues, nucleosides, free purine or pyrimidine bases, ribose or deoxyribose and phosphates.
- Purines and pyrimidines, being complementary bases, can participate in base pairing, based on the specific shapes and hydrogen bond properties.
- A nucleoside consists of a combination of a nitrogenous base and a sugar (ribose or deoxyribose). The bond between them is called the beta-glycosidic linkage.

NOTES

- RNA and DNA are polymers constituted of monomers called mononucleotide units.
- In the early 1950s, four scientists, James Watson and Francis Crick at Cambridge University and Maurice Wilkins and Rosalind Franklin at King's College, determined the true structure of DNA from data and X-ray crystallography of a molecule taken by Franklin.
- A unique feature of DNA molecule is that it is able to make exact copies of itself, or is self-replicate.
- Ribonucleic Acid (RNA) has derived its name from the sugar group present in its backbone, that is, ribose. Some of the features of DNA and RNA are common.
- DNA is present in cells in the long structures in the nucleus, also known as chromosomes. These chromosomes duplicate before the cells divide and this process is known as DNA replication.
- The formation of new DNA chains from raw material including one original DNA molecule is known as DNA replication.
- Replication occurs in three steps: Initiation, Elongation and Termination. DNA replication is semi-conservative and bidirectional in nature.
- Replication machineries include primosotors or replication enzymes: DNA polymerase, DNA helicases, DNA clamps and DNA topoisomerases, and replication proteins; for example Single-Stranded DNA Binding proteins (SSB).
- Eukaryotes initiate DNA replication at multiple points in the chromosome, so replication forks meet and terminate at many points in the chromosome.
- During the elongation phase double helical DNA structure get unwind and forms a Y- shaped replication fork, which leads to the formation of a continuous leading strand and a discontinuous lagging strand with Okazaki fragments on it.
- DNA polymerase is responsible for polymerization of deoxyribonucleotides the building blocks of DNA, during the replication of DNA.
- Prokaryotic DNA polymerase exists in two forms as core polymerase and holoenzyme.
- Core polymerase synthesizes DNA from the DNA template but it cannot initiate the synthesis alone but holoenzyme can accurately initiate DNA synthesis.
- During RNA transcription, RNA polymerase opens the double-stranded DNA to utilize one strand of the exposed nucleotides as a template for RNA synthesis.
- In most of the Prokaryotes, a single type of RNA polymerase specifies transcription of all the three RNA types.
- Eukaryotes have multiple types of nuclear RNA polymerases (I, II, III, IV and V) each responsible for synthesis of a distinct type of RNA.

NOTES

- Spindle apparatus includes the microtubules, associated proteins (kinesin and dynein), condensed chromosomes and any centrosomes or aster.
- Two models explains the organization of spindle apparatus. In the search-and-capture model, the spindle is predominantly organized by the poleward separation of centrosomal MicroTubule Organizing Centres (MTOCs).
- Spindle microtubules emanate from centrosomes and seek out kinetochores and bind a kinetochore to become stabilized and exert tension on the chromosomes.
- In self-assembly model, microtubules undergo acentrosomal nucleation among the condensed chromosomes.
- Spindle apparatus perform its three major functions: Chromosome Alignment, Chromosome Segregation and Bipolarity.
- Transcription is a process of synthesis of RNA or transcript. It is based on the complementary base pairing from DNA templates catalyzed by DNA dependent RNA polymerases.
- mRNA templates obtained from the transcription of DNA, are converted into the proteins by the process known as translation.
- Genetic code is common for all life forms and so it is universal. But there are some exceptions to this rule. For example, as UGA is a stop codon but it codes for tryptophan in Mycoplasma, Spiroplasma, and mitochondria of several species.

5.7 KEY TERMS

- **Nucleic Acids:** These are biopolymers, or large biomolecules, which are essential to all known forms of life. They help protein synthesis in the body.
- **Deoxyribonucleic Acid (DNA):** It is a molecule consists of two polynucleotide chains that coil around each other forming a double helix and carrying genetic instructions for. It is present in all living organisms as the main element of chromosomes.
- **Ribonucleic Acid (RNA):** It carries instructions from DNA to control protein synthesis. However, RNA carries the genetic information instead of DNA in some viruses.
- **DNA Template:** It is the strand used by DNA polymerase or RNA polymerase to attach complementary bases during DNA replication or RNA transcription.
- **Gene Expression:** It is the process of producing a protein from its DNA- and mRNA-coding sequences.
- **Gene:** It is a portion of DNA that contains instructions for making a Protein.
- **Tubulin:** It belongs to superfamily of globular proteins. α- and β-tubulins polymerize into microtubules, which is a major component of the Eukaryotic cytoskeleton.

5.8 SELF ASSESSMENT QUESTIONS AND EXERCISES

NOTES

Short-Answer Questions

1. Write a short note on the chemical and enzymatic hydrolysis of nucleic acids.
2. What are the properties of purines and pyrimidines?
3. How are polynucleotides formed?
4. Write a short note on the different forms of DNA structure.
5. State the role of topoisomerase and SSB proteins during DNA replication.
6. What is microtubule polymerization and depolymerization?
7. List the differences between leading and lagging strand.
8. Briefly explain the concept of genetic code.
9. Differentiate between self-assemble model and search and capture model of organization of spindle apparatus.

Long-Answer Questions

1. Explain the components of nucleic acids.
2. What is DNA replication? Explain the process of DNA replication.
3. Demonstrate the structure and roles of DNA polymerase in eukaryotes.
4. Describe the different types of spindle apparatus and their organization.
5. Explain the structure and roles of Prokaryotic and Eukaryotic RNA polymerase.
6. What is mitotic apparatus? Describe its structure and function.
7. What are MAP's? Describe their types and role in spindle formation.
8. Explain the function of different subunits of Eukaryotic RNA polymerase.
9. Describe the process of transcription in detail.

5.9 FURTHER READING

Rastogi, S. C. 2012. *Cell Biology*, 3rd Edition. New Delhi: New Age International (P) Ltd.

Vyas, S. P. and A. Mehta. 2011. *Cell and Molecular Biology*, 1st Edition. New Delhi: CBS Publisher & Distributors.

Verma, P. S. and A. K. Agarwal. 2004. *Cell Biology, Genetics, Molecular Biology, Evolution and Ecology*. New Delhi: S. Chand & Co. Ltd.

De Robertis, E. D. P. and E. M. F. De Robertis (Jr.). 2010. *Cell and Molecular Biology.* Philadelphia: Lippincott Williams & Wilkins.

Hunter, L. E. 2009. *The Process of Life: An Introduction to Molecular Biology*. USA: The MIT Press.

NOTES

Weaver, R. F. 2008. *Molecular Biology*. New York: McGraw Hill Higher Education.

Becker, W.; L. Kleinsmith; J. Hardin; and G. Bertoni. 2008. *The World of the Cell*. London: Pearson Education.

Alberts, B.; A. Johnson; J. Lewis; M. Raff; K. Roberts and P. Walter. 2007. *Molecular Biology of the Cell*. New York: Garland Publishing Inc,

Karp, G. 2007. *Cell and Molecular Biology: Concepts and Experiments*. New York: John Wiley and Sons Inc.

Gupta, P. K. 1999. *Cell and Molecular Biology*. Meerut: Rastogi Publications.

NOTES

NOTES