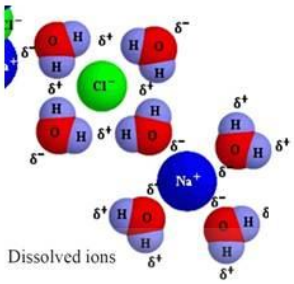
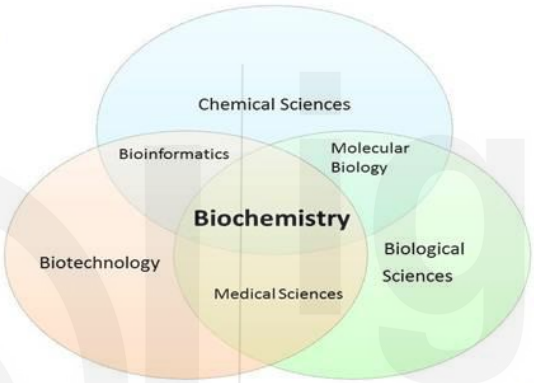
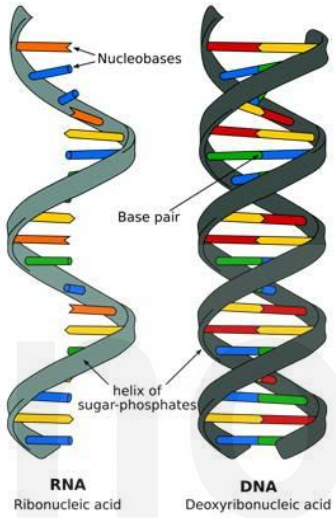
Indira Gandhi National Open University
School of Sciences

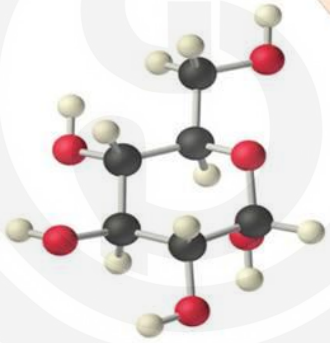
BBCCT-101

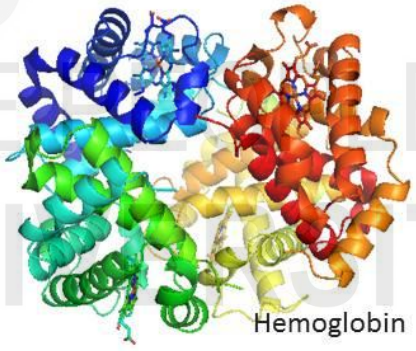
Molecules of Life











Indira Gandhi National
Open University
School of Sciences

BBCCT-101
MOLECULES OF LIFE

Course Design Committee

Prof. Bechan Sharma
Dept. of Biochemistry
University of Allahabad

Prof. Reena Gupta
Dept. of Biotechnology
H.P. University, Shimla

Prof. D. V. Devaraju
Dept. of Biochemistry
Bangalore University

Prof. K. Vali Pasha
Dept. of Biochemistry
Yogi Vemana University,
Andhra Pradesh

Dr. Suneeta Joshi
Dept. of Biochemistry
Daulat Ram College
University of Delhi

Prof. Ranjit K. Mishra
Dept. of Biochemistry
University of Lucknow

Prof. Sanjeev Puri
UIET, Panjab University

Prof. Seemi Farhat Basir
Dept. of Bio Sciences
Jamia Milia Islamia

Prof. Vijayshri
Former Director
School of Sciences, IGNOU

Dr. Parvesh Bubber
SOS, IGNOU

Dr. M. Abdul Kareem
SOS, IGNOU

Dr. Arvind Kumar Shakya
SOS, IGNOU

Dr. Maneesha Pandey
SOS, IGNOU

Dr. Seema Kalra
SOS, IGNOU

Block Preparation Team

Prof. K. Vali Pasha (Content Editor)
Dean, Faculty of Science
Department of Biochemistry
Yogi Vemana University
Kadapa-516003
Andhra Pradesh

Faculty Members, SOS, IGNOU

Dr. M. Abdul Kareem (Unit 1)

Dr. Arvind Kumar Shakya (Units 2 & 12)

Dr. Parvesh Bubber (Units 3 & 4)

Dr. Seema Kalra (Units 5 to 11)

Dr. Maneesha Pandey (Units 13 & 14)

Course Coordinator: Dr. M. Abdul Kareem

Print Production	Cover Page and Graphic Design
Sh. Sunil Kumar Assistant Registrar (Pub.) SOS, IGNOU	Dr. M. Abdul Kareem Assistant Professor SOS, IGNOU

Acknowledgement: Mr. Sumit Verma for word processing.

March, 2020

ISBN:

Further information on the Indira Gandhi National Open University courses may be obtained from the University's office at Maidan Garhi, New Delhi-110 068 or the official website of IGNOU at www.ignou.ac.in.

Printed and published on behalf of Indira Gandhi National Open University, New Delhi by Prof.Sujata Varma, Director, School of Sciences, IGNOU.

Laser Typeset by : Rajshree Computers (Near Sec. 2, Dwarka), V-166A, Bhagwati Vihar, Uttam Nagar, New Delhi-110059

Printed at :

MOLECULES OF LIFE : INTRODUCTION

Biochemistry focuses on how the origin of life takes place from simple atoms to complex organisms under the diverse environmental conditions. The aim of biochemists is to explain the mechanisms responsible for metabolic reactions occurring in a living cell. The foundation of biochemistry is based on understanding the molecules of life i.e., biomolecules; thousands of which are found in the living cells. You'll be surprised to know that the remarkable properties of living organisms arise from thousands of different biomolecules. A major objective of biochemistry is to understand the characteristic nature, properties and functions of these biomolecules. Biochemists have sought to isolate these to determine the structure, behaviour and analyze their functions.

"Molecules of Life" is one of the core courses in Biochemistry (Hons.) being offered as an elective for undergraduate students at Indira Gandhi National Open University. While preparing this course it has been kept in mind that, students joining B.Sc. (Hons.) Programme must have studied science subjects i.e., Biology, Physics and Chemistry at 10+2 level. In recent years biochemistry has made a remarkable growth that has been contributing to the understanding of many basic and advanced issues related to health, nutrition and medical sciences.

In this course the structure and functions of various biomolecules like amino acids, proteins, carbohydrates, lipids and nucleic acids have been described. In addition to these topics a separate unit on water has been included. The entire course is divided into four blocks of which Block 1 and Block 2 contain four units each, whereas Block 3 and Block 4 containing three units each.

Block 1 explains about origin of biochemistry and its progress as an interdisciplinary subject of research. It is well known fact that water acts as a universal solvent because of its unique properties. Therefore one unit is dedicated to illustrate the structure and biological significance of water. A detail discussion on structure of amino acids along with their physicochemical properties has been incorporated. A short account of peptide bond formation and its significance, structural hierarchy of proteins has been included.

In Block 2 a detailed discussion on various types of carbohydrates ranging from monosaccharides to polysaccharides has been discussed. In this block structure, properties and functions of carbohydrates have been explained in detail. A special emphasis has been given to glycobiology to explain blood group antigens.

Block 3 deals with lipids as a major source of energy and explains how essential these are in the formation of plasma membrane of cell and other sub-cellular organelles. Lipids also play a vital role in cell signaling and synthesis of steroidal hormones.

Block 4 explains the structure and functions of vitamins and nucleic acids. You will also study how DNA is acting as a genetic material. It also explains the physical and chemical changes occurring in DNA upon exposure to ultraviolet rays and chemical reagents.

The aim of this course is to generate interest in our distance learners to acquire the knowledge of biochemistry. Each unit has been written with an eye to help students learn and master the fundamentals of biochemistry. It has been tried in best manner to enable the learners cope up with the complexities and challenges of biochemistry in a simplified and systematic way. The structural outline in the beginning of each unit is a road map of the unit.

The objectives mentioned reflect the teaching and learning approaches. The running text describes and illustrates basics of biochemistry in a concise, learner friendly and interesting manner. It is supported by suitable figures, tables and several diagrams to integrate the concept of the unit. The key features and concepts have been highlighted and important points have been broken out of text and presented in boxes. A variety of teaching and learning approaches such as an inbuilt self-assessment exercises and terminal questions along with answers provided in each unit will support the learner to evaluate and meet the objectives of the given self-learning material.

Objectives

The broad objectives of this course will enable you to:

- describe the origin of biochemistry;
- explain the significance of biochemistry in understanding life;
- clarify the basic concepts of the biochemistry;
- draw the basic structures and classification of biomolecules;
- distinguish different types of biomolecules;
- explain the biological significance of biomolecules;
- illustrate and explain the importance of water and vitamins in maintaining normal metabolism; and
- enumerate salient features of nucleic acids and explain how DNA is acting as a genetic material.

We wish that after studying this Course you will acquire basic understanding of the biochemistry and its applications in different fields.

We wish you the best in this endeavour!!

Indira Gandhi National
Open University
School of Sciences

BBCCT-101
MOLECULES OF LIFE

Block

1

INTRODUCTION TO BIOCHEMISTRY

BLOCK 1: INTRODUCTION TO BIOCHEMISTRY

The first block of this course consists of 4 units dealing with the basic concepts of origin of biochemistry, water, amino acids and peptides and proteins. Unit 1 describes how biochemistry evolved from biology and chemistry to an interdisciplinary subject. A detailed discussion on chemical foundations of life has been dealt. Different units of measurement used in biochemistry have also been included.

Unit 2 deals with detailed discussion on the unique properties of water followed by its biological significance. We will also learn about the ionization of water, concept of pH and buffers. At the end of this unit detail explanation on water as a reactant and fitness of the aqueous environment has been given.

Unit 3 deals with the study of structure of amino acids and their classification along with their physico-chemical properties.

Unit 4 of this course is on peptides and proteins. The formation of peptide bond and its significance in maintaining the stability of proteins is discussed. At the end of the unit the structure and organization of proteins which are the most versatile biomolecules has been explained.

Objectives

After studying this block you should be able to:

- describe the origin of biochemistry and chemical foundations of life;
- explain the chemistry of water and the concept of pH;
- draw the structure and classify amino acids;
- distinguish the formation of peptides and proteins; and
- illustrate the structure and classify proteins.

We hope that after studying this block you will acquire understanding of the fundamentals of biochemistry.

Wishing you success in this endeavour!

UNIT 1

FOUNDATIONS OF BIOCHEMISTRY

Structure

1.1 INTRODUCTION

The term ***Biochemistry*** is self-explanatory which is evolved from two basic Science subjects i.e., ***Biology*** *and* ***Chemistry***. Biochemistry is a branch of Science that deals with **"*Chemistry of life and living systems*"**. In your earlier classes you might have studied both Biology and Chemistry as individual subjects. In this unit, you will see how biology and chemistry merged with each other leading to the genesis of biochemistry. You shall be introduced to the significance of biochemistry in understanding our health and disease status. You will be motivated by studying and understanding the Cellular and Molecular basis of life in a cell. You shall also learn how various chemical entities involve in maintaining life inside biological systems.

Besides this, you will be briefly introduced to the major historical discoveries occurred in biochemistry that enabled human beings for better understanding of the *chemistry of life*. Along with the glorious history, we shall also describe how biochemistry progressed as an interdisciplinary branch of science. The best part of biochemistry is, it includes the basic principles of all major science subjects and answers the unanswered questions of life in living systems. At the end, we shall also discuss units of measurement used to express various physicochemical properties of substances.

"Much of life can be understood in rational terms if expressed in the language of chemistry. It is an international language, a language for all of the time, and a language that explains where we came from, what we are and the physical world will allow us to go."

– Arthur Kornberg (1918-2007)

Objectives

After studying this unit, you should be able to:

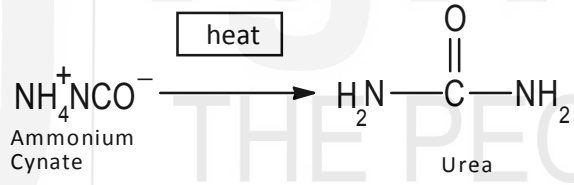
- recall the important historical milestones of biochemistry;
- define the term biochemistry;
- describe the interdisciplinary nature of biochemistry; and
- analyse basic composition and components of life.

1.2 ORIGIN OF BIOCHEMISTRY

Friedrich Wöhler

Every mother who curdles milk or prepares *dosa* dough is a biochemist! It may sound little ironical but it is true. Humans beings have been using the applications of biochemistry like fermentation, pickling, and tanning since ages, but they were not aware of science behind it. In the early nineteenth century the path breaking work done by Friedrich Wöhler a German chemist laid the stone which further directed to the development of a new branch of science that is recognised as biochemistry. The distinguishing work done by Wöhler is the *synthesis of Urea* in the laboratory without involving any living (biological) material. He did this pioneering work in the year 1828, by heating ammonium cyanate. The interesting part of this synthesis was 'production of organic substance from inorganic substance'.

Louis Pasteur



Inorganic **Organic**

Another important milestone in the history of biochemistry was the work done by Louis Pasteur, where he explained the significance of yeast cells in converting sugar to alcohol. This landmark work was performed in the mid-19th century. Pasteur explained that yeast cells encourage fermentation (alcohol production) process. He also established the fact that bacteria present in the fermentation mixture are responsible for the production of citric acid (sour taste). Since then people started avoiding usage of bacteria in fermentation mixtures. This finding led to a major breakthrough in the process of wine production. Later, in the year 1897, Eduard and Hans Buchner, the two German brothers, showed extracts of yeast cells could carry out the entire process of fermentation. Buchenr's work disclosed the possibilities for ***in vitro*** (Latin, "in glass") research.

Experiments that performed inside a test tube or outside a living system are called ***in vitro*** experiments. However, the experiments carried out in a living system are referred as ***in vivo.***

Later, in the year 1903 Carl Neuberg coined the term biochemistry. In the first half of 20th century, some of the basic reactions occurring inside biological systems were explained. One of the significant metabolic cycle was illustrated in 1937 by Hans Adolf Krebs. He described citric acid cycle for which he received Nobel Prize in physiology. Similarly, the reactions of glycolysis (splitting of sugar) were explained by Gustav Embden and Otto Meyerhof in

1940. These are some of the significant discoveries that showed a wonderful impact in the field of biochemistry.

Up to mid 20th century, it was believed that proteins are hereditary factors. But the experiments conducted by Hershey and Chase in 1952 proved that DNA is the genetic material but not proteins. Later in the year, 1953 ground-breaking discovery of the 20th century took place, when James Watson and Francis Crick described the double helical structure of DNA. This discovery helped science fraternity to understand how genetic information is stored in DNA! In the year 1968 Har Gobind Khorana, an Indian American biochemist discovered genetic code and received Nobel Prize in physiology in 1968.

Major progress in biochemistry took place in mid 20th century due to the development of several biochemical and biophysical techniques like Chromatography, NMR, X-ray diffraction, Radio isotope labeling and Electron microscopy. This journey of biochemistry is continuing and expanding everyday with the inception of cloning, genetic engineering and several other advanced molecular and cell biology techniques. The major discoveries that happened in the fields of chemistry, physics, biology and advanced cell biology allowed scientists for the better understanding of life and its chemical existence. As time progressed biochemistry evolved as an independent subject with interdisciplinary concepts, components and connections among basic sciences (Biology, Chemistry and Physics). Hence, there is no hesitation in saying that Biochemistry is an interdisciplinary subject in its nature. In next section, we will be studying more about the interdisciplinary nature of Biochemistry and how and why Biochemistry is considered as an interdisciplinary subject.

Har Gobind Khorana

1.3 BIOCHEMISTRY AS A DISCIPLINE AND INTERDISCIPLINARY SUBJECT

So far we have discussed about historical background of biochemistry, during which we came to know that biochemistry appeared as a central part of research interest in chemistry as well as biology. Apart from these two subjects, biochemistry interacts with fundamentals of physics and computer science (Fig.1.1). Now, let us discuss in detail about interdisciplinary nature of biochemistry by taking few examples and applications of core research areas in biochemistry.

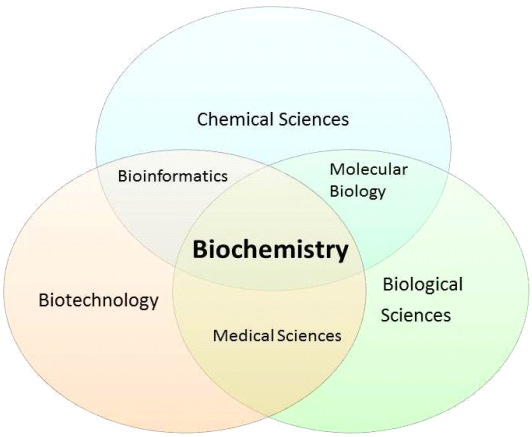


Fig. 1.1: Interaction of Biochemistry with other allied branches of science.

Structural Biochemistry: In this area of study major inputs will be taken from basics of organic along with physical chemistry. This area of study is useful in understanding how and what composes life. In this, we will be studying about biomolecules structures: functions and their interactions with each other along with thermodynamics of a living system.

Biophysics: This is another widely studied area of biochemistry, where researchers will be interested in knowing the kind of physical interactions taking place between biomolecules and the surrounding environment. These interactions may be either inside the cell or outside the cell. That enables the maintenance of life in a living system. Most of the biophysical techniques (will be discussed in detail in Course BBCCT-105, "Proteins") depend on basic physical and chemical properties of biomolecules. These techniques are frequently used in diagnostics, medical and clinical research fields for separation, isolation, purification and identification purposes.

Cell and Molecular Biology: This is another upcoming area in biochemistry research with wide opportunities for discovery and innovative research. The major principles applied in this field are from microscopy (electron, confocal and fluorescence microscopy), microbiology, cell and tissue culture.

Medical and Clinical Biochemistry: It is interesting to know that all Medical graduates and Paramedical students read biochemistry as a compulsory subject. The reason why we are emphasising this point is, to understand health and disease status of any living system knowledge of biochemistry is very much essential. In simple words, any imbalance in the normal metabolism leads to metabolic disorders causing disease (Fig.1.2). The one with sound knowledge of biochemistry can easily identify and understand the cause of disease status.

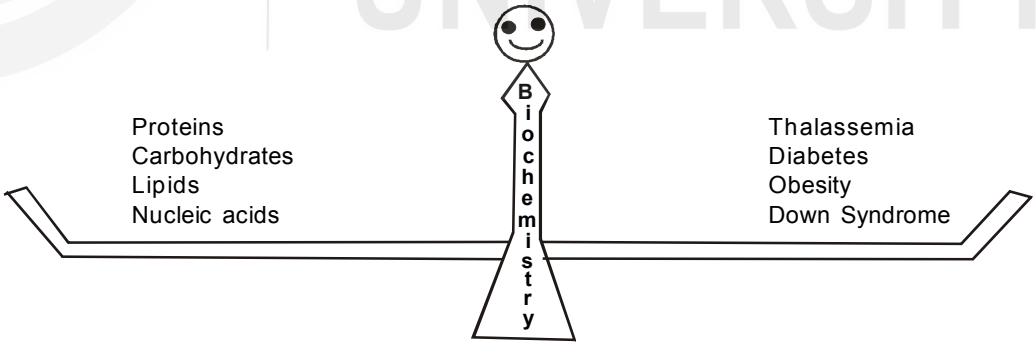


Fig.1.2: This diagram shows, how knowledge of biochemistry is useful in understanding the health and disease status of an individual. For example, detailed knowledge of lipd metabolism is needed to address the issues related to obesity.

Nutritional Biochemistry: This is another junction where biochemistry cross-talk with principles of nutrition. Our diet is majorly composed of carbohydrates, proteins, lipids and vitamins. These are the major biomolecules that are responsible for energy, structure and functioning of our body. Hence, one with adequate knowledge of biochemistry is capable of understanding reasons for

certain dietary disorders like malnutrition, kwashiorkor and marasmus. Moreover, life style disorders like Obesity and Diabetes are well managed by the sound knowledge of biochemistry.

Bioinformatics: Other upcoming and promising area which is widely involved in *drug discovery* and design. In this area by using computers, biological data can be accessed, stored and retrieved. Tools of bioinformatics are none other than bioinformatics softwares and high end computers, which are developed by using principles of maths, physics and chemistry. The major applications of bioinformatics are in the field of biochemistry and pharmacy.

These are some of the important sub-areas of biochemistry which clearly indicates that biochemistry is having interactions with almost all fields of sciences. Hence biochemistry evolved as an independent branch with interdisciplinary nature.

SAQ 1

Fill in the blanks with correct answers:

i) The term biochemistry was coined by .....

ii) .....synthesized Urea by using ammonium cyanate

iii) .....is the Indian scientist who received Nobel Prize for the discovery of genetic code.

iv) Bioinformatics is widely used in the discovery of .....

v) ..... was explained by Adolf Krebs.

1.4 CELLULAR AND CHEMICAL FOUNDATIONS OF LIFE: CHEMICAL ELEMENTS OF CELLS AND LIVING SYSTEMS

Up to now, we have discussed interdisciplinary nature of biochemistry. Now it's time to know the basics of life i.e., in which form does life exists? In your lower classes, you might have studied that cell is the basic unit of life in a living system. The hierarchy of organization in a living system is like cells to tissue, tissues to organ, organs to a system, and systems to an organism. The complexity of organization and structure slightly varies from invertebrates to vertebrates in other words from smaller organisms to higher organisms. At the same time, the composition of the cell varies from organism to organism. Fig. 1.3 illustrates the chemical composition of a normal human.

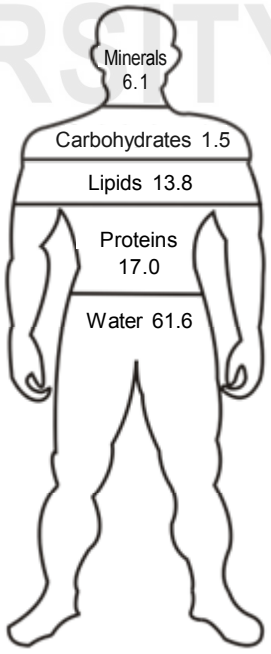


Fig.1.3: Chemical composition of a normal man (weight 65 kg).

We all know that on this earth life exists in different forms like microorganisms, plants, and animals. All these forms of life display certain common properties like motility (except plants), growth, reproduction. The fact to be remembered here is, that all these organisms are made up of similar organic and inorganic chemical entities like proteins, carbohydrates, lipids, nucleic acids and vital minerals. The crux to recognise in this section is, "these biomolecules are common in their basic composition (precursor elements like carbon, nitrogen, oxygen, phosphorous etc. refer Table 1.1) but differ in their arrangement from one organism to another".

Table 1.1: Inorganic elements and their biological functions

Elements	
Macro	**Functions**
Carbon (C), Hydrogen (H), Oxygen (O), Nitrogen (N)	Synthesis of organic compounds like proteins, carbohydrates, lipids and nucleic acids.
Sulphur (S)	A component of sulphur containing amino acids.
Sodium (Na)	Regulates a osmotic pressure in cells. Helps in the transmission of nerve impulses.
Magnesium (Mg)	Acts as a cofactor of kinase enzymes Involved in protein synthesis.
Calcium (Ca)	Required for the formation of strong bones and teeth. Helps in muscle contraction. Promotes blood clotting.
Phosphorus (P)	Synthesis of nucleotides, Synthesis of ATP. Formation of strong bones and teeth.
Chlorine (Cl)	Component of the biological buffer Synthesis of hydrochloric acid by the gastric glands in the stomach.
Potassium (K)	Regulates osmotic pressure in cells. Required in muscle contraction and transmission of nerve impulses.
Micro	**Functions**
Copper (Cu)	Cofactor for several enzymes
Iodine (I)	Involved in the synthesis of thyroid hormones.
Ferrum (Fe)	Involved in the synthesis of hemoglobin and respiratory enzymes.

After going through above table, we need to realise that forms (shape or size) of biomolecules may be different but their basic composition remains same. Let us understand this concept in detail:

Non-living to living! It might be surprising to know that, major biomolecules are synthesised from non-living inorganic chemical elements. When we observe these non-living elements they exhibit all the physical and chemical properties that are characteristic of a non-living or lifeless material. But these lifeless elements arrange to form biomolecules (Fig.1.4) like carbohydrates, proteins, nucleic acids and lipids . Their precise interaction with each other is responsible for the maintenance of life inside a cell.

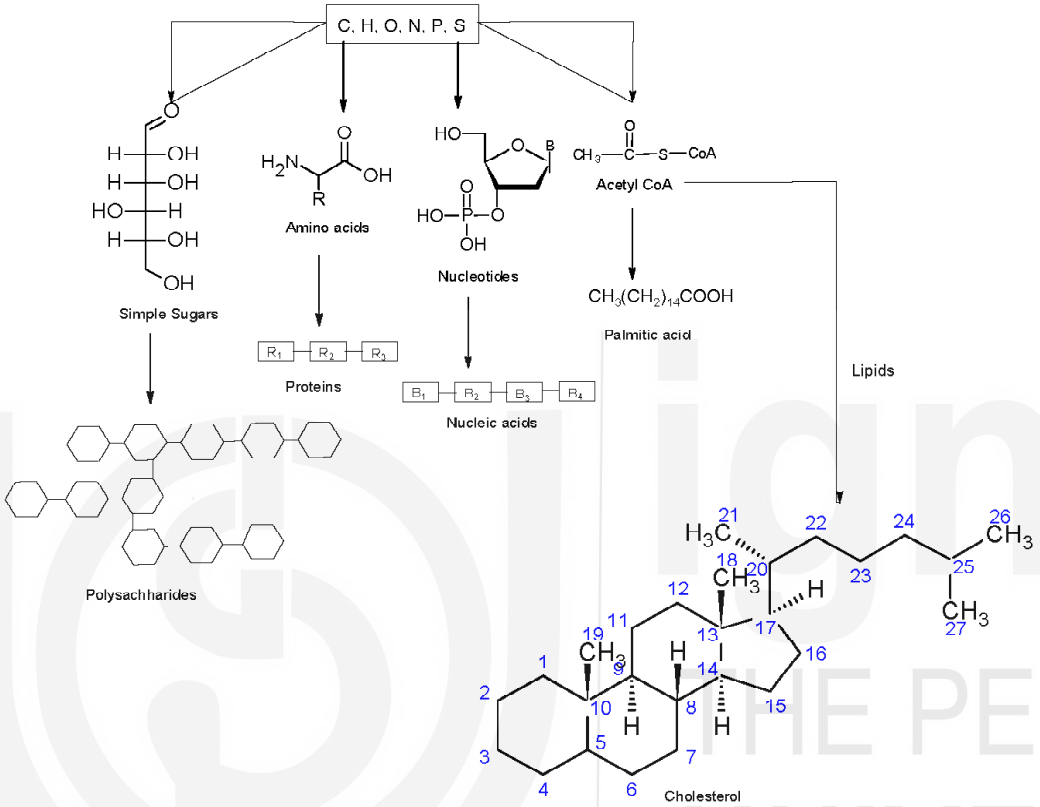


Fig.1.4: Synthesis of biomolecules from inorganic elements.

(Modified from http://www.tetonnm.com/pics/IndependentSamplePages/1-893441-42-3.pdf)

In your lower classes you might have studied that DNA acts as a hereditary (genetic) material that carries information from one generation to the next. Both animal and plant cells possess structurally similar DNA (double helix) i.e., composed of structurally similar nitrogenous bases like adenine, guanine, cytosine, and thymidine. Though these DNA's are structurally similar, they differ in their functions (expression of characters). Here we need to understand that in spite of structural similarities in DNA double helix and nitrogenous base composition, they differ in their nucleotide sequence, that is responsible for functional diversity. In nature every living cell is capable of synthesising desired biomolecule according to the genetic information available in its DNA sequence. In the process of biological synthesis of biomolecules enzyme machinery present in a living cell uses chemical elements as precursors.

Since from the beginning of this section we have been studying that, biomolecules are vital for maintenance of life, let us know some of the important biological functions exhibited by biomolecules in Table 1.2.

You will be studying in detail about the structure, classification and biological significance of vital biomolecules in next units of this course.

Important functions of Water (inorganic molecule) in maintaining life are: acts as solvent; major component of body fluids; regulates body temperature; vital molecule in digestion; absorption and respiration.

Table 1.2: important functions of biomolecules

Biomolecules	Functions
Carbohydrates	Primary source of energy.
Proteins	Form structural components like myosin of muscles and keratin of hair. Build new cells for growth & renew damaged tissue. Involved in the synthesis of blood clotting factors and hemoglobin. All enzymes are proteins.
Lipids	Source of energy. Phospholipids are major constituents of the plasma membrane. Maintenance of body temperature. Synthesis of wax.
Nucleic acids	Hereditary factor

Concept

In this picture a toy house is made by using blocks. You can dismantle and rearrange these blocks to form another shape or toy model. Imagine these blocks with elements like C, O, N, H, S, P etc. that rearrange inside living cells to form desired biomolecules.

1.5 UNITS OF MEASUREMENT IN BIOCHEMISTRY

Units of any measurement play a vital role in our routine life starting from using milk in kitchen to purchasing a designer made suit. These units of measurement facilitate scientists across the world to communicate and exchange their results with one another. The units related to scientific measurement are assigned by the International System of Units (SI). SI is the short form of the French term *Le Système international d'unités,* developed in 1960 from an older metric system. In this section you will be studying about different types of "Units" used to quantify or for standard measurement of various materials and reagents used in biochemistry.

Biochemists routinely use units such as microgram, milligram, litre etc., to express the results in a biochemical analysis. These units enable to observe and analyse the data obtained before and after conducting experiments. Table 1.3 (a-c) shows some of the basic units like weight, length and volume used frequently in practical biochemistry. However, we will be discussing the detailed terminology of specific concentration like Molarity (M) and Normality (N) in the lab course: BBCCL-102.

Table 1.3: Units of measurement for weight, length and volume

a) **Weight**

Unit (One)	Expression	Equivalent
Kilogram	kg	1000 g
Gram	g	1000 mg or 10^{-3} kg
Milligram	mg	1000 µg or 10^{-3} g
Microgram	µg	1000 ng or 10^{-6} g
Nanogram	ng	1000 pg or 10^{-9} g
Picogram	pg	1000 fg or 10^{-12} g
Femtogram	fg	0.001pg or 10^{-15} g

b) **Length**

Unit (One)	Expression	Equivalent
Metre	m	1000 mm
Millimetre	mm	1000 µm
Micrometre	µm	1000 nm
Nanometre	nm	1000 pm
Picometre	pm	1000 fm
Femtometre	fm	0.001pm

c) **Volume**

Unit (One)	Expression	Equivalent
Litre	L	1000 mL
Millilitre	mL	1000 µL
Microlitre	µL	1000 nL
Nanolitre	nL	1000 pL
Picolitre	pL	1000 fL
Femtolitre	fL	0.001pL

Biomolecular Dimensions: The dimensions of mass and length for biomolecules are expressed in '*dalton*' and 'nanometre'. Molecular mass is denoted as '*m'*, this is defined as the 'mass of one molecule'. One Dalton (Da) is equivalent to one-twelfth the mass of carbon; a kilodalton (kDa) is 1,000 daltons; a megadalton (MDa) is 1 million daltons. However, nanometre is abbreviated as 'nm'. One nm is equal to one billionth of metre i.e., 10^{-9} metre (refer length in Table1.3).

SAQ 2

Tick (✓) the correct option given in the brackets:

i) Amino acids are the building blocks of proteins (**True or False**)

ii) Polysaccharides are made up of nucleotides (**True or False**)

iii) Cell is the basic unit of life (**True or False**)

iv) 1 milligram is equal to 100 µg (**True or False**)

1.6 SUMMARY

Origin of biochemistry took place early in the 18th century. Later in the mid of 19th century biochemistry established as a subject. The word "biochemistry" was first proposed in 1903 by Carl Neuberg, a German chemist.

Biochemistry has become the language of biology and medicine.

Modern biochemistry evolved on the applications of advanced techniques to biological problems. Biochemistry deals with the structural and functional chemistry of the living organisms that include the structure of individual molecule and their molecular regulation and expression.

Owing to the interactive nature of biochemistry with various branches of science subjects it is considered as an interdisciplinary subject in nature.

It is interesting to know that biomolecules are composed of inorganic elements like C, N, H, O, S, and P. The basic unit of life is Cell which further joins to form tissue and so on to form a complete organism.

SI units are used to measure and express various properties of biomolecules.

1.7 TERMINAL QUESTIONS

1. Explain the synthesis of Urea by Wöhler.
2. Describe about interdisciplinary nature of biochemistry.
3. Write a brief note on the chemical foundations of life.
4. Enlist the major functions of biomolecules.

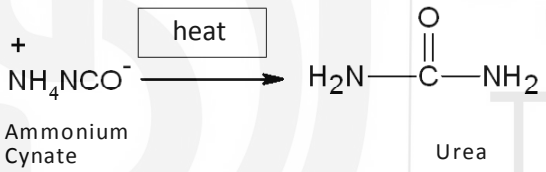
1.8 ANSWERS

Self-Assessment Questions

1. i) Carl Neuberg
   ii) Friedrich Wöhler
   iii) Hargobind Khorana
   iv) Drugs
   v) Citric acid cycle
2. i) True
   ii) False
   iii) True
   iv) False

Terminal Questions

1. In the year 1828 Wöhler synthesised urea by heating ammonium cyanate. This was the first noted reaction that proved that synthesis of organic molecules is possible from inorganic substances.



2. The core objective of biochemistry is to understand the chemical basis of life. To achieve this objective biochemistry crosstalk with many of the allied sciences like Chemistry, Biology, Physics and also Computer Science. Please see section 1.3 for complete details.
3. Every living organism shows its existence by performing certain essential activities like growth, reproduction etc. However, life is controlled by maintaining equilibrium between synthesis and degradation of biomolecules. These biomolecules, produced by chemical elements like C, N, S, P etc. Please see section 1.4 for more details.
4. Carbohydrates, proteins, lipids and nucleic acids are considered as major biomolecules. They have a significant role in maintaining life. Please refer Table 1.2 for complete details.

1.9 FURTHER READINGS

1. Lehninger: Principles of Biochemistry, David L. Nevron and Michael M. Cox. W.H. Freeman, 4th Edition, 2004. New York.
2. Harper's Illustrated Biochemistry, 29e. Robert K. Murray, David A Bender, Kathleen M. Botham, Peter J. Kennelly, Victor W. Rodwell, P. Anthony. Weil, USA.

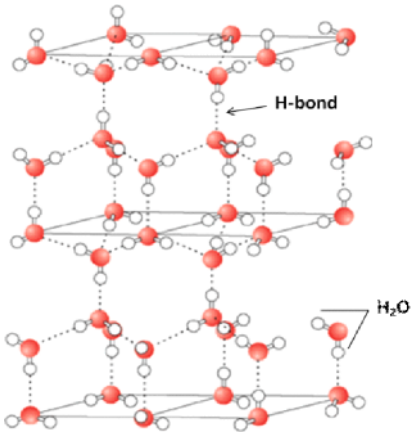
3. J. L. Jain: Fundamentals of Biochemistry, S. Chand & Company Ltd., India

4. U. Satyanarayana and U. Chakrapani: Biochemistry, UBS Publishers Distributors Pvt. Ltd., Kolkata, India.

5. Pearson Global Edition - Appling/Anthony - Cahill/Mathews - Biochemistry: Concepts and Connections. Edinburgh Gate, Harlow, Essex CM20 2IE, England, 2015.

6. Experimental Biochemistry: A Student Companion. Beedu Sasidhar Rao and Vijay Deshpande. I.K. International Pvt. Ltd. New Delhi.

7. Practical Biochemistry: For Medical, Dental and Allied Courses. 2nd edition. Dr. G. Rajagopal and Dr. B. D. Toora. Ahuja Publishing House, New Delhi.



UNIT 2

WATER

Structure

2.1 INTRODUCTION

In Unit 1, you have learned about foundations of Biochemistry. You know living organisms are composed of water, organic molecules and inorganic ions. Water is the most abundant molecule in living organisms. It is essential for life. Life first arose in an aqueous environment. Water makes up 70% or more of the body weight of all living organisms. The unique properties of this ubiquitous substance have a profound influence on the chemistry of life. However, water acts as an aqueous medium in which living organisms carry out their biological functions such as metabolism. It can dissolve or dissociate most compounds and is considered as a universal solvent. Water acts as a reactant and is also a product in many metabolic reactions. Therefore, we can say that existence of living beings would not have been possible without water.

This unit begins with a brief discussion on noncovalent interactions in aqueous systems (Sec. 2.2). Here, you will learn about structure of water, hydrogen bond, hydrophilic, hydrophobic, Vander Walls and electrostatic interactions. The Sec. 2.3 will discuss the unique properties of water. You will learn about ionization of water and the concept of pH in Sec 2.4. The Sec. 2.5 will explain you about buffers and Henderson-Hasselbalch equation. At the end of this Unit–water as a reactant and fitness of the aqueous environment will be discussed.

Objectives

After studying this unit, you should be able to:

- describe the noncovalent interactions *i.e*. hydrogen bond, hydrophilic and hydrophobic, Vander Walls and electrostatic interactions;
- explain the unique properties of water;
- explain ionization of water and define pH;
- define buffers and derive the Henderson Hasselbalch equation; and
- justify water as a reactant and fitness of the aqueous environment.

2.2 NONCOVALENT INTERACTIONS IN AQUEOUS SYSTEMS

You might have studied about water in your school chemistry courses. A molecule of water consists of two hydrogen atoms and one oxygen atom. It is represented by chemical formula H_2O. However, it constitutes an aqueous medium in which almost all biological processes take place. The solubility of a substance in water depends upon the nature of forces of the functional groups of the substance with water. The unusual properties of water make it an ideal medium for living organisms. Therefore, water is considered as a universal solvent as it dissolves many substances.

Water is the principal component of biological systems. Most functions of biomolecules depend upon aqueous medium as they interact with water molecules by weak interactions. Weak interactions, called noncovalent interactions include hydrogen bond, hydrophilic, hydrophobic, vander Waals and electrostatic interactions. These interactions are weak but collectively show significant effects in biological systems. These are involved in maintaining the structure of proteins, nucleic acids, polysaccharides and membrane lipids. Therefore, you will learn about noncovalent interactions involved in aqueous systems such as:

Hydrogen bond

Hydrophilic and Hydrophobic interactions,

Vander Waals interactions, and

Electrostatic interactions.

Let us study these interactions one by one.

2.2.1 Hydrogen Bond

Hydrogen bond is very common in biological systems. It plays a significant role in biochemical reactions as well as formation of biological structure. **Therefore, let us first understand the molecular structure of a water molecule which reveals the cause of the hydrogen bonding**.

Look at Fig. 2.1 (a) it shows the molecular geometry of a water molecule in which two hydrogen atoms are covalently bonded to one oxygen (O) atom and

each hydrogen (H) atom are at the two corners (two bonding electron pairs) and the two nonbonding electron pairs (unshared electron pairs) of oxygen atom are at the other two corners. This geometry looks like a tetrahedron. These nonbonding electron pairs repel each other. As a result, these repulsive forces act to push the hydrogens closer together. Thus, the H-O-H bond angle decreases to 104.5° as compared to 109.5° for a perfect tetrahedron (see Fig. 2.1b). Hence, the structure of water molecule acquires a bent angular geometry or V-shape structure.

In a water molecule, the oxygen atom is more electronegative than the hydrogen atom; therefore, the oxygen nucleus attracts the shared pair of electrons than the hydrogen nucleus. The unequal sharing of electrons results in **two electric dipoles** in water molecule, (one along each O-H bond). The oxygen atom bears a (δ^-) partial negative charge and each hydrogen atom bears (δ^+) partial positive charge. The presence of opposite charges in a water molecule allow it to form interactions with other water molecules. This interaction is specifically called the **hydrogen bond**. Hydrogen bond is an electrostatic interaction between partial negatively charged oxygen atom of one water molecule and partial positively charged hydrogen atom of another. Hydrogen bond with the three parallel lines as shown is Fig. 2.1(c). The distance between the participating atoms for hydrogen bonding is around 0.17 nm and distance of a covalent bond (H-O) is approximately 0.096 nm.

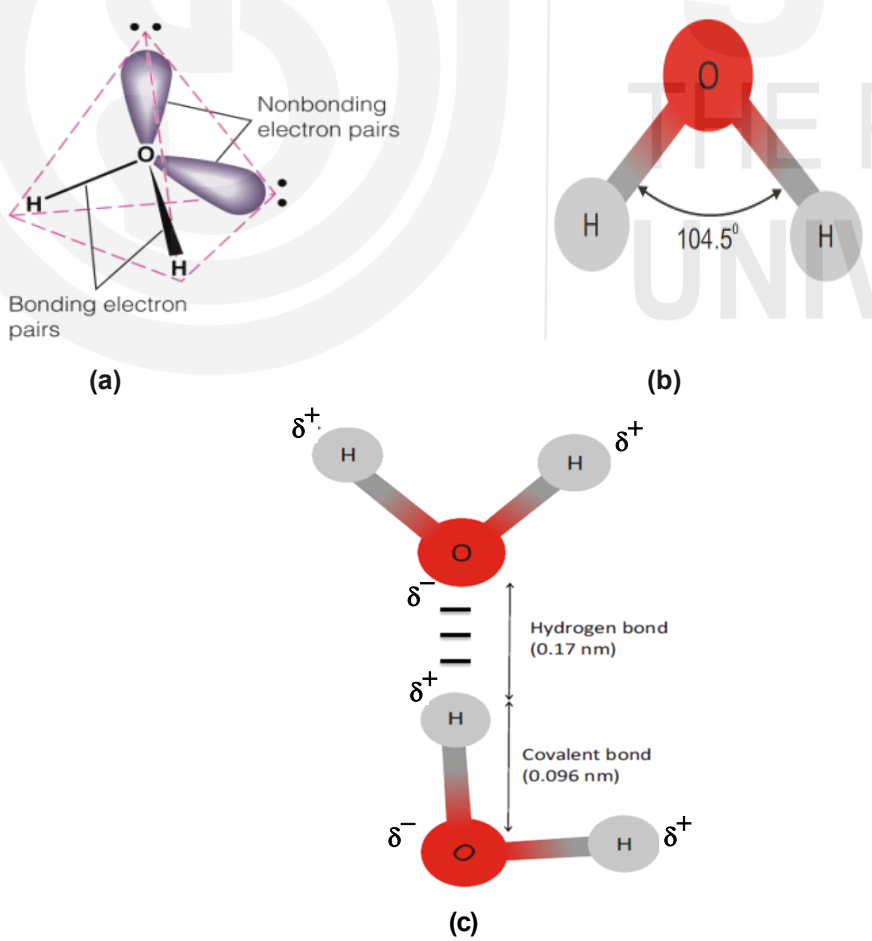


Fig.2.1: (a) Tetrahedral geometry of a water molecule; (b) H-O-H bond angle (104.5°); and (c) Hydrogen bonding between two water molecules and the symbol delta (δ) shows a partial weaker charges on a water molecule.

However, each water molecule can theoretically form H bonds with four nearby water molecules. This feature of water can be viewed in ice formation. Look at Fig. 2.2 showing the network of hydrogen bonding in the crystal structure of ice (solid state). Hydrogen bonds are weak and very unstable. They keep forming and breaking at all times and so most water molecules in liquid state are always hydrogen bonded. Hydrogen bonds provide the stability to water molecules which are accounting for unusual properties of water. Extensive hydrogen bonding is responsible for water being a liquid at room temperature.

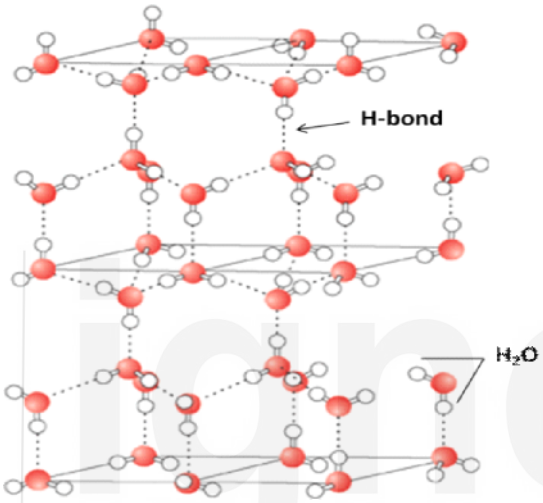
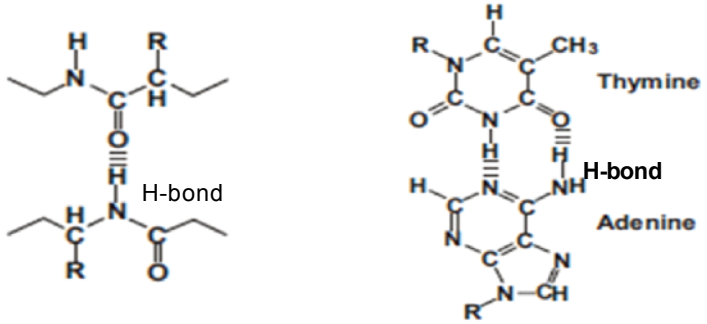


Fig. 2.2: Hydrogen bonding in ice.

However, hydrogen bonding is not unique to water. Water molecules also readily form hydrogen bonds with other kind of biological molecules such as proteins and nucleic acids. Besides, hydrogen bonding is frequently encountered between water and biomolecules and within biomolecules. The electronegative atoms, usually oxygen (O) and nitrogen (N) are involved in forming hydrogen bonds. These atoms primarily occur in amino group ($-NH_2$) and hydroxyl group ($-OH$). In biological systems, the presence of amino and hydroxyl groups in polypeptides (proteins) and nucleic acids (DNA and RNA) allows to form hydrogen bonding. Look at Fig. 2.3 showing biologically important hydrogen bonds within polypeptide chains as well as de-oxy ribonucleic acids (DNA). Therefore, hydrogen bonding is crucial for the formation of three-dimensional structures of proteins, nucleic acids and other biological molecules.



(a) Hydrogen bonding between two polypeptide chains **(b) Hydrogen bonding between nitrogenous bases of two stands of DNA**

Fig. 2.3: Hydrogen bonding between biological molecules.

2.2.2 Hydrophilic and Hydrophobic Interactions

Hydrophilic means 'water loving' and hydrophobic means 'afraid of water'. Hydrophilic interactions are found between two polar or charged molecules and they interact by forming hydrogen bonds. Hydrophilic molecules (polar) such as glucose, amino acids etc. that experience hydrophilic interactions and readily dissolve in water. On the other hand hydrophobic, (non-polar) molecules such as lipids, oils tend to aggregate together via hydrophobic interactions in order to avoid contact with water. A non-polar molecule is virtually insoluble in water and cannot form hydrogen bonds with water. Thus, they have tendency to self-associate in an aqueous environment. However, there are some compounds in biological systems which have both 'polar groups' and 'non-polar groups' are called amphipathic molecules (*amphi* meaning ''both'' and *philos* meaning ''loving'').

Amphipathic molecules exhibit both hydrophobic and hydrophilic interactions which play an important role in the formation of biological membranes. For example, phospholipid molecules are amphipathic in nature which contain polar and charged head groups with phosphate and a hydrophobic group as a non-polar tail. You can see in Fig. 2.4 how these molecules are arranged such that their polar hydrophilic regions (shown in blue color) interact with water molecule while non-polar hydrophobic regions (shown in red color) interact with each other with minimum exposure to water molecules. As a result, amphipathic molecules aggregate themselves in an aqueous solution and form micelle (spherical in shape). This shape involves in the formation of cell membrane.

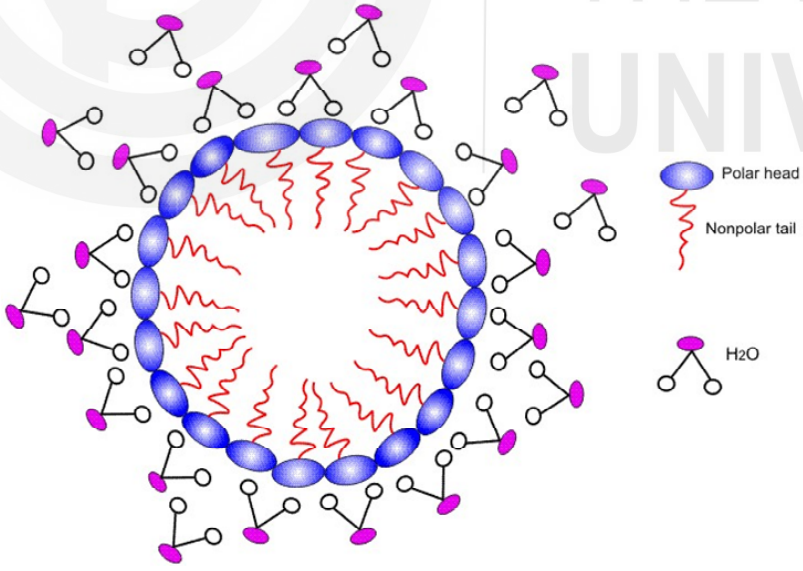


Fig. 2.4: Micelle formation

2.2.3 Vander Waals Interactions

Vander Waals Interactions are the weakest of all intermolecular attractions and arise due to attraction of the nuclei and electron clouds between atoms or molecules. As you know, electrons are negatively charged and nucleus bears positive charge, so when two uncharged atoms come close together, the

nucleus of one atom attracts the electron cloud of the other, and vice versa. As a result, the transient dipoles develop in them. Such dipoles are called **induced dipoles** and the attractions between them are called **Vander Waals interactions** as shown in Fig. 2.5. These forces require the two atoms to be present at a characteristic optimal distance called vander Waals radius**.** This interaction contributes in protein interactions with other molecules such as enzyme-substrate interaction, antibody-antigen interaction as well as protein folding.

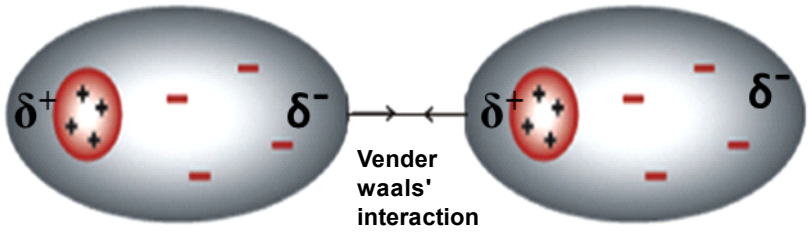


Fig. 2.5 : Vander Waals interactions. (The delta sign δ^+ and δ^- show weaker positive charge on nuclei cloud of one atom and a weaker negative charge on electronic cloud of another atom respectively).

2.2.4 Electrostatic Interactions

These interactions occur between polar or ionic compounds for example sodium chloride (NaCl) ions which are held together by strong electrostatic forces. As you know water is a polar molecule, it readily dissolves charged or polar molecules. When crystal form of NaCl (salt) dissolves in water, it dissociates into the sodium ions (Na^+) and the chloride ions (Cl^-). These opposite ions are surrounded by water molecules. Fig. 2.6 (a) shows that negative end (δ^-) of water molecules attracts the sodium ions and the positive end (δ^+) of water molecules attracts the chloride ions. Since water has a high dielectric constant, it dissolves ionic compounds such as NaCl by effectively weakening the electrostatic interactions between them. These ions then get hydrated or solvated by a large number of water molecules which surround them. In biological systems, simple inorganic ions such as sodium (Na^+), potassium (K^+), calcium (Ca^{2+}), magnesium (Mg^{2+}), and chloride (Cl^-), do not exist as free and each ion is bound by electrostatic interaction between the ion and the oppositely charged end of the water dipole. Fig 2.6 (b) shows the electrostatic interaction between a K^+ ion and a water molecule (H_2O).

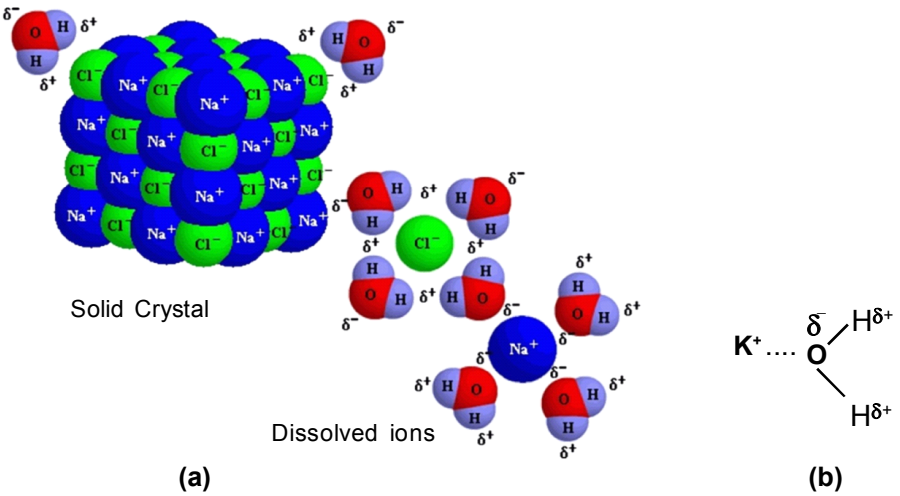


Fig. 2.6 : (a) Electrostatic interactions between NaCl and water molecules.
(b) Electrostatic interaction between K^+ ion and a water molecule.

SAQ 1

a) **Fill in the blanks with suitable words.**

i) The oxygen atom in water molecule has .....unshared electron pairs.

ii) The bent shape of a water molecule is due to ..... between nonbonding electron pairs of oxygen atom.

iii) The oxygen atom in water molecule acquires partial negative charge and hydrogen atom a partial positive charge because of high ..... of oxygen atom.

iv) A water molecule has permanent electric ....because of partial positive charge on hydrogen atoms and partial negative charge on oxygen atom.

v) In a DNA double helix the hydrogen bonds are formed between hydrogen and oxygen and between hydrogen and .....of two bases.

b) **Choose the correct answer given in the brackets.**

i) Hydrogen bonds are (weaker/stronger) than covalent bonds.

ii) A water molecule in ice can form maximum of (3/4) hydrogen bonds with neighboring molecules.

iii) Hydrogen bonds are (unstable/permanent).

c) **Fill in the blanks with suitable words:**

i) Lipid molecules have tendency to..... in an aqueous environment.

ii) The inter molecular force of attraction between nuclei and electron cloud of two atoms, results in..... interactions.

iii) Molecules that have both polar and non-polar groups are called ..... molecules.

d) **Choose the correct answer given in the brackets.**

i) Water dissolves salt by effectively (weakening/strengthening) electrostatic interaction between ions.

ii) In amphipathic molecules, (polar/non-polar) groups aggregate so that they are (in contact with/avoid of) water.

2.3 UNIQUE PROPERTIES OF WATER

The unique properties of water are due to hydrogen bonds present in it. We will now learn some of the physical properties of water and relate them to its usefulness for supporting life.

Table 2.1: Physical properties of Water

Property	Constants with units
Melting point	0°C
Boiling point	100°C
Heat of vaporization	2260 J/g
Dielectric constant	78.5 at 25°C
Density (maximum)	999.97 kg/m^3 at 4°C
Specific heat capacity	1

Hydrogen bonding between water molecules can explain the unusual nature of many physical properties listed in Table 2.1. Let us discuss them one by one:

Melting point: Unless a significant proportion of H bonds are broken the crystal structure of ice cannot be destabilised. The thermal energy required to do so accounts for the melting point. Melting of ice is a spontaneous process.

Water is expected to have a high boiling point. This is supported by the high heat of vaporisation that provides a direct measure of the energy required to overcome the attractive forces between water molecules. Many higher organisms lose excess body heat by evaporation of sweat.

The high dielectric constant of water makes it an effective solvent for dissolving polar (such as sugars) and charged molecules (amino acids such as lysine). It dissolves these compounds by reducing their electrostatic interactions and form hydrogen bonds. On the other hand, non-polar substances tend to cluster together in the presence of water.

Water has the maximum density at 4°C. When it freezes, it is less dense than liquid water. This is advantageous for the survival of aquatic organisms. When water freezes in ponds, ice floats on the surface and life still continues in the liquid water underneath.

Water has a high specific heat. It needs to absorb 4.814 Joules of heat in order to raise the temperature of 1g of H_2O by one degree Celsius (1°C). This is beneficial for organisms to withstand temperature fluctuations.

SAQ 2

Fill in the blanks with suitable words:

i) Ice floats on surface of water because it is less ..... than liquid water.

ii) Water is a good solvent for polar or ionic compounds due to its …….

iii) 1 gram of water needs 4.814 joules of heat to raise the temperature by 1^{O}C is due to .....……

iv) Unique properties of water are due to ..... ……

2.4 IONIZATION OF WATER

You might have learnt about ionization of water in your 10+2 chemistry courses. You know that water molecules ionize to a limited extent. This ionization or dissociation produces H^+ (proton) and OH^- (hydroxy) ions which is given below:

$$H_2O \rightleftharpoons H^+ + OH^- \qquad \text{Eq. (1.1)}$$

$$H_2O + H^+ \rightleftharpoons H_3O^+$$

In reality protons (H^+) are not ''free'' in solution and they immediately associate with another water molecule to form hydronium ions (H_3O^+). The hydronium ions are routinely represented as H^+. So we can write H^+ in place of H_3O^+ .

The equilibrium constant (K_{eq}) for Eq. 1.1 is represented as:

$$K_{eq} = \frac{[H^+], [OH^-]}{[H_2O]} \qquad \text{Eq. (1.2)}$$

K_{eq} = equilibrium constant for dissociation of water. From electrical conductivity measurement of water, the value of K_{eq} is 1.8×10^{-16} M (moles L^{-1}) at 25°C.

$[H^+]$, $[OH^-]$ and $[H_2O]$ in brackets represent molar concentration.

Molar: The term molar or molarity is a unit used to express the concentration of a substance dissolved in a solution. It is denoted by M. Therefore, 1 M is gram molecular weight of a substance per litre of solution (g mole $^{-1}$)

Let us now define another useful constant called the ionic product of water, it is represented by K_w. The molecular weight of water is 18 g mole^{-1}. Therefore, the concentration of water at 25°C is grams of H_2O in 1 Litre (L) is divided by gram molecular weight, *i.e.*,1000/18 = 55.5 moles L^{-1}. Thus, the molar concentration of pure H_2O is 55.5 M (molar).

Substituting the values for K_{eq} and molar concentration of H_2O in the Eq. (1.2)

$$1.8 \times 10^{-16} = \frac{[H^+]\,[OH^-]}{55.5}$$

Rearranging, we gets

$$[H^+]\,[OH^-] = (55.5)\,(1.8 \times 10^{-16})$$

To simplify,

$$[H^+]\,[OH^-] = 1.0 \times 10^{-14}$$

The ionic product $[H^+]\,[OH^-]$ of water is represented by K_w, thus

$$K_w = [H^+]\,[OH^-] = 1.0 \times 10^{-14} \text{ at } 25\ ^\circ\text{C} \qquad \text{Eq. (1.3)}$$

According to the Eq. 1.3, the product of hydrogen ion concentration and hydroxyl ion concentration is the constant value of 1.0×10^{-14} at 25°C. The value of K_w can be used to calculate the concentration of H^+ and OH^- of water.

$$K_w = [H^+]\ [OH^-] = 1.0 \times 10^{-14}$$

$$[H^+]^2 = 1.0 \times 10^{-14}$$

or

Solving for $[H^+]$, we get

$$[H^+] = 1.0 \times 10^{-7} \qquad \text{Eq. (1.4)}$$

or $\quad [H^+] = [OH^-] = 1 \times 10^{-7}$ Moles/L

The Eq. 1.4 states that the concentrations of H^+ and OH^- of pure water at 25°C are equal (10^{-7} Moles/L). When concentrations of $[H^+]$ and $[OH^-]$ are equal in any given solution, as in pure water, the solution will be neutral. In acidic solution the concentration of H^+ ions is more than the concentration of OH^- ions. Similarly, alkaline solution will have higher concentration of OH^- ions than the H^+ ions. As an example let us calculate hydrogen ion concentration in a given solution using the ionic product of water (K_w).

Example 1: (a) What is the concentration of H^- ions in solution of 0.01 M NaOH.

$$K_w = [H^+]\ [OH^-] = 1.0 \times 10^{-14}\ M^2$$

Solving for $[H^-]$

$$[H^+] = \frac{K_w}{[OH^-]} = \frac{1.0 \times 10^{-14}\ M^2}{0.01\ M} = \frac{10^{-14}\ M^2}{10^{-2}\ M}$$

To simplify, the concentration of H^- ions will be 1×10^{-12} M in a given solution.

(b) Calculate the concentration of OH^- in a solution of 0.1M HCl.

$$K_w = [H^+]\ [OH^-]$$

Solving for $[OH^-]$

$$[OH^-] = \frac{K_w}{[H^+]} = \frac{1.0 \times 10^{-14}\ M^2}{0.1\ M} = \frac{10^{-14}\ M^2}{10^{-1}\ M}$$

To simplify, the concentration of OH^- ions will be 1×10^{-13} M in a given solution.

2.4.1 pH

Recall from your school chemistry books the concept of pH. The ionization products of water have an important role to define the pH of any solution. The

pH is defined as the negative logarithm of the hydrogen ion (concentration in moles/ litre) and can be represented by the following expression:

$$pH = -\log [H^+] = \log \frac{1}{[H^+]}$$

where $[H^+]$ is hydrogen ion concentration in mol/L.

In 1909, **S.P.L Sorenson** introduced the concept of pH to express hydrogen ion concentration of aqueous solutions.

You know that pure water contains an equal amount of hydrogen ions and hydroxyl ions,

$$[H^+] = [OH^-] = 1 \times 10^{-7} \text{ M at } 25°C.$$

The pH of pure water at 25°C can thus be calculated as follows,

$$pH = \log \frac{1}{1.0 \times 10^{-7}} = \log (1.0 \times 10^7)$$

$$= \log 1.0 + \log 10^7 \qquad (\log 1 = 0)$$

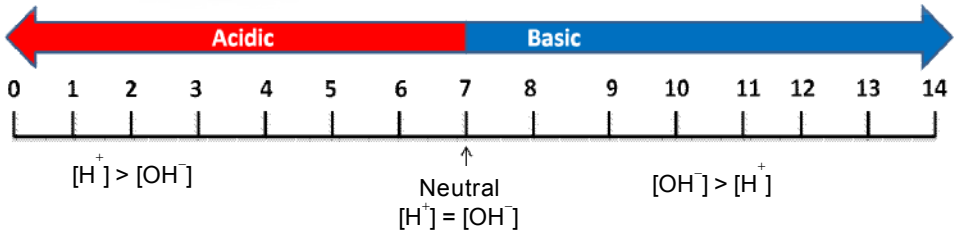
$$= 0 + 7$$

pH = 7

Therefore, pH of pure water is 7.0 at 25°C. Sometimes, pOH is used to describe basicity or OH^- concentration

$$pOH = -\log [OH^-]$$

Therefore, the acidic or basic nature of a solution is experssed in terms of **pH scale**, ranges from 0 to 14 as shown in Fig. 2.7.



pH of 7 is a neutral

pH < 7 is an acidic

pH > 7 is a basic

Fig. 2.7: pH scale.

On pH scale, a pH of 7.0 represents neutral solution in which the number of hydrogen ions is balanced by the same number of hydroxyl ions. The pH of an acidic solution is less than 7 and that of a basic or alkaline solution is more than 7. **There is an inverse relation between pH value and hydrogen ion**

concentration. The lower pH value corresponds to higher H^+ concentration $[H^+]>[OH^-]$ and the higher pH value corresponds to higher concentration of hydroxyl ions $[OH^-]>[H^+]$.

However, most of the biochemical reactions are sensitive to pH. The biomolecules such as proteins and nucleic acids carry functional groups that will be charged or neutral depending on the pH.

The changes in pH may influence the catalytic activity of enzymes as well as functioning of the cells. Table 2.2 enlists the pH values of some cellular organelles and body fluids of living system.

Table 2.2: pH values of some important organelles and body fluids

Cellular Organelles and body fluids	pH
Mitochondria matrix	7.5
Cytosol	7.2
Nucleus	7.2
Blood	7.4
Peroxisomes	7.0
Urine	6.0
Saliva	6.3-6.8
Human skin	5.5
Lysosomes	4.5
Gastric juice	1

SAQ 3

Fill in the blanks with correct words:

i) Hydronium ions are routinely represented as ..... ions.

ii) Lower pH value indicates .....concentration of H^+ ions.

iii) The value of Kw of water is.....

iv) The molar concentration of pure water is .....

v) pH of a solution indicates the acidic or.....nature of solution.

vi) pH scale ranges from ..... to ..... .

2.5 BUFFERS

A buffer is an aqueous solution that resists a change in pH when small quantities of an acid or a base are added to it. The property of resistance is called buffer action. Most commonly used buffers are made up of a conjugate acid-base pair. Acidic buffer is a mixture of a weak acid and its conjugate base e.g. acetate buffer system. Alkaline buffer contains a weak base and its salt. If

hydrogen ions are added to the buffer solution, they are neutralized by the base and similarly on addition of hydroxyl ions they will be neutralized by the acid. As a result of these neutralization reactions, the pH of buffer solution, remains unchanged.

In general, an acid is a proton donor and a base is a proton acceptor. An acid and its corresponding base are called conjugate acid-base pair. For example, acetic acid is a weak acid, which dissociates into a proton and acetate ions and vice versa in an aqueous solution as follows.

$$\underset{\text{HAc}}{CH_3COOH} \rightleftharpoons \underset{\text{Ac}^-}{H^+ + CH_3COO^-}$$

To simplify, acetic acid has been represented by HAc while acetate has been represented by Ac^-. This is a reversible reaction where acetic acid acts as a proton donor and the acetate acts as a proton acceptor.

The ionization of acetic acid is designated by dissociation constant (*K*a), thus

$$Ka = [H^+] [Ac^-] / [HAc]$$

This conjugate acetic acid-acetate pair acts as a buffer system if we add small amounts of either acid or bases to this buffer solution. The acid and base are themselves ionized and releases protons and hydroxyl ions respectively in the buffer solution. Look at Fig. 2.8 to understand how this buffer system works.

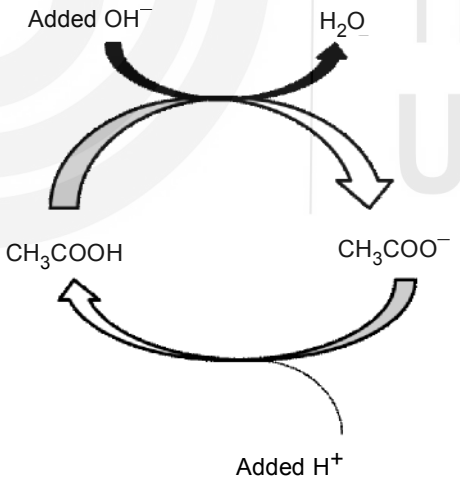


Fig. 2.8: Acetic acid-acetate buffer system.

If small amount of **H^+** added to the buffer solution, it will react with acetate ion and form acetic acid.

$$CH_3COO^- + H^+ \rightleftharpoons CH_3COOH$$

Likewise, if small amounts of **OH^-** added to the buffer solution, H^+ of acetic acid will react with the hydroxyl ion and form H_2O and acetate ion.

$$CH_3COOH + OH^- \rightleftharpoons CH_3COO^- + H_2O$$

In the above reaction, water is formed and no free H^+ or OH^- ions are available. Thus, the acetic acid-acetate buffer system acts as a conjugate acid-base pair and tends to resist change in pH if little amount of acid or base is added to it. Thus, a buffer system is capable of absorbing the free H^+ or OH^-.

The buffering capacity of any buffer system depends on two reversible reaction equilibria occurring in a solution of nearly equal concentrations of a proton donor and its conjugate proton acceptor. The final outcome is a small change in both the pH and the ratio of conjugate acid and its base.

Buffers play important roles in maintaining the pH of cells and cellular organelles to carry out their biological processes. The important biological buffers in the human body are phosphate buffer and bicarbonate buffer. The phosphate buffer is an intracellular buffer in the cells which is maximally effective at pH 6.86 and it has buffering range from 5.9 to 7.9. The bicarbonate buffer is a physiological buffer system in human blood which plays a crucial role to maintain the blood at pH 7.4.

Let us learn about the equation which has significant role for buffers.

2.5.1 Henderson-Hasselbalch Equation

Buffer system is effective at a narrow pH range. In general, the buffering region is one unit on either side of pKa. When pH = pKa, the ratio of conjugate acid to conjugate base is equal and at this pH the buffering power is maximal.

Henderson-Hasselbalch equation is important equation to express the relationship between pH, pK_a and conjugate acid-base pair in a buffer system. Let us derive this equation:

Consider a weak acid (HA) that ionizes into a proton (H^+) and a base (A^-) in an aqueous solution:

$$HA \rightleftharpoons H^+ + A^-$$

The equilibrium constant for this ionization is often represented by dissociation constant (*Ka*), thus

$$K_a = \frac{[H^+]\ [A^-]}{[HA]} \qquad \text{Eq. (i)}$$

p*K*$_a$: Is the negative logarithm of proton dissociation constant, and it is equal to the pH, at which half of the acid has dissociated.

Solving $[H^+]$ in the Eq. (i),

$$\frac{1}{[H^+]} = \frac{1}{[Ka]} + \frac{[A^-]}{[HA]} \qquad \text{Eq. (ii)}$$

Taking the logarithm at both sides in the Eq. (ii),

$$\log \frac{1}{[H^+]} = \log \frac{1}{[Ka]} + \log \frac{[A^-]}{[HA]}$$

by substitution of **log 1/[H$^+$] = pH** and **log 1/ [*Ka*] = p*K*a, thus** we obtain

$$pH = pK_a + \log \frac{[A^-]}{[HA]} \qquad \text{Eq. (iii)}$$

In general it can be represented as:

$$pH = pK_a + \log \frac{[Base]}{[Acid]}$$

The Eq. (iii) is known as the **Henderson-Hasselbalch equation,** this used to calculate the pH of a buffer solution and the ratio of conjugate base and conjugate acid at a given pH.

SAQ 4

Fill in the blanks with correct words:

i) Buffers are made up of ……

ii) The ability of a buffer solution to resist the pH change in a solution is known as..... ……..

iii) Acids are ……and bases are ……

iv) Bicarbonate buffer system maintains the pH at ..... in human blood.

v) Henderson-Hasselbalch equation defines the relationship between .....and pK_a of a .....

vi) The ratio of conjugate acid and conjugate base in a solution can be determined by .....

2.6 WATER AS A REACTANT AND FITNESS OF THE AQUEOUS ENVIRONMENT

So far you have learned about the biological importance of water, pH and buffers in maintaining life. In this section, you will study about water as a reactant and the fitness of the aqueous environment.

1. Water as a reactant: Water not only acts as a solvent b is often actively participates as a reactant in various biochemical reactions. These reactions include condensation, hydrolysis, and oxidation-reduction reactions. According to biological needs water acts as a proton donor and proton acceptor. This ability of water is very significant in maintaining life.

Let us discuss the ability of water as a reactant with following examples:

(i) Fig. 2.9 shows the formation of adenosine tri phosphate (ATP) from adenosine di phosphate and inorganic phosphates (P_i) is an example of condensation reaction in which a molecule of water is released. The reverse of this reaction, ATP is break down into ADP and Pi by the addition of water, which is a hydrolytic reaction.

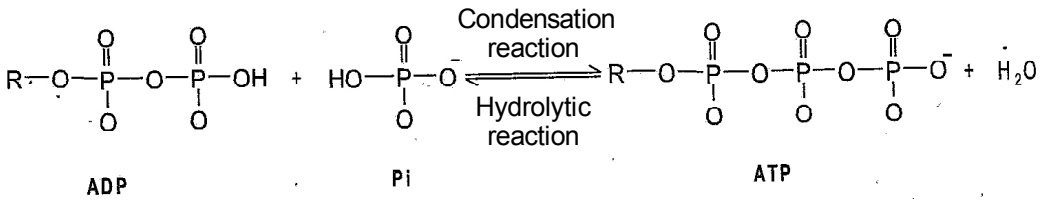


Fig. 2.9: ATP formation and its breakdown (R = Adenosine).

(ii) In animals, water (H_2O) and Carbon dioxide (CO_2) are the end products of oxidation of glucose and fatty acids. The dissolved CO_2 in blood reacts with water to form bicarbonic acid (H_2CO_3). Where water is not only acting as a substrate but also serves as a proton donor which helps in maintaining the blood pH at 7.4. The reversisble of this reaction is catalyzed by the enzyme carbonic anhydrase.

$$CO_2(d) + H_2O \underset{}{\overset{\text{Carbonic anhydrase}}{\rightleftharpoons}} H_2CO_3$$

Where (d) denotes dissolved CO_2 in Blood

(iii) In plants, water donates electron (e^-) to chlorophyll and itself undergoes ioxidizes to produce molecular oxygen during photosynthesis.

$$2H_2O + 2A \xrightarrow{\text{light}} O_2 + 2AH_2$$

In above reaction, "A" is an electron-accepting species and water serves as electron donor.

Inside cell extended networks of H bonded water molecules with other biomolecules exist. This will allow interactions to occur over distances. This network has been used by plants to transport dissolved nutrients during transpiration from roots to the leaves. Water molecules are tightly bound to certain proteins such as cytochromes-f by hydrogen bonding and they assist in the movement of protons through the membrane.

2. Water as a fitness of the aqueous environment: Most living organisms require water more than any other substance. Water is the only essential substance that supports life on the earth as well as in the aqueous environment.

Recall the unique properties of water such as high specific heat and high heat of vaporization which are dissussed in the section 2.3.

In general, these two properties are very significant in maintaining aquatic life, specifically in maintaining relatively constant body temperatures of aquafic organism irrespective of their surrounding environment.

Have you ever noticed that small organisms float or walk on water surface? The floating on surface of water is due to the high surface tension of water. The high degree of internal cohesion of liquid water, due to hydrogen bonding.

Water has low density on freezing. It has maximum density at 4°C and it becomes less dense below 4°C. Water supports life at very low temperatures in aqueous environment because the density of ice, is less than that of liquid water. This is an important natural phenomena that occurs in the aquatic ecosystem. The floating layer of ice on the top effectively insulates the water below it, allowing most aquatic organisms survive in extreme winter.

Therefore, we can say that the unique properties of water have a determinative role in the process of biological evolution.

SAQ 5

Fill in the blanks with correct words:

i) Water not only acts as a …….. but it also serve as a ……..

ii) Water serves as a…….. and a .....

iii) In plant photosynthesis, water molecule under goes oxidotion to release .....

iv) The unique properties of water such as ……. are crucial to support aquatic life in an aqueous environment.

v) Water has a high ……which allow small organism to float on ……

vi) Water has lower and maximum density at ….°C.

2.7 SUMMARY

In this Unit you have studied:

Water is the main constituent of biological systems. It is a polar molecule due to the presence of opposite charges. The presence of opposite charges on each water molecule allows formation of hydrogen bonding. Hydrogen bonding property of water makes it a unique molecule. Hydrogen bonding in water accounting for several physical properties such as melting point, high boiling point, and high dielectric constant. Hydrogen bonding property of water determines solubility of many substances. Due to polar nature of water, it forms hydrogen bonds with other polar or charged molecules. Noncovalent interactions are crucial for the formation of biological structures specially proteins and DNA.

The ionic products of water [H^+] and [OH^-] define the acidic and basic nature of a substance in an aqueous system. Weak acid-conjugate base pairs act as a buffer system. Henderson-Hasselbalch equation is useful for determining the quantitative relationships between pH, pK_a and the ratio of acids and bases in buffer systems. Biological buffers are essential in maintaining constant pH in the human body. Important biological buffers such as phosphate and bicarbonate buffers which are vital to maintain pH intracellula and in blood respectivley.

Water molecules participate in various metabolic reactions where they serve as proton donor and proton acceptor. Owing to the physical properties of water it acts as a vital molecule in maintaining life both terrestial and aquatic environments.

2.8 TERMINAL QUESTIONS

1. Describe the structure of water.
2. Explain polar nature of water molecule.
3. Define the terms hydrophilic, hydrophobic and amphipathic.
4. What do you understand by the term buffer?
5. Explain the Henderson-Hasselbalch equation and its use.

2.9 ANSWERS

Self-Assessment Questions

1. (a) (i) Two (ii) electrostatic repulsion
   (iii) electronegativity (iv) dipole moment
   (v) nitrogen
   (b) (i) weaker (ii) 4 (iii) unstable
   (c) (i) aggregate (ii) Vander Waals (iii) amphipathic
   (d (i) weakening (ii) non-polar, avoid of
2. (i) dense
   (ii) polar nature
   (iii) high specific heat
   (iv) Hydrogen bonding
3. (i) H_3O^+ (ii) high (iii) 1×10^{-14} M^2.
   (iv) 55.5 moles/ liter or 55.5 M. (v) basic (vi) 0 to 14.
4. (i) weak acids or bases and their conjugates.
   (ii) buffer action.
   (iii) proton donors, proton acceptors.
   (iv) 7.4
   (v) pH, conjugate acid-base pair.
   (vi) Henderson-Hasselbalch equation.

5. (i) universal solvent, reactant.

   (ii) proton donor, proton acceptor.

   (iii) oxygen (O_2)

   (iv) high specific heat, heat of vaporization and low density.

   (v) surface tension.

   (vi) 4°C.

Terminal Questions

1. Water has bent shape structure due to presence of two nonbonding electron pairs (unshared/lone pairs) on oxygen atom. These electrons tend to arrange themselves symmetrically at the vertices of a regular tetrahedron around the oxygen atom. Due to the electronic repulsion, the bond angle of H-O-H decreases to 104.5^0 as compared to 109.5° of a perfect tetrahedron.

2. In a water molecule, the oxygen atom has more electronegativity as compared to hydrogen atoms resulting in electronegativity difference. Therefore, oxygen atom tends to withdraw the shared electrons towards itself of each hydrogen atom. As a result, oxygen atom bears a partial negative charge (δ^-) and each hydrogen atom bears partial positive charge (δ^+). The presence of opposite charges on a water molecule is said to be a permanent polar molecule. These opposite charge on a water molecule allow it to form hydrogen bonding between water molecules. Being a polar molecule, water dissolves many polar, uncharged, ionic and charged molecules by forming hydrogen bonding with them. Hence, it is considered a biological solvent.

3. **Hydrophilic** molecules form hydrogen bonding with water molecules. Water readily dissolves polar and ionic molecules such as sugar (glucose), amino acids and sodium chloride etc.

   Hydrophobic molecules do not make hydrogen bonding with water molecules. These molecules such as lipids and hydrocarbons are nonpolar. Such molecules avoid interacting with water molecules and aggregate in aqueous solution.

   Amplipathic Molecules are having both the hydrophilic and hydrophobic groups are called amphipathic molecules.

4. A buffer system comprises of conjugate acid-base pair of a weak acid or a weak base. This system is able to resist a change in pH following addition of small quantities of strong acid or base.

5. Henderson-Hasselbalch equation is useful to expres the quantitative relationship between pH and pK_a of the weak acid and conjugate acid-base pair. The following forms of equation have several applications as:

   $pH = pK_a + \log [base/acid]$ (useful to calculate the pH of the buffer solution)

   $pH - pK_a = \log [base/acid]$ (useful to calculate the ratio of concentration of base to acid)

   $pK_a = pH + \log acid/base$ (useful to calculate the pK_a of weak acid in a buffer system)

   $pK_a - pH = \log [acid/base]$ (useful to calculate the ratio of concentration of acid to base)

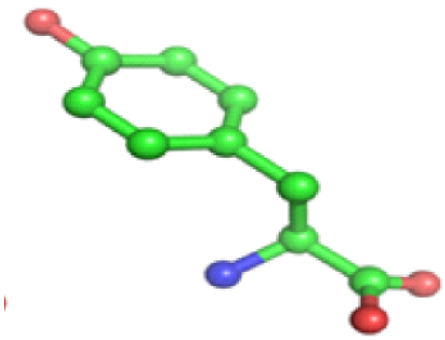
2.10 FURTHER READINGS

1. Albert L. Lehninger: Principles of Biochemistry, Worth Publishers, Inc. New York, 1984.

2. Harper's Illustrated Biochemistry, 29e. Robert K. Murray, David A Bender, Kathleen M. Botham, Peter J. Kennelly, Victor W. Rodwell, P. Anthony Weil, USA.

3. J. L. Jain: Fundamentals of Biochemistry, S. Chand & Company Ltd., India

4. U. Satyanarayana and U. Chakrapani: Biochemistry, UBS Publishers Distributors Pvt. Ltd., Kolkata, India.



Tyrosine

UNIT 3

AMINO ACIDS

Structure

3.1 INTRODUCTION

In the previous units, you learnt about the role of biomolecules such as water. This unit will introduce you to another important class of biomolecules known as proteins, which are made up of basic units known as amino acids. Hence amino acids are considered as building blocks of proteins. You may also recall from your school textbooks in chemistry and biology about proteins. Protein is derived from Greek word *protios* which means “holding the first place”. They are the most abundant macromolecules in the cell and constitute about 50 % of its dry weight.

You should know that amino acids are essential for life. They are joined together by peptide bonds to form the basic structure of proteins. There are about twenty amino acids that are used to form the proteins. It is rather interesting that cells can produce proteins with dissimilar properties and activities by joining same amino acids in many different combinations and sequences. Protein structure ranges from a few to several thousand amino acid residues. These structures are organized hierarchically from primary, secondary, tertiary to quaternary structure.

Amino acids are ancient molecules as several amino acids are the organic compounds that appeared early in the earth’s history.

Objectives

After studying this unit, you will be able to:

- describe the fundamental unit of proteins – the amino acid;
- categorize amino acids on the basis of their properties;
- discuss the optical properties of amino acids; and
- explain the pivotal role of amino acids as ampholytes.

3.2 STRUCTURE OF AMINO ACIDS AND CLASSIFICATION

You have learnt about the structure of amino acids and proteins in school in your chemistry and biology books. In this section we are going to discuss in detail about the structure of amino acids found in proteins.

3.2.1 Structure

Proteins consist of one or more long chains of amino acid residues. Amino acids are composed of carbon, hydrogen, oxygen and nitrogen atoms. However, other atoms are also found in their side chains. Look at the Fig 3.1. It shows the general formula for the amino acids. Amino acids consist of an amino group ($-NH_2$) and a carboxylic acid functional group (-COOH) along with a side chain (-R) specific to each amino acid. Carbon atom next to the carboxyl group is denoted as α-carbon. You will also see that around the α-carbon, other than carboxyl there are three more groups namely hydrogen, amino and R group. The R-group represents the side chains of amino acids which can be as simple as hydrogen atom (H) or a methyl group ($-CH_3$) or a complex structure. Amino acids having a carboxyl group and an amino group bonded directly to the alpha (α) carbon are called as alpha amino acids. In the side chain of amino acids (Lysine, Table 3.1) that have a carbon chain attached to the α-carbon, the carbons are labeled in order α, β and so on.

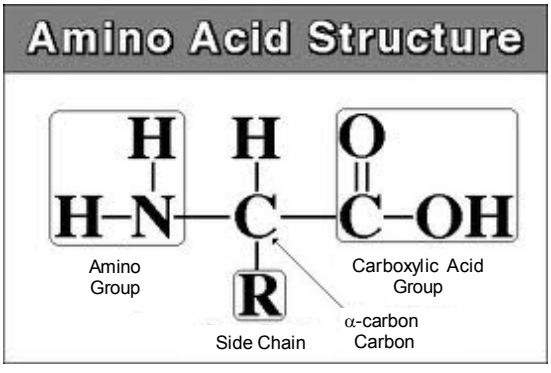


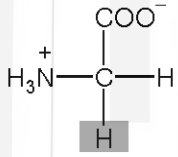
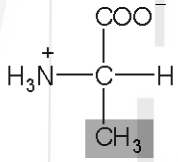
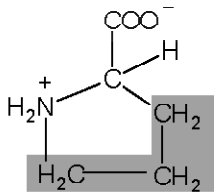
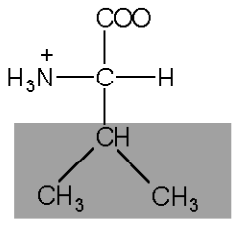
Fig. 3.1: Structure of an Amino Acid.

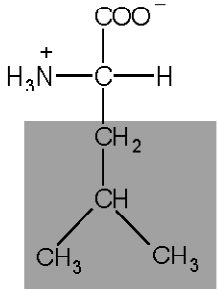
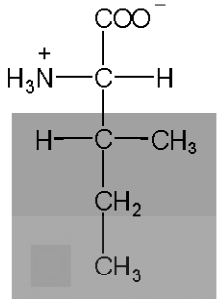
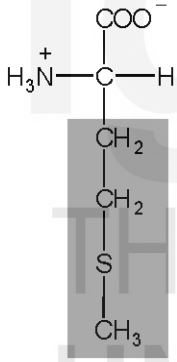
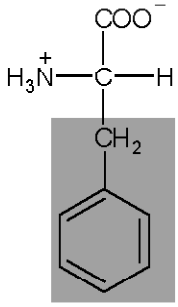
Out of 500 amino acids known so far, 23 are found in proteins. They are called proteinogenic (protein building) amino acids. These amino acids constitute the monomer units to form proteins. The twenty of the twenty three proteinogenic amino acids are encoded directly by the triplet codons in the genetic code. They are known as **standard amino acids.** These amino acids can be condensed into a polypeptide through a process called translation. Look at the

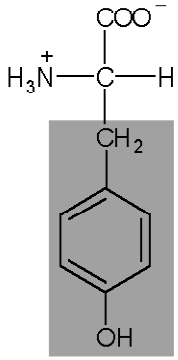
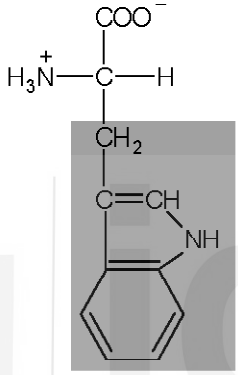
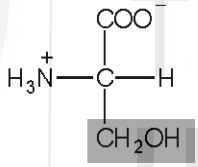
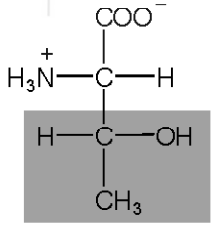
Table 3.1 carefully. It enlists standard amino acids in several categories. Standard amino acids are distinguished from other amino acids that are modified after synthesis of proteins or other amino acids which are present in living organisms but are not found in proteins (non-standard amino acids) [we will discuss them later in section 3.2.4].

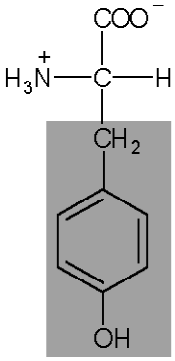
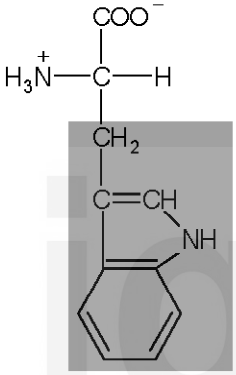
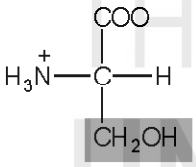
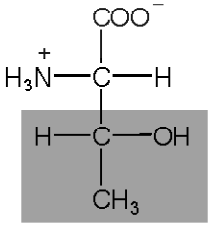
Selenocysteine is a special case. It is a rare amino acid which is introduced during protein synthesis. Often considered as 21st amino acid, it contains selenium instead of sulfur atom in the cysteine amino acid.

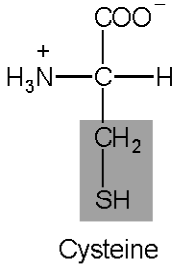
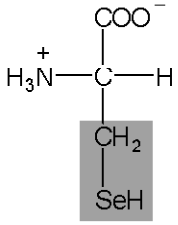
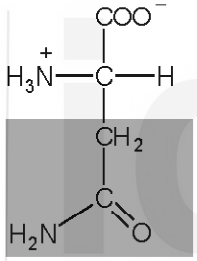
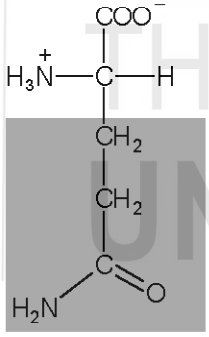
Table 3.1: Properties and Conventions Associated with the Common Amion Acids Found in Proteins

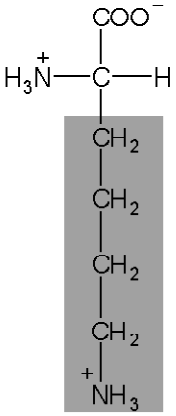
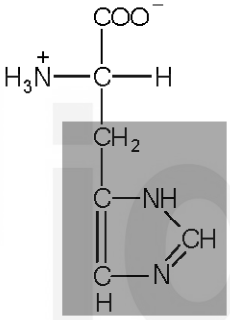
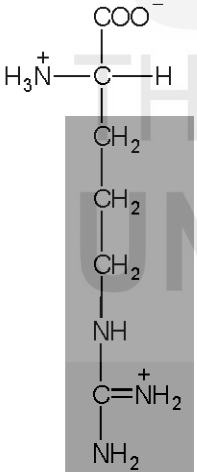
				pK_a values			
Amino acid	***Abbreviation/ Symbol***	***M_r***	***Structure***	***pK_1 (-COOH)***	***pK_R (-NH^+_3)***	***pK_R (R group)***	***pI***
Nonpolar, aliphatic R groups							
Glycine	Gly G	75	 Glycine	2.34	9.60		5.97
Alanine	Ala A	89	 Alanine	2.34	9.69		6.01
Proline (imino acid)	Pro P	115	 Proline	1.99	10.96		6.48
Valine	Val V	117	 Valine	2.32	9.62		5.97

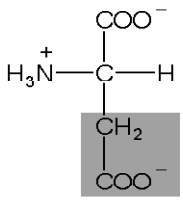
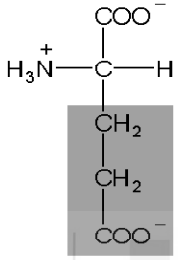
Leucine	Leu L	131	 Leucine	2.36	9.60		5.98
Isoleucine	Ile I	131	 Isoleucine	2.36	9.68		6.02
Methionine	Met M	149	 Methionine	2.28	9.21		5.74
Aromatic R groups							
Phenylalanine	Phe F	165	 Phenylalanine	1.83	9.13		5.48

Tyrosine	Tyr Y	181	 Tyrosine	2.20	9.11	10.07	5.66
Tryptophan	Trp W	204	 Tryptophan	2.38	9.39		5.89
Polar, uncharged R groups							
Serine	Ser S	105	 Serine	2.21	9.15		5.68
Threonine	Thr T	119	 Threonine	2.11	9.62		5.87

Tyrosine	Tyr Y	181	 Tyrosine	2.20	9.11	10.07	5.66
Tryptophan	Trp W	204	 Tryptophan	2.38	9.39		5.89
Polar, uncharged R groups							
Serine	Ser S	105	 Serine	2.21	9.15		5.68
Threonine	Thr T	119	 Threonine	2.11	9.62		5.87

Cysteine	Cys C	121	 Cysteine	1.96	10.28	8.18	5.07
Selenocysteine	Sec U	168.05	 Selenocysteine				
Asparagine	Asn N	132	 Asparagine	2.02	8.80		5.41
Glutamine	Gln Q	146	 Glutamine	2.17	9.13		5.65
Positively charged R groups							

Lysine	Lys K	146	 Lysine	2.18	8.95	10.53	9.74
Histidine	His H	155	 Histidine	1.82	9.17	6.00	7.59
Arginine	Arg R	174	 Arginine	2.17	9.04	12.48	10.76

Negatively charged R groups							
Aspartate	Asp D	133	 Aspartate	1.88	9.60	3.65	2.77
Glutamate	Glu E	147	 Glutamate	2.19	9.67	4.25	3.22

Key convention: Do you notice second and third column in the Table 3.1. You will see that each of the twenty amino acids is assigned three letter abbreviations (upper or lower case) and one letter code (upper case). The three letter abbreviations and one letter code act as shorthand to indicate the composition and sequence of large number of amino acids in proteins. The three letter code consists of first three letters of the name of amino acid. The abbreviations and one letter code is of immense usage in bioinformatics as it reduces the size of data files. Moreover, a transparent key convention for assigning letter code to each amino acid helps to make the purpose of notation easier, quicker and worth remembering. So you are expected to remember single letter and three letter abbreviations for amino acids. For single letter amino acids, you will notice tyrosine what is denoted by letter Y because the letter T is assigned to threonine amino acid.

Only L-α-amino acids occur in proteins. Several free L-α-amino acids play important roles in metabolic processes. D-amino acids are few and occur as small peptides in bacterial cell walls and antibiotics.

SAQ 1

Fill in the blanks:

a) ..... is the simplest amino acid.

b) ..... is an α-imnio amino acid.

c) ..... is the sulfur containing amino acid.

d) ..... and..... ..... are aromatic amino acids.

3.2.2 Stereoisomerism

The α-carbon atom of amino acids except *glycine* is bonded to four different substituents (a carboxyl group, amino group, R group and a hydrogen atom) in a tetrahedral arrangement. Therefore α-carbon is a chiral centre and asymmetric in all of them except glycine.

Molecules with asymetric or chiral centres are not superimposable on their mirror image in the same way that a left hand is not superimposable on its mirror image, a right hand.

Thus, two unique spatial arrangements of four groups around the α-carbon of amino acids gives rise to two possible **stereoisomers** (Fig. 3.2a & b). Special nomenclature of D or L-system is used to specify the absolute configurations of the substituent's of the asymmetric carbon atom. If functional groups of amino acids match with those of the configuration of L-glyceraldehyde (for more details, please refer to Unit 5), these amino acids are called L-α-amino acids. *Those amino acids having $-NH_3$ group to the right hand side are designated as D-forms and those having $-NH_3$ group to the left are designated as L-forms.* Look at Fig 3.2, the L and D configurations of amino acids is not superimposable. The stereoisomers which are non-superimposable mirror images of each other are known as **enantiomers**.

Two amino acids threonine and isoleucine $2^n = 2^2 = 4$ have two asymmetric carbon atoms each and thus have optical isomers.

All amino acids except glycine have a chiral centre and are optically active (rotate the plane of polarized light). Some of the amino acids are ***dextrorotatory*** (rotating plane polarized light to the right, clockwise direction) and some ***levorotatory*** (rotating plane polarized light to the left). Many L-α-amino acids rotate the plane of polarized light to the right (***dextrorotatory)*** while others rotate the plane of polarized light to the left (***levorotatory)***. Similarly D-isomers can be either ***dextrorotatory*** or ***levorotatory.*** Dextrorotatory molecules are designated by prefix "(+)-" or "d-" signs while levorotatory molecules are designated by prefix "(-)-" or "l-" signs respectively. The "d" and "l" prefixes are distinct from the uppercase "D" and "L" prefixes that are based on actual configuration of enantiomer. Due to confusion and its inherent limitations, DL-nomenclature is not used nowadays. Another concept of **RS notation*** of isomers is currently being used and it is proving more useful for defining the asymmetry of biomolecules. If any atom (other than hydrogen) or group on the asymmetric carbon is on the right hand side, that asymmetric carbon is designated as R (Latin rectus, ''right'') and if any atom (other than hydrogen) or group on the asymmetric carbon is on the left hand side, then that asymmetric carbon is designated as S (Latin sinister, ''left''). The configurations of an isomer of isoleucine amino acid (L-isoleucine) having two asymmetric carbons (C-2, C-3) will be represented as (2 S),(3 S) – isoleucine.

The symbols D and L refer to the absolute configuration of four groups around the chiral carbon atom and do not identify the property of light rotation.

***RS system** describes the **absolute configuration of individual chiral centres**. To give R/S prefix to a compound certain rules have to be followed:

a) In case of a tetrahedral carbon atom the four substituents are assigned priorities; the highest priority is given to the group having higher atomic weight if all the four atoms are different.

b) When all the atoms are not different then other atoms are looked into. The priorities of some of the commonly encountered functional groups in biomolecules are: SH>OH>NH_2> COOH>CHO>CH_2OH> CH_3>H.

c) Once the priorities are assigned then the molecule is oriented such that the smallest ranked group is facing away from the reader and we move from the highest to the lowest ranked order.

d) The chiral centre is labelled R if the direction taken is clockwise and S if counter clockwise. According to this system the standard reference D glyceraldehyde is R and that of L isoleucine is (2S), (3S) isoleucine.

Fig. 3.2a: Stereoisomerism in amino acid (geometric formulas).

Solid wedge shaped bonds project out of the plane of paper, the dashed bonds behind it.

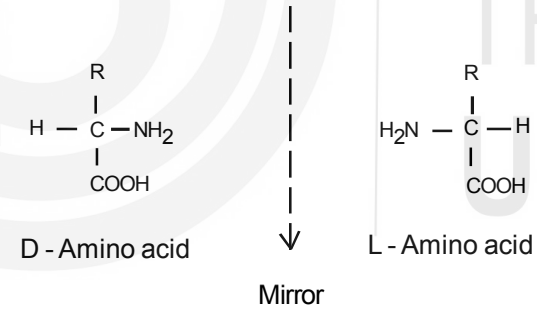


Fig. 3.2b: Stereoisomerism in amino acid (Fischer projections).

In a Fischer convention, bonds that extend above the page are represented by horizontal lines and bonds that extend below the page are represented by the vertical lines.

SAQ 2

Draw the structure of Isoleucine. How many chiral centers does it have? Draw its optical isomers.

3.2.3 Classification Based on R-group

There are several ways to classify amino acids, but the simplest way of classification is based on structure and general chemical properties of their

side chains or R groups. The chemical properties of the R groups are responsible for several characteristics of amino acids such as chemical reactivity, ionic charge and hydrophobicity. Amino acids are grouped into five main classes based on the polarity, size and shape of R-groups present in their side chains (Fig. 3.3).

1) **Non-polar aliphatic R groups:** Glycine, Alanine, Proline, Valine, Leucine, Isoleucine and Methionine amino acids belong to this group. These amino acids have aliphatic side chains which are non ionizable and hydrophobic in nature. Glycine is the simplest of twenty amino acids; its smallest side chain makes little involvement to the hydrophobic interactions. In a protein, it provides minimum steric hindrance and allows flexibility. Alanine is considered to be the parent of all other amino acids except glycine. All of them can be derived by replacing one or two H atoms of the methyl group bonded to the α-carbon atom. The side chains of the non polar aliphatic amino acids stabilize protein structure by hydrophobic interactions.

Amino acids having a secondary amine group are called imino acids. Proline is formally not an amino acid but imino acid.

It is different from other standard amino acids structurally because its secondary amino (imino) group is formed by ring closure between the R group and the amino nitrogen. It is the only cyclic amino acid. The structure of proline confers rigidity to the peptide chain in proteins.

2) **Aromatic R groups:** Phenylalanine, Tyrosine and Tryptophan amino acids have bulky, aromatic group as side chains. They belong to this category and have phenyl, phenolic and indole rings respectively as their R-groups. The hydroxyl group of tyrosine can form hydrogen bonds with water. These amino acids are responsible for UV absorption of proteins at 280 nm.

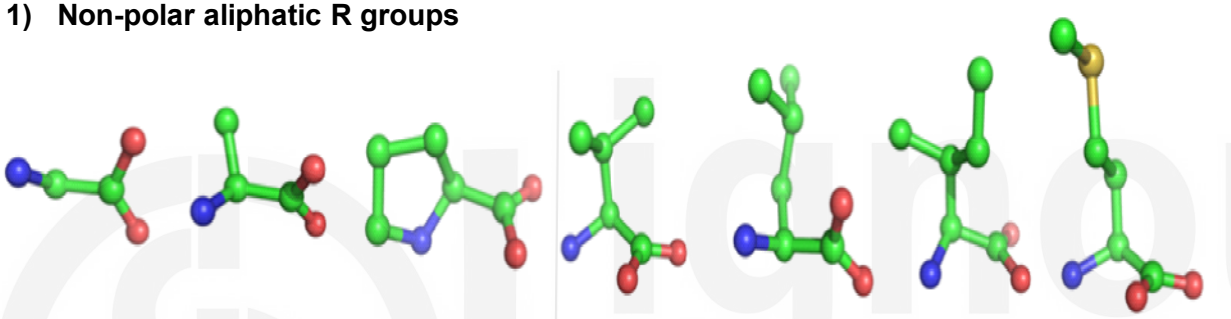
3) **Polar uncharged R groups:** Serine, Threonine, Cysteine, Asparagine and Glutamine amino acids belongs to this group. These amino acids have polar and ionizable groups in their side chains such as Serine (hydroxyl, -OH), Threonine (hydroxyl, -OH), Cysteine (sulphydryl, -SH), Asparagine (amide, -CO-NH_2) and Glutamine (amide, -CO-NH_2). The polar functional groups of these amino acids participate in hydrogen bonding with water.

4) **Positively charged R groups:** Lysine, Histidine and Arginine are positively charged amino acids, as they contain *basic* groups in their side chains. The side chains groups of Histidine, Lysine and Arginine are called imidazole, amino and guanidinium respectively. The pI (isoelectric point, discussed in the next section 3.3 of unit) of histidine is near physiological pH, hence it works as biological buffers in proteins. All basic proteins are rich in these amino acids in their sequences. These amino acids mostly occur on the protein surface to interact with water molecules.

5) **Negatively charged R groups:** Aspartate and Glutamate amino acids have *acidic* group (carboxylic, -COOH) in their side chains. All acidic proteins are rich in these amino acids in their sequences. These amino acids mostly occur on the protein surface to interact with water molecules. In proteins these side chain form strong ionic bond with basic amino acids, which stabilize the protein structures.

You might appreciate a colorful representation of the three dimensional structures of amino acids (Fig. 3.3) along with their three letters and one letter code. Carbon atoms (green), Nitrogen (blue), Oxygen (red) and Sulfur (yellow) are shown in the Figs. Hydrogen atoms are omitted from the Figs.

1) Non-polar aliphatic R groups

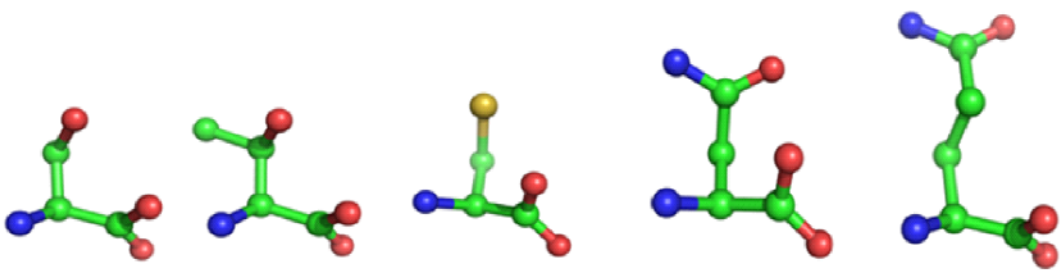


Glycine, Gly (G) | Alanine, Ala (A) | Proline, Pro (P) | Valine, Val (V) | Leucine, Leu (L) | Isoleucine, Ile (I) | Methionine, Met (M)

2) Aromatic R groups

Phenylalanine, Phe (F) | Tyrosine, Tyr (Y) | Tyrptophan, Trp (W)

3) Polar uncharged R groups



Serine Ser (S) | Threonine, Thr (T) | Cysteine, Cys (C) | Asparagine, Asn (N) | Glutamine, Gln (Q)

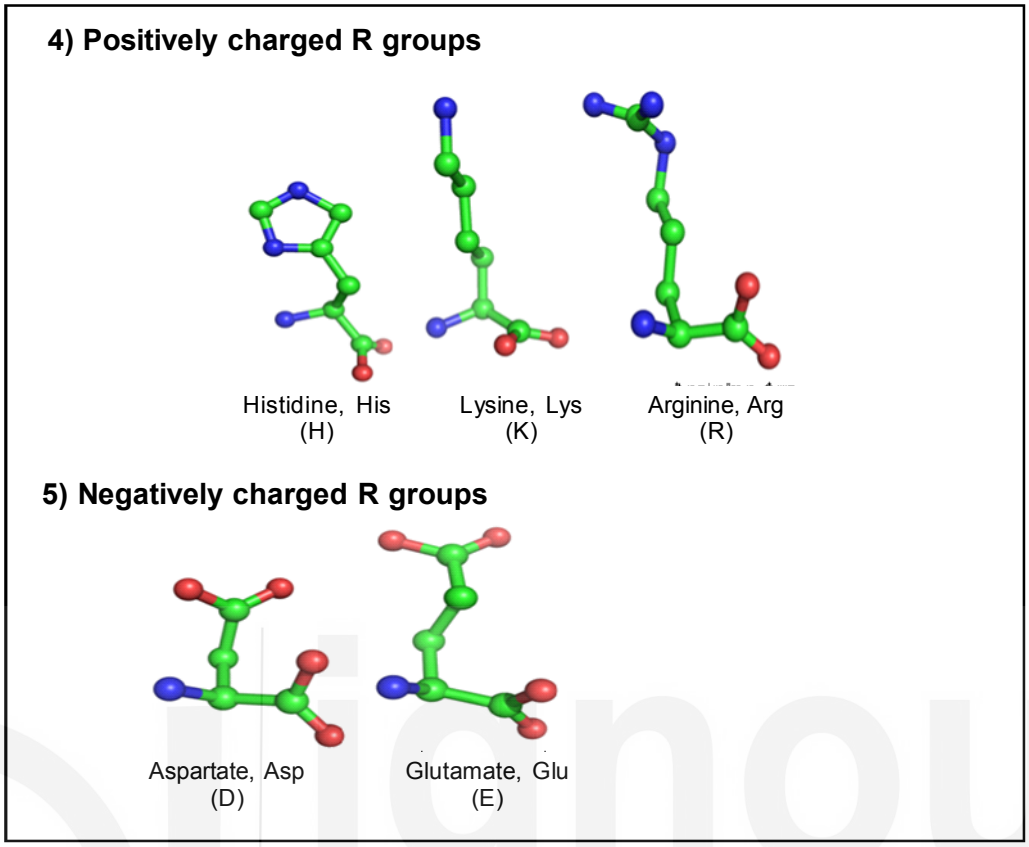


Fig. 3.3: Classification of Amino Acids on R-Group.

SAQ 3

What is the importance of R-group in an amino acid molecule?

From a nutritional perspective, some amino acids are absolute dietary necessities if normal growth is to be maintained. Dietary protein is the main source of amino acids. Amino acids taken in the diet must be modified to provide the adequate levels of all the amino acids to meet the requirements of the human body.

Essential Amino acids: Animals either lack the ability to synthesize certain amino acids or they make them in insufficient quantities than required during phases of active growth. These amino acids are classified as essential amino acids. Almost half of the amino acids required for protein synthesis fall in this category and they have to be provided in the diet. They include leucine, isoleucine, lysine, methionine, phenylalanine, threonine, tryptophan, valine, and histidine. Arginine is semi essential. Limitation of even one amino acid leads to a negative nitrogen balance. The amino acid that can be synthesized by the animals are called as non-essential amino acids.

3.2.4 Non-Standard Amino Acids

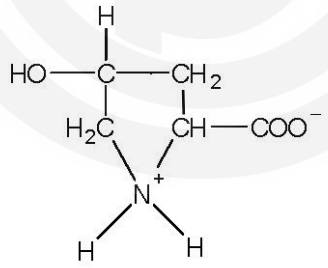
As we have mentioned before that in addition to twenty common amino acids present in proteins several other amino acids exist in the living organisms. These amino acids occurs naturally in cells but do not participate in protein synthesis and are called non-standard amino acids.

They are either absent in proteins or are formed by modification of standard amino acids in a peptide molecule.

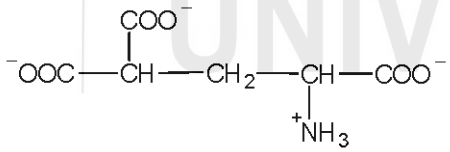
Non-standard amino acids found in proteins

Non standard amino acids (Fig. 3.4) are generally formed by post-translational modification during protein synthesis. These modifications are often necessary for the function or regulatory role of a protein. Some of the examples include 4-hydroxyproline (derivative of proline) and 5-hydroxylysine (derivative of lysine) found in collagen of connective tissues. Desmosine is formed from four lysine residues and is present in elastin protein. The carboxylation of glutamate to form γ-carboxy glutamate found in blood clotting protein prothrombin allows better binding of calcium cations. Some of the amino acids are modified transiently to alter protein function. The addition of acetyl, methyl, adenylyl or ADP-ribosyl groups to amino acids can increase or decrease a protein activity. Phosphorylated amino acids such as serine, threonine and tyrosine are often used to regulate the activity of several proteins.

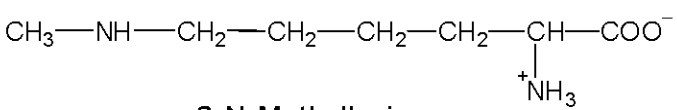
There are several additional amino acids (~300) in cells which are not found in proteins (non-protein amino acids) but play a significant role in metabolism or exist as natural products. Amino acids such as ornithine and citrulline are important metabolic intermediates in the urea cycle. β-Alanine, an isomer of alanine is a constituent of vitamin pantothenic acid and of coenzyme A. Quaternary amine creatine occurs naturally in vertebrates and helps to supply energy to all cells in the body primarily muscle. Nonprotein amino acid, γ-aminobutyrate is found in free form in the brain.



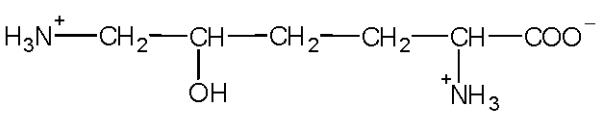
4-Hydroxyproline



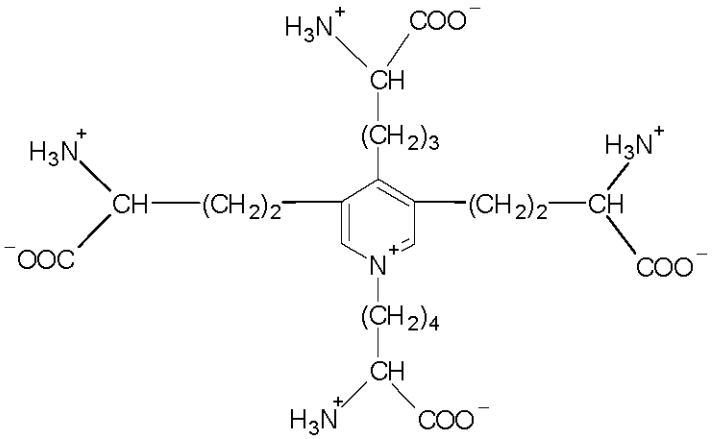
γ-Carboxyglutamate



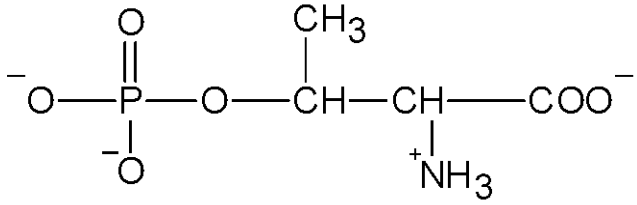
6-N-Methyllysine



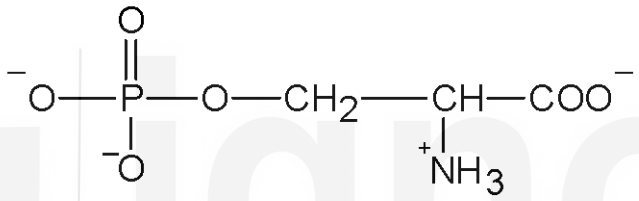
5-Hydroxylysine



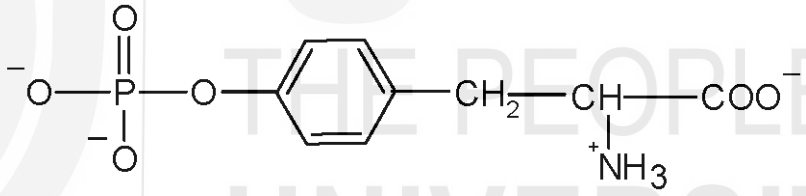
Desmosine



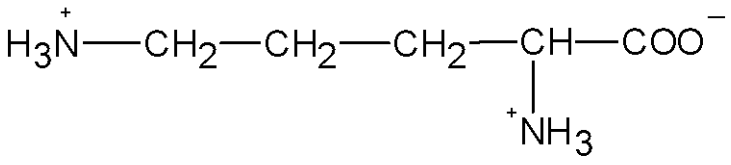
Phosphothreonine



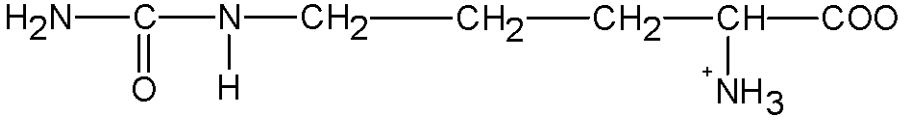
Phosphoserine



Phosphotyrosine



Ornithine



Citrulline

Fig. 3.4 : Chemical structures of modified amino acids.

Biologically Active Amino Acids

Other than their role as a component of protein, amino acids perform several biological roles such as:

1) Tyrosine amino acid is a precursor of catecholamine neurotransmitters dopamine, epinephrine and norepinephrine.

2) Thyroxine (derivative of tyrosine) and Indole acetic acid (derivative of tryptophan) are very well known hormones which regulates the metabolism of cell.

3) Glycine is a precursor of porphyrins such as heme.

4) Arginine, citrulline and ornithine participate in the urea cycle for disposal of nitrogenous waste.

5) Ornithine and S-adenosylmethionine are precursors of polyamines.

6) Aspartate, glycine and glutamine are precursors of nucleotides.

Non-Protein amino acids also perform several significant roles as metabolic intermediates such as biosynthesis of α-amino butyric acid (GABA). However, not all the functions of other abundant non protein amino acids are known.

SAQ 4

What are the amino acids with one letter abbreviation of W and S? Are they essential or non-essential amino acids? What types of side groups are present in their side chains?

3.3 PROPERTIES OF AMINO ACIDS

3.3.1 Amino Acids as Acids and Bases

In order to understand the acid-base properties of proteins, we will begin with the acid base behavior of their constituent amino acids. When an amino acid lacking ionizable R group is dissolved in water at neutral pH, it exists as the dipolar ion, a **zwitterion** and can act as acid as well as base (Fig. 3.5). It bears no net charge as it has equal number of oppositely charged ions. Such amino acids exhibiting acid-base character (amphoteric) are called **ampholytes**. The amino and carboxyl groups of amino acids as well as ionizable R groups of amino acids function as weak acids and bases.

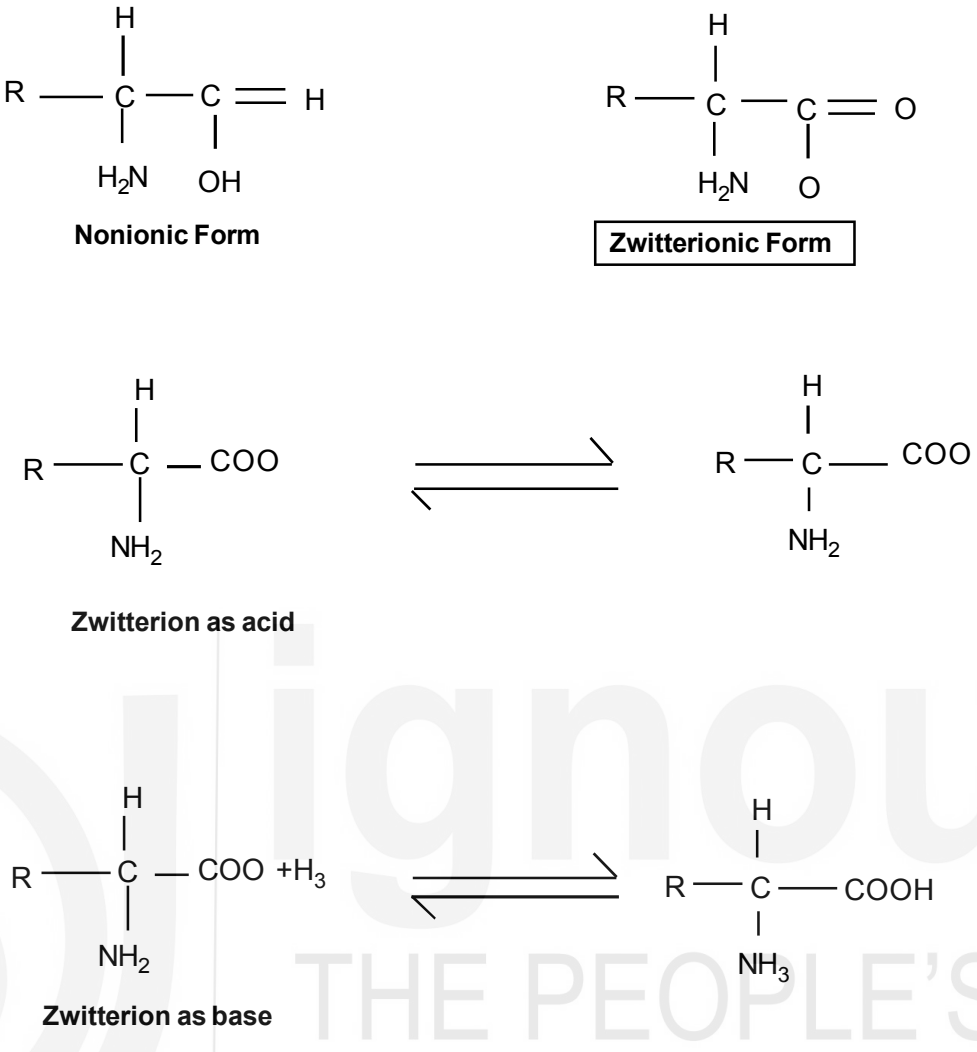


Fig. 3.5 : Nonionic and Zwitterionic forms of amino acids.

R-COOH and R-NH_3^+ are considerably weak acids. However, R-COOH is stronger acid than R-NH_3^+. At the physiological pH of 7.4, carboxyl groups and amino groups are completely ionized and are present as R-COO^- and R-NH_3^+ respectively.

3.3.2 Titration Curves of Amino Acids

In the previous unit you have been introduced to the concept of ionization. The behavior of ionizable groups in amino acid with respect to pH change can be depicted by acid-base titration. Acid-base titrations involve steady addition or removal of protons. The carboxyl and amino group of the amino acid loses proton on titration with a strong base such as sodium hydroxide. Glycine amino acid lacks R group. Therefore behavior of COOH and NH_2 group can be observed without interference of R group in the titration curve (Fig. 3.6); do you notice the plot that has two distinct stages? What are these two stages? At a very low pH, glycine exists in fully protonated form and has a net charge of +1. All the molecules of glycine are present as $^+H_3NCH_2COOH$. However at the mid point of first stage of titration, equimolar concentrations of proton donor ($^+H_3NCH_2COOH$) or proton accepter ($^+H_3NCH_2COO^-$) forms of the glycine

exists. At this point, pH is equal to pK_a of the protonated group being titrated. The pK_a is a measure of the acid strength of weak acids. The pH of glycine at midpoint is 2.34. So pK_a of COOH group is 2.34 (labeled pK_1 in the plot). As titration proceeds, removal of first proton is fully complete. At this stage, glycine is in fully ionized form ($^+H_3NCH_2COO^-$) as a zwitterion (dipolar ion) with no net electric charge. The pH of glycine at this point is 5.97 (labeled as pI in the plot). Second stage of titration involves removal of proton from the NH_3 the second group of glycine. The second mid point of titration for this stage is 9.60. The pK_a for the NH_3 group is 9.6 (labeled as pK_2 in the plot). The titration gets complete at pH 12. Most of the molecules at this stage will be negatively charged and exist as ($H_2NCH_2COO^-$). The carboxyl group of glycine is 100 times more acidic than the carboxyl group of acetic acid (pK_a 4.76).

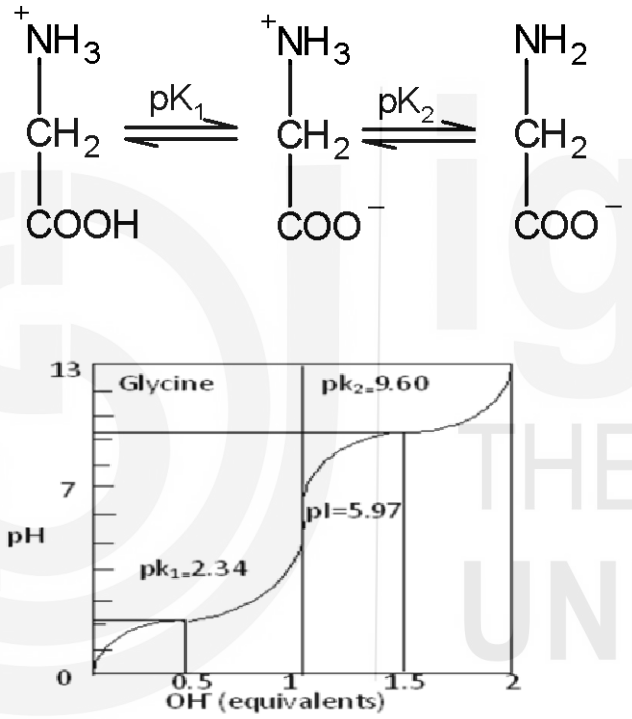


Fig. 3.6 : Titration curve of 0.1 M Glycine at 25°C.

The characteristic titration curve of the amino acid helps us to determine:

1. Quantitative measurement of the pKa of the ionizable groups (carboxyl and amino) of amino acids.

2. Regions of buffering power: There are two regions of buffering power in the amino acid, one is centered on 2.34 and other one is around 9.60.

3. Relationship between net electric charge of the amino acid and pH of the solution. The characteristic pH where amino acid exists in the dipolar form ($^+H_3NCH_2COO$-) with no net charge is called as **isoelectric point** or **isoelectric pH,** designated as pI. The pI of the glycine will be:

pI= ½ (pK_1 + pK_2) = ½ (2.34 + 9.6) = 5.97

At pH above its pI, glycine (negatively charged) will move towards the anode and at a pH below its pI (positively charged) it will move towards the cathode.

The pK_a values of free amino acids depend on the ionizable R groups. The polar environment favors the charged form ($R\text{-}COO^-$ or $R\text{-}NH_3^+$) and the non-polar environment favors the uncharged form (R-COOH or $R\text{-}NH_2$). The α-amino and carboxyl group of amino acids loses their charge on their incorporation in the peptide bonds. Therefore other than N-terminal and C-terminal residues, the ionizable R groups of amino acids such as histidine, lysine, arginine, aspartate, glutamate, cysteine and tyrosine influences their pK_a values. The pK_a values of these amino acids in proteins will be different than when they are in free form due to the influence of protein microenvironment. Glycine does not act as a buffer at physiological pH. **Histidine is the only amino acid that has significant buffering at neutral pH.** All amino acids that do not have an ionisable R group like glycine have similar titration curves. Amino acids with ionisable R groups such as histidine or glutamic acid have three stages corresponding to the three functional groups.

SAQ 5

pK_1 of valine is 2.2 and pK_2 is 9.7. Find out pI of Valine.

3.3.3 Physical Properties

Amino acids are colorless crystalline substances. Their shape varies from sharp needle shape (tyrosine) to hexagonal plates (cystine). Amino acids can be sweet (glycine and alanine) to bitter (arginine) or tasteless (tyrosine). Amino acids have high melting points and are soluble in the polar solvents (water and ethanol) because of their charged functional groups and insoluble in the non-polar solvents such as benzene, hexane and ether. The hydrophobic (water repelling) and hydrophilic (water loving) order of some amino acids is as given below:

Hydrophobic (non-polar) amino acids:

Val > Leu> Phe > Cys> Ala > Gly

Hydrophilic (polar) amino acids:

Arg > Lys > Asp > Asn > Ser > Thr

Amino acids do not absorb the visible light and are therefore colorless. Tyrosine and tryptophan absorb high wavelength ultraviolet light (250-290 nm). It helps in the detection and determination of protein quantity in different solutions. The light absorption property is extensively used to determine the concentration of biomolecules in the solution.

Amino acids along with their amino group, carboxyl group and R-group can undergo several chemical reactions. Peptide bond and disulphide bond formation affects the protein structure and their formation.

3.4 SUMMARY

Amino acid is the most basic fundamental unit of proteins. Twenty common amino acids constituting the proteins contain a α- carboxyl group, an α-amino group and a distinctive R group. These amino acid residues are categorized in five different classes on the basis of their polarity and charge of their R group. Amino acids exhibit both acids as well as base properties and are often called ampholytes. Amino acids also exhibit characteristic titration curves. Long chains of amino acids constitute polypeptides which form the basic structural unit of proteins. Primary structure of amino acids leads to folding of polypeptides in secondary structures which are further folded in supersecondary structures. Therefore we can say that amino acids play a central role in biochemistry.

3.5 TERMINAL QUESTIONS

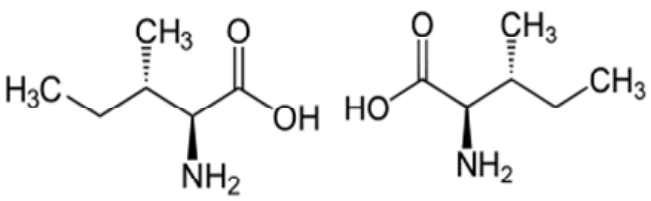
1. Define the term zwitterion. When does pH becomes pI?
2. How many amino acids are known to form proteins in living organisms?
3. How amino acids can be classified on their chemical nature?
4. How amino acids act as acid as well as base?

3.6 ANSWERS

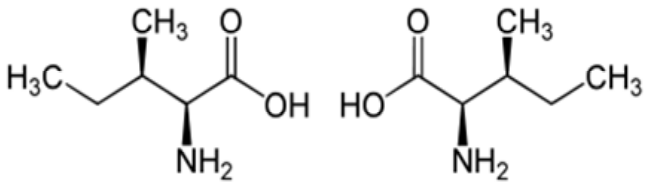
Self-Assesment Questions

1. a) Glycine b) Proline c) Cysteine/Methionine
   d) Phenylalanine, Tyrosine

2. Isoleucine is an essential amino acid that has a chiral side chain. Its structure is given in Table 3.1. Four stereoisomers of isoleucine are possible. They are as given below:



L-isoleucine (2*S*,3*S*) and D-isoleucine (2*R*,3*R*)



L-*allo*-isoleucine (2*S*,3*R*) and D-*allo*-isoleucine (2*R*,3*S*)

3. The R-Group is the variable part of an amino acid molecule. The R-group in the amino acid can be a methyl group or hydrogen or a complex cyclic chain. The R-group can dictate the primary, secondary, tertiary and quaternary structure of proteins.

4. W the one letter code stands for Tryptophan amino acid and S letter code stands for polar amino acid Serine. Tryptophan is an essential amino acid and Serine is a non-essential amino acid. The side chains in both the amino acids are polar and neutral.

5. 5.95

Terminal Questions

1. Zwitterion is a dipolar ion with a positive as well as negative charge but it is a neutral molecule. Amino acids are well known examples of zwitterions because they contain acidic as well as basic group in a single molecule (an ammonium ion and a carboxylate ion). At very low pH, ionizable groups are fully protonated and at very high pH these groups are unprotonated. The predominant ionic form of molecules in solution depends on the pH. The characteristic pH where amino acid exists in the dipolar form ($^{+}H_3NCH_2COO^{-}$) with no net charge is called as **isoelectric point** or **isoelectric pH** designated as pI for example pI of the alanine is 6.02.

2. There are twenty different known amino acids that form proteins related to the genetic code of the living organisms. There are still many other amino acids that are not yet known.

3. Amino acids can be classified into three groups depending on their chemical nature:

   a) Neutral amino acids

   b) Acidic amino acids

   c) Basic amino acids

   a) Neutral amino acids: They are largest group of amino acids and are further subdivided

      i) Aliphatic amino acids
      Mono amino mono carboxylic amino acids:
      Glycine, Alanine, Valine, Leucine, Isoleucine

      ii) Hydroxy amino acids
      Serine, Threonine

iii) Aromatic amino acids
Phenylalanine, Tyrosine

iv) Heterocyclic amino acids
Tryptophan, Histidine

v) Imino acids
Proline

vi) Sulphur containing amino acids
Cysteine, Methionine

b) Acidic amino acids: These amino acids have two carboxyl groups - COOH and one -NH_2 group and are also known as monoamino dicarboxylic acids such as Aspartic acid/Asparagine and Glutamic acid/Glutamine.

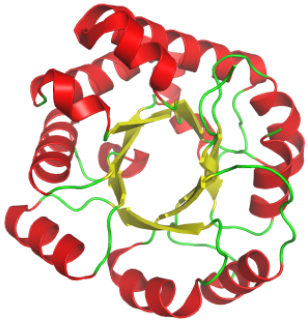
c) Basic amino acids: These amino acids have one carboxyl groups - COOH and two-NH_2 groups and are also known as diamino monocarboxylic acids such as Arginine and Lysine.

4. Please see section 2.3.1

3.7 FURTHER READINGS

1. Lehninger: Principles of Biochemistry, David L. Nevron and Michael M. Cox. W.H. Freeman, 4th Edition, 2004. New York.
2. Harper's Illustrated Biochemistry, 29e. Robert K. Murray, David A. Bender, Kathleen M. Botham, Peter J. Kennelly, Victor W. Rodwell, P. Anthony Weil, USA.
3. Donald J Voet Principles of Biochemistry, Jophn Wiley and Sons, Inc, USA.
4. J. L. Jain: Fundamentals of Biochemistry, S. Chand & Company Ltd., India.
5. U. Satyanarayana and U. Chakrapani: Biochemistry, UBS Publishers Distributors Pvt Ltd., Kolkatta, India.
6. Thomas M. Devlin: Textbook of Biochemistry, John Wiley and Sons, Inc. Danvers, MA, USA.

UNIT 4



Tertiary Structure of Protein

PEPTIDES AND PROTEINS

Structure

4.1 INTRODUCTION

Proteins are large complex molecules and play several vital roles in the body. These roles includes regulatory, extracellular as well as intracellular and many others. How will you understand the function of proteins? The answer lies in the structure of proteins. In the previous unit you have learnt that the amino acids are building blocks of proteins. A linear chain of amino acid residues is known as polypeptide. The amino acids are joined together by the peptide bond. Proteins contain at least one big polypeptide. The smaller polypeptides with amino acid sequences generally less than 20-30 amino acid residues are simply called as peptides or sometimes oligopeptides.

The genetic information stored in the DNA is expressed in the form of proteins in the cell. Each protein has a specific amino acid sequence that is specified by the unique nucleotide sequence of the gene encoding the protein. The process of synthesizing the protein is known as translation.

The synthesis of proteins in vivo occurs in a systematic manner in the order of the sequence of nucleotides in a gene. Since there are twenty different amino acids, for a protein with n amino acid residues; there can be 20^n possible sequences. Each amino acid has characteristic structure and function. Its place in the structure of proteins influences the function of protein as well as its properties. The characteristics of an individual protein are therefore determined by its amino acid composition. Several combination of amino acids leads to wide variety of proteins and adds to the diversity of their functions. Protein structures are organized hierarchically from primary, secondary, tertiary to quaternary structures.

Objectives

After studying this unit, you will be able to:

- describe the peptide bond;
- illustrate the structure of proteins;
- explain the organization of proteins; and
- discuss the macromolecular complexity of proteins.

4.2 PEPTIDES

As you know functional groups in an organic molecule determine which reaction they may undergo. Amino acids with their functional groups such as carboxyl, amino and R groups can undertake several chemical reactions. Reactions due to amino group include transamination, amide bond formation (amidation) and oxidative deamination; for carboxylic acid groups, reactions include formation of esters (esterification), amides and acid anhydrides; for R groups such as –OH and –SH reactions include oxidation and esterification. The most important reaction of amino acids is the condensation reaction that leads to the formation of a peptide bond.

4.2.1 Peptide Bond

The peptide bond is a covalent bond between α-amino group of one amino acid and α-carboxyl group of another forming an amide linkage (CO-NH). Peptide bond formation is a condensation reaction which involves removal of water (dehydration). Look at Fig. 4.1. Two amino acids are joined covalently through an amide linkage termed as **peptide bond**.

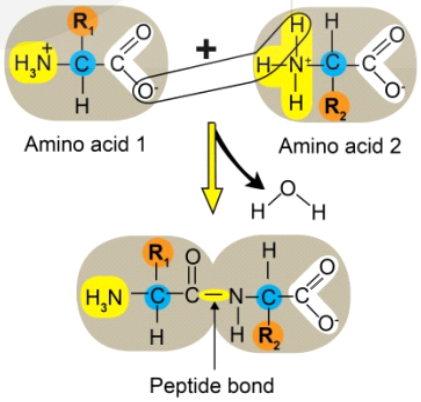


Figure. 4.1: Formation of a Peptide Bond.

When two amino acids molecules are linked, the product is called a dipeptide. On addition of more amino acids the chain length increases e.g. tripeptide contains three amino acid residues and a tetrapeptide; four and so on. When more amino acids are joined, the structure is called oligopeptide. Joining of many more amino acids leads to formation of a polypeptide.

Amino acid polymers can also be classified on the basis of their molecular weight or the numbers of amino acid residues they contain. Molecules with molecular weight ranging from several thousand to several million daltons are called polypeptides. Molecules with low molecular weights, fewer than 50 amino acids are called as peptides. Proteins may have thousands of amino acid residues and consist of one or more polypeptide chains.

Pauling and Corey, two well known scientists who analyzed the geometry and dimensions on the peptide bond laid the basis of our current understanding of protein structure. X-ray diffraction studies of crystals of amino acids characterize the peptide bond (Fig. 4.2) as rigid and planar (flat). All the six atoms of the peptide group lie in a single plane. The C–N bonds formed by joining of two amino acids are shorter than other types of C–N bonds. The peptide bond has a partial double bond character. It shows resonance or partial sharing of two pairs of electrons between the carbonyl oxygen and the amide nitrogen. There is little freedom or no rotation around this bond. Rotation is permitted about the N–C_α and the C–C_α bonds. The rigidity of the peptide bond has serious structural consequences for the protein structure.

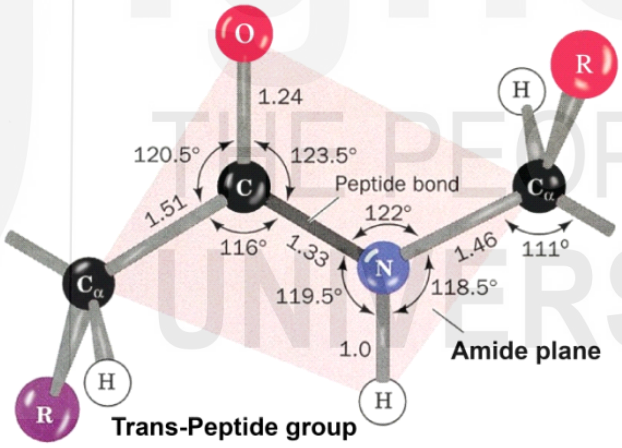


Fig. 4.2: Structure of Peptide Bond.

Each amino acid in a polypeptide chain is called residue. The terminal amino acid residue with free amino group is called as N-terminal residue and is written on the left hand side in an amino acid sequence. It refers to the start of a protein or polypeptide. Amino acid with free carboxyl group is known as C-terminal residue is written on the right hand side in an amino acid sequence. When the protein is translated from messenger RNA, it is synthesized from N-terminus to C-terminus. The convention for writing polypeptide sequence is to put the N-terminus on the left and write the sequence from N-terminus to C-terminus. The names of intermediary amino acid residues are written in the same sequence as they are placed. For example, glutathione is a tripeptide of non-protein origin with glutamic acid (N-terminus), cysteine and glycine (C-terminus) residues. It is written as Glu-Cys-Gly.

SAQ 1

Answer the following:

a) Do the –R-groups bound to the central carbons participate in the peptide bond?

b) Do the –H groups bound to the central carbons participate in the peptide bond?

c) Do the amine and the carboxyl groups attached to central carbons participate in the union between amino acids?

4.3 PROTEIN STRUCTURE

Proteins are physically and functionally complex macromolecules that perform multiple roles. What is it that makes one protein an enzyme, another a hormone or an antibody? How do they differ? What do you think? The answer to these questions lies in the protein structure. All the twenty amino acids do not occur in equal amounts in all the proteins. The amino acid composition of proteins is variable. Some of the amino acids may be present only once or absent from particular proteins while some of the amino acids may be repeated many times. The smaller amino acids predominate in most of the proteins. Some of the proteins contain a single polypeptide chain but many proteins are multisubunit. The multisubunit proteins are more complex in structure and contain two or more polypeptide chains. The individual polypeptide chains in the multisubunit proteins may be identical or different. If the subunits (individual polypeptide chain) are identical the protein is known as **oligomeric** and the repeating subunit or individual polypeptide chain is called as **protomer**. Hemoglobin has four polypeptide chains (two identical α and two identical β). It can be considered as a tetramer of four subunits or dimer of $\alpha\beta$ protomer. If the two polypeptides are linked covalently by disulphide bonds such as in insulin, they are generally called chains and not subunits.

Proteins do not have uniform regular structures. Proteins can be described in terms of levels of organization in primary, secondary, tertiary and quaternary structures.

4.3.1 Primary Structure

The amino acid sequence of a protein is unique to that protein, and defines the structure and function of the protein. The primary structure of proteins refers to amino acid sequence of the polypeptide chain. It is held together by covalent peptide bonds, which are formed during the process of protein biosynthesis or translation. The counting of residues conventionally begins from the N-terminal end (NH_2-group, the end where the amino group is not involved in a peptide bond). A specific sequence of nucleotides in DNA is transcribed into mRNA,

which is read by the ribosome in a process called translation. The sequence of a protein can be determined by methods such as chemical methods like Edman degradation or tandem mass spectrometry technique. However, the primary sequence can be read directly from the sequence of the gene using the genetic code. Post-translational modifications such as disulfide formation, phosphorylations and glycosylations are usually also considered a part of the primary structure, and cannot be read from the gene as they occur after translation is completed.

The primary structure is written in form of the amino acid sequence of the polypeptide such as for given polypeptide: (N-terminal) H_2N-GIVEQCCTSICSLYQ-LENYCN-COOH (C-terminal).

Please recall from the previous unit, the one letter code for each amino acid, you will notice that the N-terminal amino acid residue is Glycine and C-terminal is Asparagine.

SAQ 2

Answer the following:

a) ..... is the repeating unit in proteins.

b) Primary structure of proteins represents 3-D structure of proteins (True/ False).

c) A dipeptide has two amino acids and one peptide bond (True/False)

d) Name one chemical method to determine the primary sequence of proteins?

4.3.2 Secondary Structure

Secondary structure refers to highly regular local sub-structures. Two main types of secondary structure observed in proteins are, the alpha helix and the beta sheet. These were suggested in 1951 by Linus Pauling. These secondary structures are defined by patterns of hydrogen bonds between the main-chain peptide groups present in same chain or in different chains. They have a regular geometry, being constrained to specific values of the dihedral angles Ψ and ϕ on the Ramachandran plot (for details, please refer to Unit-8 of the course C-3, Proteins).

Both the alpha helix and the beta-sheet represent a way of saturating all the hydrogen bond donors and acceptors in the peptide backbone. Some parts of the protein are ordered but do not form any regular structures. They should not be confused with random coil, an unfolded polypeptide chain lacking any fixed three-dimensional structure. Several sequential secondary structures may form a “supersecondary unit”.

α-helix: It is most common secondary structure found in proteins (Figs. 4.3 and 4.4). It has following salient features:

The amino acids in an α–helix are arranged in a right-handed **helical** structure where each amino acid residue corresponds to a 100° turn in the helix.

It is most common and also stable conformation asit makes optimum use of H-bonds possible using main chain atoms.

The helix has 3.6 residues per turn, and a translation of 1.5 Å (0.15 nm) along the helical axis.

The pitch of the alpha-helix (the vertical distance between one consecutive turn of the helix) is 5.4 Å (0.54 nm) which is the product of 1.5 and 3.6. The N-H group of an amino acid forms a hydrogen bond with the C=O group of the amino acid *four* residues earlier; this repeated $i + 4 \rightarrow i$ hydrogen bonding is the most prominent characteristic of a α-helix. In other words i^{th} amino acid is hydrogen bonded to 5^{th} ($i+4^{th}$) amino acid peptide bond atom.

The torsion angles for α -helix are, ϕ= -60° and Ψ= -45° to -50°, which lies in the left side, lower quadrant of the Ramachandran plot.

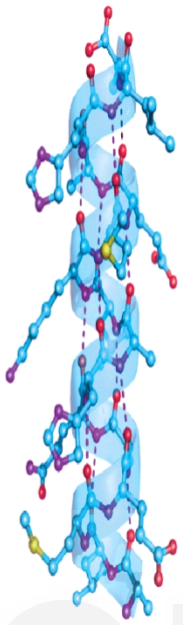


Fig. 4.3: Ball and stick model of α-helix.

Proline is absent in α-helices as it is no H-bond donor because of imino group and cannot form H-bond and is called helix disruptor.

When viewed from top, helix is seen in form of helical wheel in which side chains are protruded outside of the central core (Fig. 4.4).

An α -helix in polypeptides may form with either L-or D-amino acids but not with a combination of both the amino acids. α-helix occurs in α-keratin found in skin, hair, nails etc.

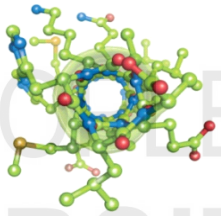


Fig. 4.4: View of α-helix from one end, looking down the longitudinal axis.

β -sheet: The β-sheet (Fig. 4.5) or β-pleated sheet is another form of regular secondary structure found in proteins. It is less common than the alpha helix.

β-sheets consist of β-strands connected by at least two or three backbone hydrogen bonds, forming a generally twisted, pleated sheet.

A β-strand is a stretch of polypeptide chain typically 3 to 10 amino acids long with backbone in an almost fully extended conformation.

β-sheet is formed by placing several β-strands adjacent to each other either in parallel or anti-parallel way (Fig. 4.6).

The backbone oxygen and nitrogen atoms of strands forming sheet are connected through hydrogen bonds in a β-sheet.

Side chains of consecutive amino acids are protruding above and below the plane of the sheet.

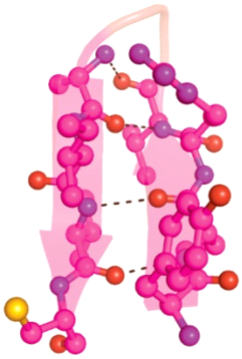
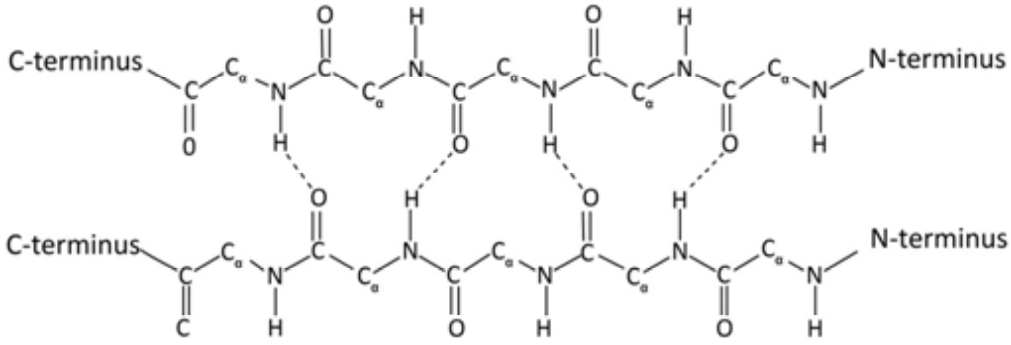


Fig. 4.5: β-Sheet.

Parallel β Sheet



The higher-level association of β–sheets has been implicated in formation of the protein aggregates and fibrils observed in many human diseases, notably the amyloidosis such as Alzheimer's disease.

Antiparallel β Sheet

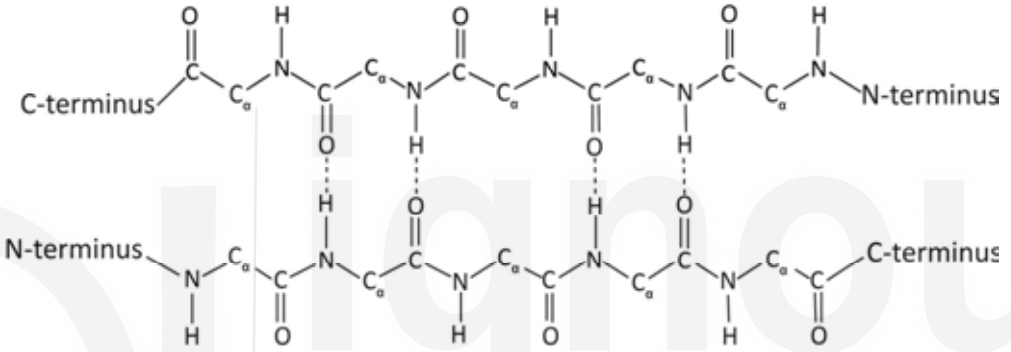


Fig. 4.6: Parallel and anti parallel β-Sheet Structures in Polypeptides.

The pleating causes the distance between $C^{\alpha i}$ and $C^{\alpha i+2}$ to be approximately 6 Å, rather than the 7.6 Å (2 × 3.8 Å) expected from two fully extended trans peptide.

The energetically preferred dihedral angles near (ϕ, ψ) = (–135°, 135°) for β-sheet structures diverge significantly from the fully extended conformation (ϕ, ψ) = (–180°, 180°).

Multiple β –sheets provide strength and rigidity in structural proteins such as silk fibroin.

SAQ 3

Answer the following:

a) Secondary structure of proteins is not maintained by hydrogen bond (True/False)

b) Alpha helix is the most common and stable form of polypeptide chain (True/False)

c) In alpha helix, the bonding is between the carbonyl oxygen of one peptide bond with the N-H group of a fourth amino acid away (True/False)

d) β-sheet is formed by placing several β-strands adjacent to each other either in ..... or ..... way (Fill in the Blanks)

4.3.3 Tertiary Structure

It refers to three-dimensional structure of protein molecules (Fig. 4.7). Primary sequence of amino acid folds into secondary structures and all secondary structural elements (helices, sheets, turns and coils) come together through various non covalent interactions and get folded into the tertiary structure. This structure is functionally active as well as compact globular in nature. The folding of secondary structure into tertiary is driven by the hydrophobic interactions, but the structure is stable only when the parts of a protein domain (section of protein that can exist independently and sufficient to perform its function) are locked into place by *specific* tertiary interactions, such as salt bridges, hydrogen bonds, and the tight packing of side chains and disulfide bonds.

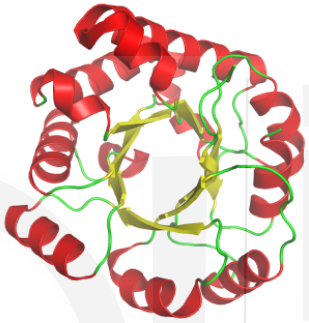


Fig. 4.7: Tertiary Structure of Protein.

The disulfide bonds are extremely rare in cytosolic proteins, since the cytosol is generally a reducing environment. In secreted proteins that do not spend time in the cytoplasm, disulfide bonds between cysteine residues help to maintain the protein's tertiary structure. In water soluble globular proteins such as myoglobin, tertiary interactions are stabilized by folding of non-polar amino acids in the hydrophobic interior away from the surrounding aqueous hydrophilic medium. The protein's aqueous surface is enriched by charged or hydrophilic residues. A variety of common and stable tertiary structures appear in a large number of proteins that are unrelated in both function and evolution– for example, many proteins are shaped like a TIM barrel named after the enzyme triosephosphateisomerase, a conserved metabolic enzyme. A TIM barrel fold is a tertiary structure which contains eight anti parallel β-strands in a circle, each connected with other through a α-helix in between.

4.3.4 Quaternary Structure

Quaternary structure refers to the arrangement of multiple folded protein in a multiple subunit complex.

It reflects the fourth level of complexity in protein structure (Fig. 4.8). Large proteins having molecular weight of > 100 kD consists of two or more polypeptide chains. Each of the polypeptide unit is called as subunit. These subunits may have identical or different primary structure. The spatial arrangement or association of these subunits constitutes quaternary structure. Proteins having more than one subunit are known as oligomers and the identical subunits are called as protomers. In most of oligomeric proteins,

protomer subunits are symmetrically arranged. Multi subunit proteins perform diverse functions in which one subunit may be regulatory in nature and other may be catalytic one. Hemoglobin, DNA polymerase and ion channels are well known examples of multiple subunit proteins with quaternary structure (Fig. 4.9).

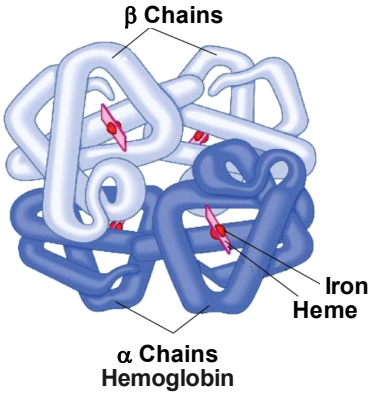


Fig. 4.8: Quaternary Structure of a hemoglobin protein.

(Source: website http://bio1151b.nicerweb.com)

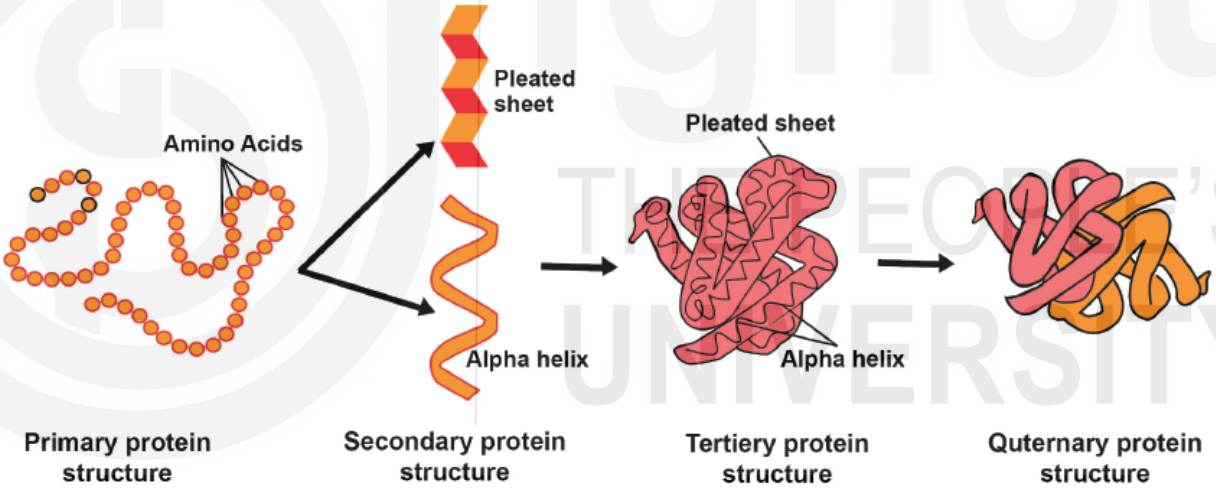


Fig. 4.9: Protein Architecture-Stages of Protein folding.

4.4 SIMPLE, CONJUGATED AND DERIVED PROTEINS

Proteins are often bound to ligands such as cofactor and coenzymes or to other macromolecules. On the basis of their composition, proteins are classified in three categories.

Simple proteins: These proteins are made up of amino acids joined together by peptide bonds. On treatment with acids, proteins are hydrolyzed and liberate the constituent amino acids. These proteins are further divided on the basis of their solubility. Pancreatic ribonuclease A (RNase A) is an example of simple protein as it yields only amino acids upon hydrolysis.

Conjugated proteins: These proteins contain permanently associated chemical components along with amino acids. These chemical components or non-protein substances conjugated with simple proteins are called as prosthetic group. The prosthetic group may be either a metal or a compound. A protein without its prosthetic groups is called as apoprotein. A protein molecule combined with its prosthetic group is called holoprotein. Conjugated proteins are further divided into several categories on the basis of chemical nature of the prosthetic group for example metalloproteins, glycoproteins, phosphoproteins, lipoproteins and nucleoproteins. Enzymes perform their biological activity with the help of prosthetic groups.

Derived proteins: These are not naturally occurring proteins and are derived from simple proteins by the action of heat, enzymes and chemical agents. This group also includes artificially produced polypeptides. Peptones, peptides and proteoses are well known examples of derived proteins.

SAQ 4

Answer the following:

a) The arrangement of four polypeptide subunits of the hemoglobin (Hb) represent ..... structure of protein. (Fill in the Blanks).

b) The non-protein substances conjugated with simple proteins are called as ..... (Fill in the Blanks).

c) Which term in protein structure refers to the arrangement of different protein subunits in a multiprotein complex.

d) Identify the term in protein structure that refers to the overall three dimensional shape of a protein.

4.5 SUMMARY

The term peptide is referred to short amino acid oligomers which lack a stable three dimensional structure. Polypeptide is a single linear unbranched chain of amino acids but often in an undefined conformation. Polypeptides folds in a stepwise process and it may involve multiple pathways. The term protein is referred to biomolecule in a stable conformation. Peptides and proteins can be distinguished on the basis of size. Each protein is made up of at least one polypeptide chain. Primary structure of amino acids leads to folding of polypeptides in secondary structures which are further folded in supersecondary structures. The most common secondary structures in a protein are α-helix and α-sheets. Three dimensional shape of protein molecule is its tertiary structure. The shape of protein molecule is determined by several bonding interactions between side chains of the amino acids. The three dimensional structures of proteins reflect its function. Several proteins are made up of multiple polypeptide chains often called as protein subunits. Arrangement of two or more polypeptide subunits in a protein in three

dimensional complexes constitutes quaternary structure. Proteins participate in almost every biochemical process in the body. Protein misfolding is the molecular basis for wide range of human diseases such as Type 2 diabetes, Alzheimer's disease, Huntington and Parkinson's diseases.

4.6 TERMINAL QUESTIONS

1. In a large polypeptide predict the folding of following amino acids pairs in the protein-outside or interior and why?

   a) Pair of valine amino acids b) Aspartic acid and serine c) Alanine and isoleucine

2. What is an oligopeptide? How it is different from a polypeptide?

3. Does proteins with same number of different amino acids identical?

4. Do you agree with the statement "Protein molecules are immensely diverse"? If yes, give your rationale.

4.7 ANSWERS

Self-Assessment Questions

1. a) No b) No c) Yes
2. a) peptides, b) False, c) True, d) Edman Degradation Reaction Method
3. a) False, b)True, c) True, d) parallel and antiparallel.
4. a) quaternary structure b) prosthetic group c) quaternary structure d) tertiary structure.

Terminal Questions

1. The amino acids of a polypeptide affect the shape of the protein. The pairs will behave as given below:

   a) Nonpolar amino acids such as valine will be repelled by water, so portion of polypeptide chain containing these amino acids will be folded toward the interior away from the hydrophilic environment.

   b) Serine is a polar amino acid and aspartic acid is a charged in nature. Pair of these amino acids will be attracted to water, so portion of polypeptide chain would be facing toward the outside.

   c) Non polar amino acids such as alanine and isoleucine are water repelling, therefore portion of polypeptide chain would be folded in toward the interior away from the watery environment.

2. Oligopeptide is a peptide made of few amino acids (oligo = few) while polypeptides are long chain peptides with many amino acids.

3. No not necessarily. The proteins having same number of different amino acids may or may not be identical because sequences in which amino acids are connected may be different. The protein folding may vary with the arrangement of the sequence of amino acids. Numeric similarity for constituent amino acids does not guarantee identical proteins.

4. The genetic code of living beings codifies twenty different amino acids. Several numerous combinations of amino acids are part of the polypeptide chain and proteins. Therefore protein molecules are immensely diverse and can assume many different configurations to play varied roles in cells and tissues.

4.8 FURTHER READINGS

1. Albert L. Lehninger: Principles of Biochemistry, Worth Publishers, Inc. New York, 1984.

2. Harper's Illustrated Biochemistry, 29e. Robert K. Murray, David A Bender, Kathleen M. Botham, Peter J. Kennelly, Victor W. Rodwell, P. Anthony Weil, USA.

3. Donald J Voet Principles of Biochemistry, Jophn Wiley and Sons, Inc, USA.

4. J. L. Jain: Fundamentals of Biochemistry, S. Chand & Company Ltd. India.

5. U. Satyanarayana and U. Chakrapani: Biochemistry, UBS Publishers Distributors Pvt Ltd. Kolkatta, India.

6. Thomas M. Devlin: Textbook of Biochemistry, John Wiley and Sons, Inc.Danvers, MA, USA.

ignou
THE PEOPLE'S
UNIVERSITY