NATIONAL OPEN UNIVERSITY OF NIGERIA

SCHOOL OF SCIENCE AND TECHNOLOGY

COURSE CODE: BIO213

COURSE TITLE: CHEMISTRY OF AMINO ACIDS AND PROTEINS

COURSE GUIDE

BIO213
CHEMISTRY OF AMINO ACIDS AND PROTEINS

Course Team Mr. Timothy Bulus (Writer) - KSU

NATIONAL OPEN UNIVERSITY OF NIGERIA

National Open University of Nigeria
Headquarters
14/16 Ahmadu Bello Way
Victoria Island
Lagos

Abuja Office
5, Dar es Salaam Street
Off Amina Kano Crescent
Wuse II
Abuja.

E-mail: centrainfo@nou.edu.ng
URL: www.nou.edu.ng

Published By:
National Open University of Nigeria

First Printed 2011

ISBN: 978-058-285-1

CONTENTS **PAGE**

Introduction

Biology is the science of living things. The functions of various cells are only possible because of molecular interaction among several biosubstances that exist in these cells. The major classes of these biosubstances/biomolecules (also called macromolecules because of their sizes) include: proteins, carbohydrates, fats and oil, and nucleic acids. The understanding of their structures will aid you to understand their functions. Proteins, which constitute a very large family of macromolecules, will be the focus of this course.

This course, the **Chemistry of Amino Acids and Proteins**, covers mainly the chemistry of amino acids (the building blocks of proteins) and their polymers (Peptides and Proteins). The course will help you to appreciate how the structural features of these three molecules affect biological activity of an organism. The purpose underlying the study of this course, therefore, is to develop a greater comprehension of how the structures of biomolecules affect their overall cellular functions.

What you will Learn in this Course

The course consists of study units and a Course Guide. This Course Guide tells you briefly what the course is about, what course material you will be using and how you can use the materials. In addition, it sets some general guidelines on the amount of time you are likely to spend on each unit of the course in order to complete it successfully.

The Course Guide gives you guidance in respect of your Tutor-Marked Assignment which would be made available in an assignment folder.

There will be regular tutorial classes that are related to the course. It is advisable for you to attend these tutorial sessions. The course will prepare you for the challenges you will meet in trying to link structures and functions of biomolecules.

Course Aims

The aim of this course is simply to provide you with the understanding of the structure, function, relationship of proteins/peptides and their building blocks (Amino acids).

Course Objectives

In order to achieve this aim, the course has a set of objectives. Each unit has specific objectives which are included at the beginning of the unit.

You are expected to read these objectives before you study the unit.

You may wish to refer to them during your study to check on your progress. You should always look at the unit objectives after completion of each unit. By doing so, you would have followed the instructions in the unit.

Below are the comprehensive objectives of the course as a whole. By meeting these objectives, you should have achieved the aims of the course. Therefore, after going through this course you should be able to:

- state the relationship between amino acids, peptides and proteins
- explain the basic structural features of commonly occurring amino acids
- relate properties of amino acids to their functional groups
- explain how amino acids are linked via peptide bonds
- identify similarities of properties in amino acids and peptide
- differentiate between peptides and proteins
- explain the various criteria used for the separation/purification and sequencing of peptides and proteins
- explain the roles played by covalent and non-covalent forces in stabilising different level of protein structure
- identify the various functional and structural classes of proteins
- explain the nature of enzymes and enzyme-catalysed reactions

Working through this Course

To complete this course, you are required to read each study unit, read the textbooks and read other materials which may be provided by the National Open University of Nigeria.

Each unit contains Self Assessment Exercise and at certain points in the course, you would be required to submit assignments for assessment purposes.

At the end of the course there is a final examination. The course should take you about a total of 17 weeks to complete. You will find below all the components of the course, what you have to do and how you should allocate your time to each unit in order to complete the course on time and successfully too.

This course entails that you spend a lot of time to read and practice. For easy understanding of this course, I will advise that you avail yourself the opportunity of attending the tutorial sessions where you will have the opportunity to compare your knowledge with that of other people, and also have your questions answered.

The Course Materials

The main components of this course are:

1. The Course Guide
2. Study Units
3. References/ Further Reading
4. Assignments
5. Presentation Schedule

Study Units

The study units in this course are as follows:

Module 1 Amino Acids

Unit 1 Amino Acids as Building Blocks of Proteins
Unit 2 Classifications of Amino Acids
Unit 3 Properties of Amino Acids
Unit 4 Chemical Reactions of Amino Acids
Unit 5 pH, pK_a and Buffering Capacity of Amino Acids

Module 2 Peptides

Unit 1 Peptides: Formation and Nomenclature
Unit 2 Properties, Examples and Functions of Biological Peptides
Unit 3 Separation of Peptides I
Unit 4 Separation of Peptides II
Unit 5 Sequencing of Peptides

Module 3 Proteins

Unit 1 Proteins:Nature, Properties, Examples and Biological Functions
Unit 2 Structural Levels of Proteins
Unit 3 Stability of Proteins- The Role of Noncovalent Forces
Unit 4 Classification of Proteins
Unit 5 Enzymes, Vitamins and Co-enzymes

For easy presentation of the course to you, we have divided the course into three modules of five units each; with each module covering an important aspect of the course. Module one focuses on amino acids.

Module two and three cover the polymers, the peptides and the proteins respectively.

In Module One, the first unit focuses on the general structural features of amino acids. The second unit deals with the classification of amino acids. The third and fourth units discuss the key properties and the chemical reactions found among amino acids respectively. Unit five covers pH, pK_a and the buffering potential of amino acids.

Module Two is mainly concerned with peptides as a polymer of amino acids. In unit one of this module, the definition of peptide, formation of peptide bond and the nomenclature of peptides are dealt with. The second unit deals with the properties, examples and functions of biologically active peptides. The third and fourth units of this module are concerned with the various techniques for the separation/purification of peptides. The last unit covers the sequencing of peptides.

Module Three focuses on proteins as polymers of amino acids. The nature, properties, examples and biological function are discussed in unit one. Unit two deals with the different hierarchy of protein structure. Unit three focuses on folding, stability and unfolding (denaturation) of proteins with particular reference to the role played by weak noncovalent forces. The fourth unit covers the classification of proteins while the fifth unit focuses on enzymes, a major class of protein.

Each unit consists of one or two weeks' work and includes an introduction, objectives, reading materials, conclusion, summary, tutor-marked assignment, references and further reading. The unit directs you to work on exercises related to the required reading. In general, these exercises test you on the materials you have just covered or require you to apply it in some way. They assist you to evaluate your progress and to reinforce your comprehension of the material. The TMAs will help you in achieving the stated learning objectives of the individual units and of the course as a whole.

Presentation Schedule

Your course materials have important dates for early and timely completion and submission of your TMAs and attending tutorials. You should remember that you are required to submit all your assignments by the stipulated time and date. You should guard against falling behind in your work.

Assessment

There are three aspects to the assessment of the course. The first is made up of Self Assessment Exercises, the second consists of the Tutor-Marked Assignments and the third is the written examination/end of course examination.

You are advised to practise the exercises. In tackling the assignments, you are expected to apply information, knowledge and techniques you gathered during the course. The assignments must be submitted to your facilitator for formal assessment in accordance with the deadlines stated in the presentation schedule and the assignment file. The work you submit to your tutor for assessment will account for 30% of your total course work. At the end of the course you will need to sit for a final or end of course examination. The examination will account for 70% of your total course mark.

Tutor-Marked Assignment

The TMA is a continuous assessment component of your course. It accounts for 30% of the total score. You will be given four (4) TMAs to answer. Three of these must be answered before you are allowed to sit for the end of course examination. The TMAs will be given to you by your facilitator and returned after you have done the assignment.

Assignment questions for the units in this course are contained in the assignment file. You will be able to complete your assignment from the information and material contained in your reading, references and study units. However, it is desirable in all degree level of education to demonstrate that you have read and researched more into your references, which will give you a wider viewpoint and may provide you with a deeper understanding of the subject.

Make sure that each assignment reaches your facilitator on or before the deadline given in the presentation schedule and assignment file. If, for any reason, you cannot complete your work on time, contact your facilitator before the assignment is due to discuss the possibility of an extension. Extension will not be granted after the due date unless in exceptional circumstances.

Final Examination and Grading

The end of course examination for this course has a value of 70% of the total course work. The examination will consist of questions, which will reflect the type of self-testing, practice exercise and tutor-marked assignment problems you have previously encountered. All areas of the course will be assessed.

You should use the interval between finishing the last unit and sitting for the examination to revise the whole course. You will find it useful to review your self-test, TMAs and comments on them before the examination. The end of course examination covers information from all parts of the course.

Course Marking Scheme

Assignment	**Marks**
Assignments 1-4	Four assignments, best three marks of the four count at 10% each-30% of course marks
End of course examination	70% of the overall course marks
Total	100% of course materials

Facilitators/Tutors and Tutorials

There are 16 hours of tutorials provided in support of this course. You will be notified of the dates, time and location of these tutorials as well as the name and phone number of your facilitator, as soon as you are allocated a tutorial group.

Your facilitator will mark and comment on your assignments, keep a close watch on your progress and any difficulties you might face and provide assistance to you during the course. You are expected to mail your Tutor-Marked Assignments to your facilitator before the schedule date (at least two working days are required). They will be marked by your tutor and returned to you as soon as possible.

Do not delay to contact your facilitator by telephone or e-mail if you need assistance. The following might be circumstances in which you would find assistance necessary. Thus, you would have to contact your facilitator if:

- you do not understand any part of the study or the assigned readings
- you have difficulty with the self-tests
- you have a question or problem with an assignment or the grading of an assignment.

You should endeavour to attend the tutorials. This is the only chance to have face-to-face contact with your course facilitator and to ask questions which are answered instantly. You can raise any problem encountered in the course of your study. To benefit from the course tutorials, you should prepare a question list before attending them. You will learn a lot from participating actively in discussions.

Summary

Science is the systematic study of nature. The **Chemistry of Amino Acids and Proteins** is a course that provides you with the molecular explanation of some concepts in living systems. The course focusses on

the understanding of the basic chemistry of amino acids and their polymers, peptides and proteins. The course will answer questions such as: What is the chemical nature of these molecules? What are their properties? How are they related structurally and functionally? How are the polymers formed from their monomeric units? What are some modern separation techniques employed in the study (purification and sequencing) of these biomolecules? Can they by any means be classified?

Apart from the academic benefit of studying this course, you will find the principles taught in this course quite helpful in your everyday life. It is also hoped that this course will improve your quality of decision, particularly as is affect your health and nutrition.

Finally, I wish you success in the course and I hope that as you give your best to this course, you will find it both interesting and useful.

Course Code BIO213
Course Title Chemistry of Amino Acids and Proteins

Course Team Mr. Timothy Bulus (Writer) - KSU

NATIONAL OPEN UNIVERSITY OF NIGERIA

National Open University of Nigeria
Headquarters
14/16 Ahmadu Bello Way
Victoria Island
Lagos

Abuja Office
5, Dar es Salaam Street
Off Amina Kano Crescent
Wuse II
Abuja.

E-mail: centrainfo@nou.edu.ng
URL: www.nou.edu.ng

Published By:
National Open University of Nigeria

First Printed 2011

ISBN: 978-058-285-1

CONTENTS **PAGE**

MODULE 1 AMINO ACIDS

UNIT1 AMINO ACIDS AS BUILDING BLOCKS OF PROTEINS

CONTENTS

1.0 INTRODUCTION

Proteins as one of the major classes of foods (others being carbohydrates, fat and oils and vitamins) are very essential for the normal functioning of all biological systems. In living organisms, proteins constitute one of the major biomolecules and they are known to perform various functions such as transportation of molecules, structural component of hair and nails, and other catalytic roles e.t.c. All proteins are polymers of only 20 amino acids. This unit examines some key structural features of amino acids.

2.0 OBJECTIVES

By the end of this unit, you should be able to:

- define amino acids
- explain how amino acids occur in nature
- use the concepts of chirality to describe the structure of amino acids
- write the structure of some common amino acids.

3.0 MAIN CONTENT

3.1 Definition of Amino Acids

Essentially, amino acids are the building blocks of proteins.

3.2 Occurrence of Amino Acids in Nature

Most amino acids occur naturally in conjugated form as components of proteins. Those that occur commonly in proteins are called common amino acids. These are twenty in number. Table 1 gives the names, one-letter and three-letter abbreviations of the twenty common amino acids.

Table 1: The 20 Common Amino Acids

Amino Acids	Three-letter abbreviations	One-letter symbol
1. Glycine	Gly	G
2. Alanine	Ala	A
3. Proline	Pro	P
4. Valine	Val	V
5. Leucine	Leu	L
6. Isoleucine	Ile	I
7. Methionine	Met	M
8. Phenylalanine	Phe	F
9. Tyrosine	Tyr	Y
10. Tryptophan	Trp	W
11. Serine	Ser	S
12. Threonine	Thr	T
13. Cysteine	Cys	C
14. Asparagine	Asn	N
15. Glutamine	Gln	Q
16. Lysine	Lys	K
17. Arginine	Arg	R
18. Histidine	His	H
19. Aspartate	Asp	D
20. Glutamate	Glu	E

Certain amino acids and their derivatives, however, are not found in proteins. They nonetheless perform useful biological roles. These are called uncommon amino acids as shown in Table 2.

Table 2: Uncommon Amino Acids

Uncommon amino acids	Biological function
γ –Aminobutyric acid (GABA)	Transmit signal in the nerves
Histamin	Transmit signal in the nerves
Serotonin	Transmit signal in the nerves
β -Alanine	Component of panthothenic acid (a vitamin)
Epinephrine (also called Adrenaline)	Acts as a hormone
Penicillamine	Constituent of penicillin antibiotics
Ornithine	Intermediate for urea formation
Homoserine	Metabolic intermediates
Citrulline	Precursor of arginine

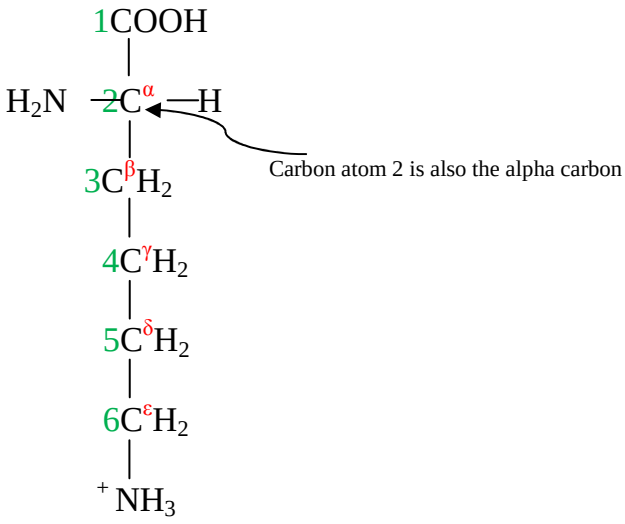
3.3 Basic Structure of Amino Acids

For you to fully grasp the structure of amino acids, you will need to know:

(1) how carbon atoms in organic compounds are named;
(2) the concept of chirality in organic compounds; and
(3) the concept of stereoisomerism

3.3.1 How Carbon Atoms in Amino Acids are Named

Carbon atoms in an amino acid, just as in any organic compound, are conventionally named either by using Greek letters such as alpha (α), beta (β), gama (γ), delta (δ), epsilom (ε) e.t.c. or by a simpler Arabic numerals numbering convention (1, 2, 3, 4…). Carefully study the structure of lysine given below and you will see that carbon atom 2 can also be referred to as alpha carbon atom (C_α).

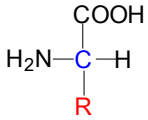


Structure of lysine (Lys) showing various ways of numbering the carbon atoms

3.3.2 Chirality in Amino Acids

Chiral compounds are compounds which possess at least one chiral carbon atom or chiral centre. A chiral carbon or centre is a carbon atom to which four different functional groups are covalently linked. For a typical amino acid, the four different functional groups include:

1. Carboxyl group (-COOH) or its carboxylate ion form ($-COO^-$)
2. Amino group ($-NH_2$) or its ammonium ion form ($-NH_3^+$)
3. Hydrogen atom (H)
4. A variable side chain (Also called R group). The **R group** gives each amino acid its identity. All amino acids with the exception of glycine possess a chiral cetre at the alpha carbon (C_α).



You should note that the **R group** (red in the structure above) must not be –COOH, H_2N, or H for the alpha carbon (blue) to be a chiral centre.

Again, most of the amino acids (except proline) are alpha amino acids, which means that the amino group is linked to the same carbon atom to which the carboxyl group is attached (the alpha carbon, C_α).

3.3.3 The Concept of Stereoisomerism

Stereochemistry of biomolecules deals with the three-dimensional arrangement of constituent atoms of the molecule in space. Each fixed spatial arrangement of atoms in a molecule is called the molecule's **configuration**. A Carbon compound may have different forms (configurations) in which the various atoms are arranged differently in space. Such forms of the compound are referred to as its **stereoisomers**.

You have already learnt that the alpha carbon (C_α) of amino acids, due to its tetravalency, is easily bonded to four atoms. When all these atoms are different, the carbon atom is said to be chiral or asymmetric and the number of stereoisomers that can occur in such a chiral compound is mormally given by $\boldsymbol{2^n}$, where ***n*** is the number of chiral carbon(s) in the compound.

For example, Glycine has only one stereoisomer because its *n=0,*
Alanine has one chiral carbon hence; the number of its tereoisomers is 2^1 i.e 2.

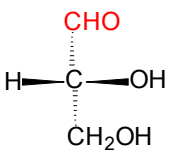
In stereochemistry, stereoisomers that are mirror images of each other are called **enantiomers** while pairs of stereoisomers that are not mirror images of each other are called **diastereomers.** You may consult a chemistry text for in-depth notes on this. Meanwhile, let us consider the D- and L- configurations of amino acids.

D- and L-Configurations of Amino acids

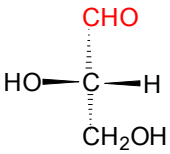
In 1891, Emil Fischer proposed the **D, L** system of nomenclature for asymmetric/chiral compounds (such as amino acids and sugars). The system specifies the absolute configuration of the substituents around the chiral centre with reference to **glyceraldehyde**, a three-carbon sugar.

By this system, stereoisomers that have configuration similar to that of D-glyceraldehyde are designated L, and those with configuration identical to L-glycceraldehyde are designated L. This can be done by simply aligning the carboxyl group (COOH) of the amino acid with the aldehyde group (CHO) of glyceraldehydes as illustrated for Alanine and Tyrosine below:

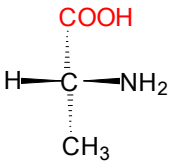
Using Perspective Formulas



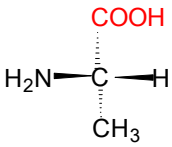
D-Glyceraldehyde



L-Glyceraldehyde



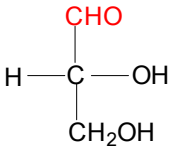
D-Alanine



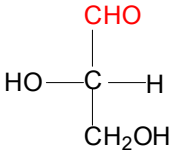
L-Alanine

Please note that the amino acids which occur in proteins are usually the L- isomers and only living cells produce this form. This is because biochemical reactions are stereospecific. For example, enzymes are specific in their actions, recognising only one isomer and not the other.

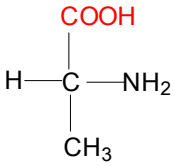
Using Fischer Projection Diagrams



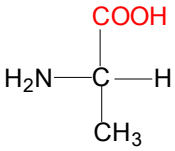
D-Glyceraldehyde



L-Glyceraldehyde



D-Alanine



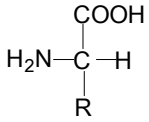
L-Alanine

In short, an amino acid is L if its α-amino group is oriented to the left, and is D if it is oriented to the right.

3.3.4 Structures of Some Amino Acids

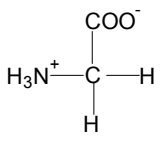
For you to understand the structures of the common amino acids and be able to write them yourself, it is important that you remember the following key points:

- All amino acids are similar structurally
- They all have carboxyl, amino group, hydrogen atom and a side chain (R group) covalently linked to the alpha carbon atom as shown below.

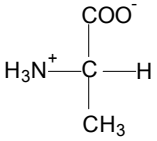


- They differ mainly in the nature of their side chains (R groups); the R groups in amino acids vary in structure, size/length and electric charge.
- Each amino acid can be remembered for a particular feature in its side chain: Glycine, the amino acid with the simplest structure, has H as its side chain; Alanine contains a methyl group ($-CH_3$)in place of R; Valine has an isopropyl R group; Phenylalanine and tyrosine have aromatic /benzene ring in their side chain; tryptophan has a heterocyclic structure with an indole group; Cysteine and methionine both have sulphur as part of their side chain e.t.c.
- At physiological pH, amino acids are ionic in nature with their amino and carboxyl groups preferring ammonium ($^+NH_3$) and carboxylate (COO^-) states rather than amino (NH_2) and carboxyl (COOH) states respectively. Amino acids are therefore drawn showing these ionic forms.

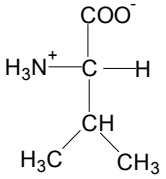
You shall learn in detail the structures of all the amino acids in unit 2 of this module. For now, use the information above to learn to write the Fischer projection structures of the following simple amino acids. They are only six in number.



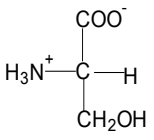
L-Glycine (Gly)



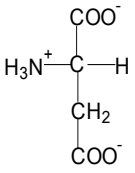
L-Alanine (Ala)



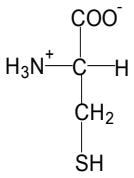
L-Valine (Val)



L-Serine (Ser)



L-Aspartate (Asp)



L-Cysteine (Cys)

4.0 CONCLUSION

Amino acids are the building blocks of proteins. All common amino acids share some identical structural features.

5.0 SUMMARY

In this unit, you have learnt that:

- amino acids are the building blocks of proteins
- twenty amino acids occur commonly as part of proteins, although some uncommon amino acids do also occur in nature
- each of the twenty common amino acids is identified with the aid of a three letter or one letter abbreviation
- carbon atoms in amino acids are identified using Arabic numerals or Greek lettering
- all common Amino acids with the exception of glycine are chiral compounds
- the structures of the common amino acids are essentially similar except with respect to their side chains.

6.0 TUTOR-MARKED ASSIGNMENT

1a. What are amino acids?
b. Name any five common and five uncommon amino acids. Also give their three-letter abbreviations.

2a. Is Glycine a chiral compound? Give your reason.
b. With the aid of a diagrammatic illustration only, identify a chiral centre in a named amino acid.
c. Name any four functional groups that could be found in amino acids.

3. Write the structures of the following amino acids:

a. Alanine.
b. Serine.
c. Cysteine.

7.0 REFERENCES/FURTHER READING

Thomas, M. Devlin (ed) (2006). *Textbook of Biochemistry with Clinical Correlations*. Canada: Wiley-Liss.

David, L. Nelson & Michael M. Cox (2005). *Lehninger Principles of Biochemistry*. New York: W. H. Freeman and Company.

UNIT 2 CLASSIFICATION OF AMINO ACIDS

CONTENTS

1.0 INTRODUCTION

In Unit 1, you learnt that amino acids are the building block of proteins.

In addition, you learnt to classify amino acids into two groups on the basis of their occurrence in nature. The common amino acids occur mainly as components of proteins while the uncommon ones, though not part of proteins, have different biologically important roles. In this unit, you will learn to classify amino acids on the basis of other criteria.

2.0 OBJECTIVES

By the end of this unit, you should be able to:

- mention some criteria for classifying amino acids
- classify amino acids on the basis of their structures
- explain the structures of all the common amino acids
- identify polar and nonpolar amino acids
- classify amino acids on the basis of their nutritional requirement.

3.0 MAIN CONTENT

3.1 Criteria for Classification of Amino Acids

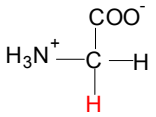
Each of the twenty common amino acids of protein has a unique side chain but some amino acids could share identical features. Amino acids have been classified on the basis of the following criteria:

i. Functional group of the side chain.
ii. Polarity
iii. Nutritional requirements

3.2 Classification Based on Structure

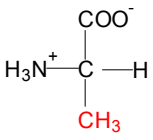
On the basis of the type of functional groups present in their side chains, amino acids can be placed into nine classes. Note that you are expected to study the structure of each amino acid taking into consideration their uniqueness.

1. Amino acid having hydrogen atom as R group-Gly

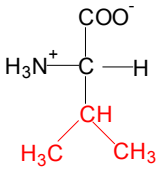


L-Glycine

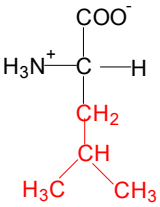
2. Amino acids possessing unsubstituted aliphatic chain as their R group-Ala, Val, Leu,
   Ile



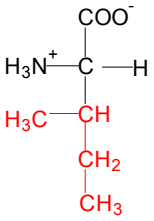
L-Alanine



L-Valine

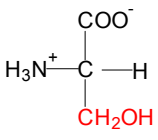


L-Leucine

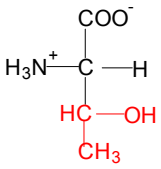


L-Isoleucine

3. Amino acids whose R groups possess aliphatic chain bearing a hydroxyl group-Ser,
   Thr

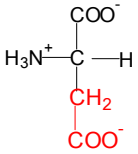


L-Serine

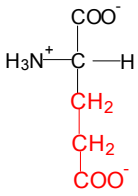


L-Threonine

4. Amino acids whose R group is an aliphatic chain terminating in an acidic carboxyl
   Group-Asp, Glu

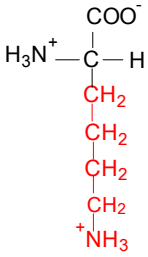


L-Aspartate

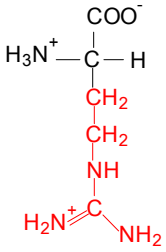


L-Glutamate

5. Amino acids whose side groups are aliphatic chain terminating in basic amino group-
   Arg, Lys

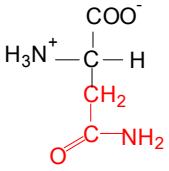


L-Lysine

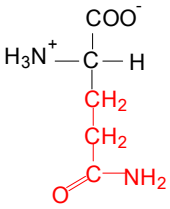


L-Arginine

6. Amino acids whose side chains are aliphatic chain terminating in amide group-Asn, Gln

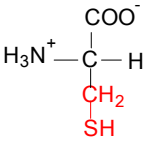


L-Asparagine

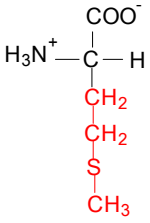


L-Glutamine

7. Amino acids with sulfur-containing aliphatic R groups- Cys, Met

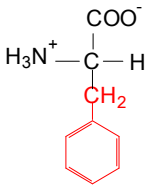


L-Cysteine

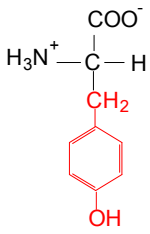


L-Methionine

8. Amino acids whose side chains terminate in an aromatic ring- Phe, Tyr

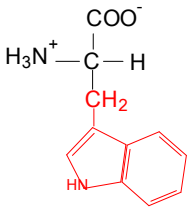


L-Phenylalanine

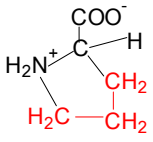


L-Tyrosine

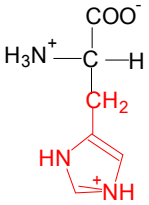
9. Amino acids whose side chains terminate in a heterocyclic ring- Trp, Pro, His



L-Tyrptophan



L-Proline



L-Histidine

3.3 Classification Based on Polarity

Amino acids could also be classified according to the polarities of their side chains. Polarity is perhaps the most useful criterion of amino acid classification. This is because protein folding (a concept you will learn about in module 3 of this course) is governed by the tendency of amino acid residues of protein to interact with water. In protein conformation, hydrophobic (a term used to mean water-hating) side chains are removed from contact with water while the hydrophilic (water-loving) ones interact freely with aqueous environment. These interactions are in harmony with the second law of Thermodynamics.

This classification scheme places each amino acid into either of the two major classes below:

(1) Those with nonpolar R groups
(2) Those with polar R groups

3.3.1 Amino Acids with Nonpolar Side Chains

Glycine, Alanine, Valine, Leucine, Isoleucine, Methionine, Proline, Phenylalanine and Tryptophan have nonpolar R group. The table below gives the description of the nature of their individual side chains which account for their nonpolarity.

Nonpolar Amino acid	Nature of R group
Gycine, Alanine, Valine, Leucine, Isoleucine	Have aliphatic side chain
Methionine	Has a thio ether end
Proline	Has a cyclic pyrrolidine side chain
Phenylalanine	Has a phenyl moiety in its side chain
Tryptophan	Has an aromatic indole group

3.3.2 Amino Acids with Polar Side Chains

This class is further divided into two sub-classes:

(a) Those with uncharged polar R groups

Serine and **Threonine** - are polar because of their hydroxyl groups.
Asparagine and **Glutamine** – the polarity is due to the presence of amide-bearing R groups.
Tyrosine – has an OH functional group attached to benzene ring. This OH is hydrophilic.
Cysteine – its thiol or (SH) group is responsible for its polarity.

(b) Those with charged polar R groups

At physiological pH values, **Lysine**, **Arginine** and **Histidine** are positively charged due to their terminal ammonium, guanidinium, and the imidazolium groups respectively. These three amino acids are basic amino acids since their side chains have net positive charges at neutral pH.

Aspartic acid and Glutamic acid are negatively charged above pH 3. These amino acids are also referred to as acidic amino acids because their side chain contain acidic carboxyl group.

3.4 Classification Based on Nutritional Requirements

Some, but not all, of the 20 common amino acids can be synthesized by the body and as such, they may not be included in diet. On the basis of their requirements in our diet, amino acids have been grouped into:

(a) Essential amino acids, and
(b) Nonessential amino acids

Those amino acids that can be synthesized in the body are called **non-essential amino acids** as they may not be included in the diet. On the

other hand, those amino acids that the body cannot synthesize *de novo* (from existing molecules) must be supplied in the human/animal diet.

Such amino acids are referred to as the **essential amino acids**. If only one of these essential amino acids is lacking from the diet, the synthesis of new protein by the body will be affected. Most bacteria and plants can synthesize all the common amino acids. The table below gives a summary of essential and nonessential amino acids for human/Albino rats.

Nonessential Amino acids	Essential Amino acids
Alanine	Histidine
Asparagine	Isoleucine
Aspartate	Leucine
Glutamate	Lysine
Serine	Methionine
	Phenylalanine
	Threonine
	Tryptophan
	Valine

4.0 CONCLUSION

In this unit, you have learnt that the chemical nature of amino acids' side chain provides a useful basis for its classification.

5.0 SUMMARY

In this unit, you have learnt that:

- amino acids vary in the types of functional groups present in their side chains
- some amino acids have side chains that allow them to interact with water (hydrophilic) while those of others make them to avoid water molecules (hydrophobic)
- amino acids can also be classified as essential and non-essential depending on their requirement in diet.

6.0 TUTOR-MARKED ASSIGNMENT

1a. Name the amino acids that have sulfur in their R groups.
b. Why do you think glutamate and aspartate are acid amino acids?

2a. Sate any five functional groups that could account for polarity of amino acids.
b. Classify the following amino acids as polar or non polar:
Glu, Phe, Val, Cys, Tyr, Asn.

3. What do you understand by the terms essential and nonessential amino acids?

7.0 REFERENCES/FURTHER READING

Thomas, M. Devlin (Ed) (2006). *Textbook of Biochemistry with Clinical Correlations*. Canada: Wiley-Liss.

David, L. Nelson & Michael, M. Cox (2005). *Lehninger Principles of Biochemistry*. New York: W. H. Freeman and Company.

UNIT 3 PROPERTIES OF AMINO ACIDS

CONTENTS

1.0 INTRODUCTION

Like any other organic compound, amino acids exhibit some physical and chemical properties that stem from their overall chemical structures.

In this unit, we shall focus basically on their general physical properties:

Their optical activity, absorption of light by amino acids and their ability to exist in ionic states.

2.0 OBJECTIVES

By the end of this unit, you should be able to:

- discuss how amino acids respond to heat
- state which amino acids are optically active and why
- list the amino acids that absorb at ultraviolet light and account for this property; and
- explain the ionic properties of amino acids.

3.0 MAIN CONTENT

3.1 Properties of Amino Acids

Amino acids possess a number of physical and chemical properties that are closely related to their overall structural features.

3.2 Physical Properties

The physical properties of amino acids include:

i. At ordinary temperature amino acids are white crystalline solids.
ii. All amino acids have high melting points (more than 200 ^{o}C).
iii. They decompose when heated at high temperature, rather than melting.
iv. They are stable in aqueous solution, and with few exceptions, they can be heated as high as 120 ^{0}C for a short periods without decomposition, even in acid or alkaline solution.
v. All amino acids are soluble in water and alcohol but insoluble in nonpolar solvents.

3.3 Optical Activity of Amino Acids

Apart from glycine, all other amino acids possess a centre of asymmetry at the α-carbon (chiral centre). Such molecules are optically active, i.e. they are able to rotate the plane of plane- polarized light when they are placed in a **polarimeter,** an instrument which measures the specific rotation of an optically active substance (see fig.1 below). Amino acid is said to be **dextrorotatory** if it rotates the plane to the right (clockwise from the point of view of the observer). On the other hand, if the direction of rotation is to the left (anticlockwise), the amino acid is said to be **laevorotatory**. Dextrorotatory molecules are designated by the prefix '**+** or ***d***' while laevorotatory enantiomers are designated with the prefix '**-** or ***l***'.

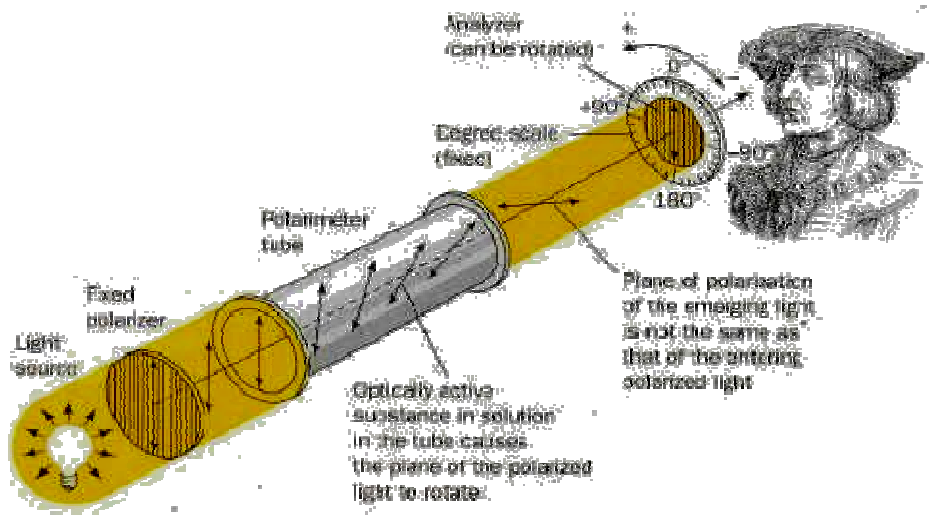


Fig. 1: Polarimeter
Source: *Biochemistry***,** Voet and Voet

3.4 Absorption of Light by Amino Acids

Phenylalanine, tryptophan and tyrosine absorb light strongly in the ultraviolet region (λ=100 to400 nm) as shown in figure 2 below. This property is attributable to the presence of aromatic nucleus in their side chains. As you will learn in modules 2 and 3, the presence of these amino acids in polypeptides, is being used to quantify polypeptides. Amino acids are not known to absorb visible light (λ=400 to 800 nm) and are therefore colourless.

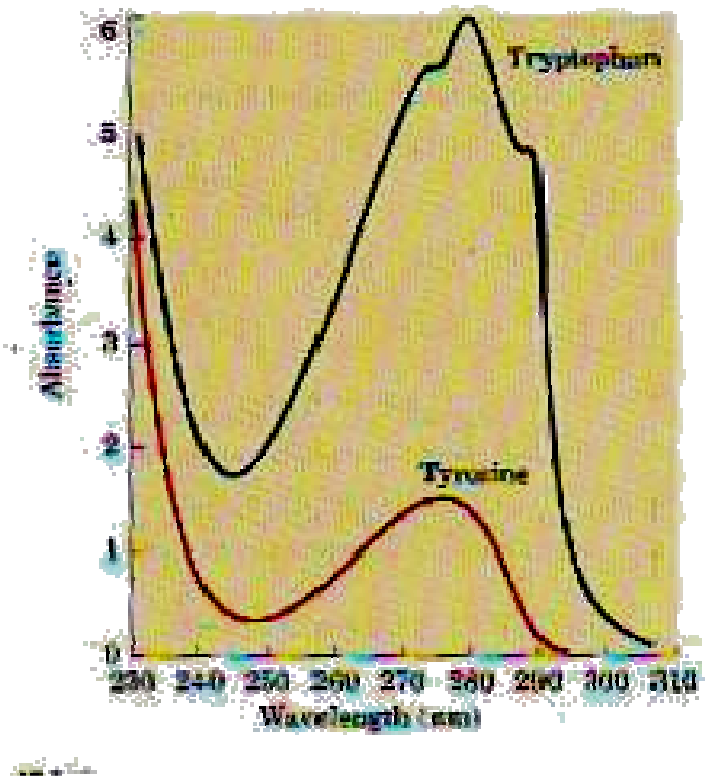
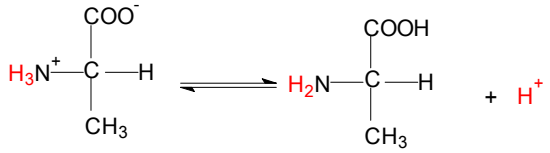


Fig. 2: **Absorption of UV Light by Aromatic Amino Acids**
Source: Lehninger, *Principles of Biochemistry*

3.5 Acid – Base Property of Amino Acids

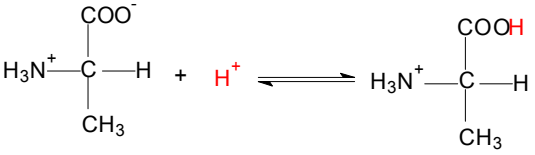
In solution, an amino acid exhibits ionic property. Both the amino and the carboxyl groups attached to its alpha carbon exist in their ionic states (COO^-, $^+NH_3$), thereby making it a dipolar ion, or a **zwitterion** (a German term for 'hybrid ion'). This zwitterion can act as an acid or as a base.

As an acid (a proton donor):



zwitterionic form of L-Ala

As a base (a proton acceptor):



zwitterionic form of L-Ala

3.5.1 Titration Curve of Simple Amino Acids

A simple amino acid, like glycine or alanine, is monoamino monocarboxylic α-amino acid. Such amino acid is diprotic. i.e., its fully protonated form ($^{+}H_3N$-CH_2-COOH) can yield two protons- one from the -COOH group and the other from the -NH_3^+ group.

In the course of acid-base titration, these protons are gradually being removed. Figure 3 below is a characteristic titration curve of a diprotic amino acid.

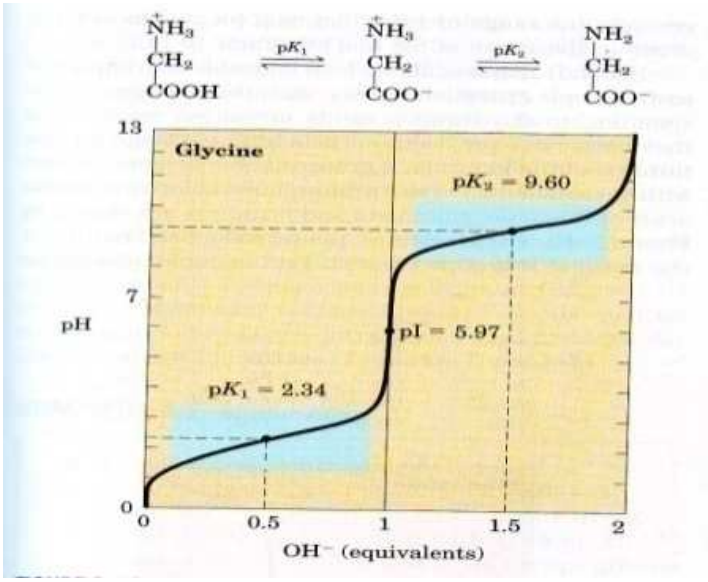


Fig. 3: Titration Curve of a Diprotic Amino Acid

Source: Lehninger

From the above curve, one can identify the following salient features:

i. The two distinct stages on the curve correspond with deprotonation of the proton-donor groups.
ii. At the start of the titration (at low pH) the amino acid is predominantly a cation (a positive ion) and at the end, it is an anion (a negative ion).
iii. The midpoint of any titration is a point of inflection where the pH is equal pK_a of the protonated group being titrated. For a diprotic amino acid, carboxyl group is the first to release its proton with a pK_a value of about 2.34. Around 9.60 you find another point of inflection corresponding to the pK_a of the amino group.
iv. Another important point of inflection exists at pH 5.97 of the titration curve. This pH is referred to as the pI (isoelectric point) of the amino acid. At its pI, an amino acid is dipolar (zwitterion) and is electrically neutral (net charge = 0).

Suffice it to say that due to variation in the structures of their side chains, some amino acids may carry additional protonated groups in their R-groups with the resultant variation in the nature of their titration curves. In addition, the pI values of amino acids have been employed in the analytical separation of mixture of amino acids.

4.0 CONCLUSION

In this unit, you have learnt why chirality, optical activity and ionic states are very important properties of amino acids.

5.0 SUMMARY

In this unit, you have learnt that:

- amino acids possess physical properties
- amino acids, with the exception of Glycine, are able to rotate the plane of plane polarized light-they are said to be optically active
- aromatic amino acids (Phe, Tyr, and Trp) strongly absorb ultraviolet light
- in solution, amino acids are ionic and this confer on them acid-base property.

6.0 TUTOR-MARKED ASSIGNMENT

1. State any three (3) properties of amino acids and explain any two.
2. Explain the following with respect to amino acids:

a. Polypeptides absorb light in the UV region.
b. An amino acid can act as proton donor as well as proton acceptor in solution.

7.0 REFERENCES/FURTHER READING

Thomas, M. Devlin (Ed.) (2006). *Textbook of Biochemistry with Clinical Correlations*. Canada: Wiley-Liss.

David, L. Nelson & Michael, M. Cox (2005). *Lehninger Principles of Biochemistry*. W. H. New York: Freeman and Company.

David, E. Metzler. (no date). Biochemistry, the Chemical Reactions of Living Cells, Volume I.

UNIT 4 CHEMICAL REACTIONS OF AMINO ACIDS

CONTENTS

1.0 INTRODUCTION

All alpha carboxyl and alpha amino groups of amino acids share common chemical reactivity. Amino acid side chains, on the other hand, exhibit specific chemical reactions depending on the type of functional groups they carry. In this unit, we shall focus on the chemical reactivity of amino acids.

2.0 OBJECTIVES

By the end of this unit, you should be able to:

- discuss some reactions that are common to amino acids due to their α-Carboxyl group
- explain some reactions that are common to amino acids due to their α-Amino group
- describe reactions that are specific for different functional groups in amino acids' side chains.

3.0 MAIN CONTENT

3.1 Chemical Reactions of Amino Acids

You may ask: what is the relevance of chemical reactions in the study of biological phenomena? Let us consider few of such reasons. First, taking

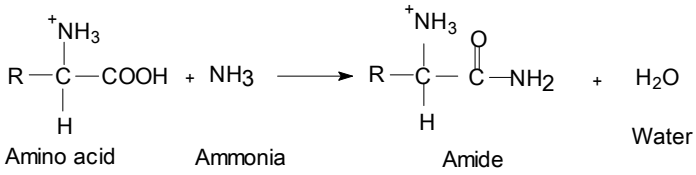
advantage of chemical reactivity of certain amino acids has permitted chemical modification of proteins. Secondly, qualitative detection and quantitative measurement of amino acids and proteins were possible because scientists have taken advantage of coloured solution formed when these acids are made to react with certain chemical reagents.

Another very important reason is that the biological function of proteins is known to be closely linked with the chemical behavior of their **R** groups.

3.2 Reactions of α-Carboxyl Group

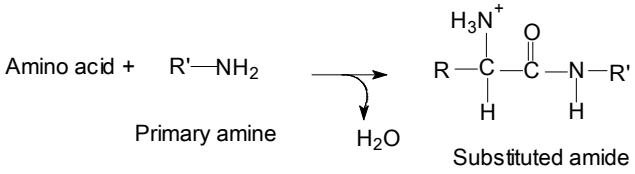
The α-carboxyl group can react with ammonia and primary amines to yield unsubstituted and substituted amides respectively.

i.

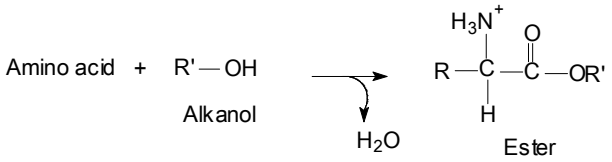


[Note: this reaction is very important for the formation of peptide bonds]

ii.

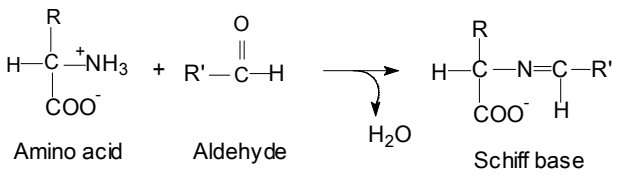


iii. Esterification in the presence of strong acid

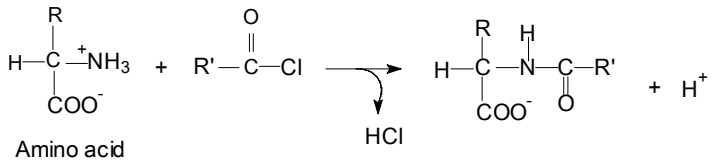


3.3 Reactions of α-amino Group

i. Free α- amino groups of amino acids may react with aldehydes to form Schiff bases.



ii. Acylation: amino groups can be acylated with acid anhydrides and acid halides.

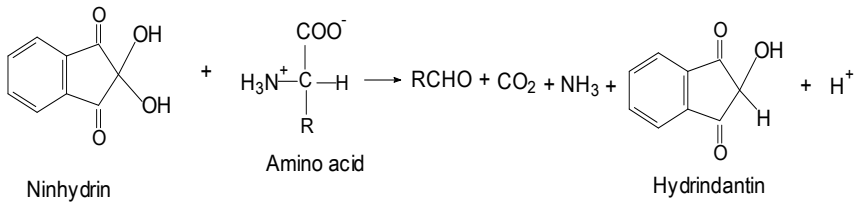


iii. Ninhydrin Reaction

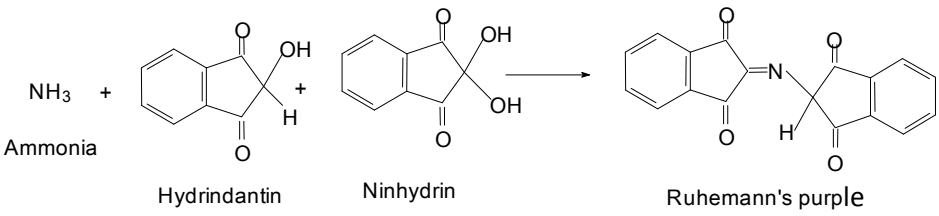
Amino acids, through the amino groups can react with ninhydrin (or triketohydrindene hydrate), a very strong oxidizing agent. This agent oxidatively deaminate the α- amino functional group of the amino acids.

The products of this reaction are aldehyde, ammonia, carbon dioxide, and hydrindantin.

Step1



Step 2



The importance of this reaction is in its application in the spectrophotometric quantification of amino acids at 570 nm. [α-imino acids**,**e.g Proline and Hydroxyproline, yield yellow complex instead of purple, with absorption maxima at 440nm].

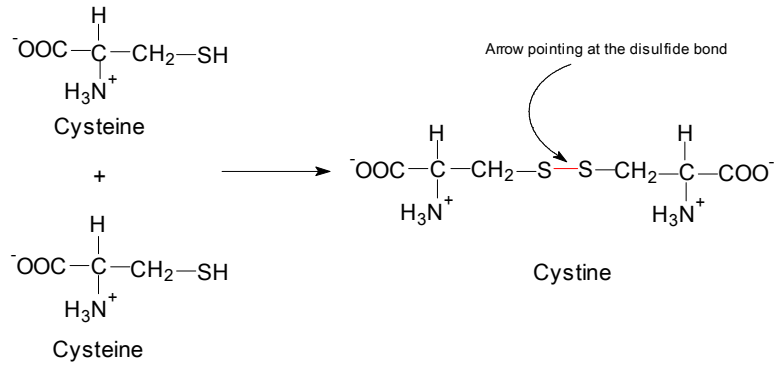
3.4 Reactions of Specific Functional Groups of the Side Chain

Specific reaction of amino acids is important for degradation, sequencing and chemical synthesis of peptides and proteins. These reactions are useful in identifying specific functional groups and hence specific amino acids possessing such functional groups.

3.4.1 Reactions of Sulfyhydryl Group

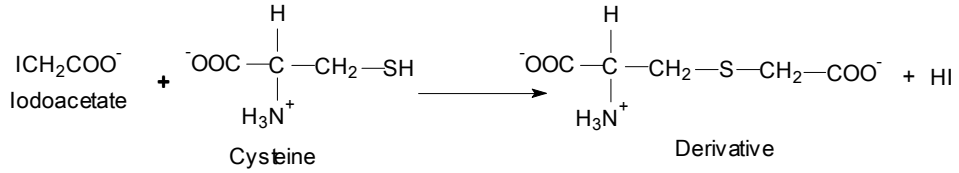
i. Disulfide bond (bridge) formation

The thiol groups of cysteine can react with one another to form disulfide species.



ii. Reactions of cysteine with iodoacetic acid.

Cysteine is also known to react with iodoacetic acid to yield S–carboxymethyl cysteine derivatives.



3.4.2 Xanthoproteic Reactions of Aromatic Ring

Amino acids with aromatic nuclei form yellow nitro derivatives on heating with concentrated nitric acid. This reaction is based on the nitration of benzene ring with concentrated HNO_3 yielding yellow derivatives of nitrobenzenes. The reaction is given by tyrosine, tryptophan, and polypeptide which contain these amino acids.

Phenylalanine is more difficult to nitrate and so requires H_2SO_4 as a catalyst.

3.4.3 Millon's Reaction

The amino acid called tyrosine, a hydroxybenzene-radical containing compound, reacts with Millon's reagent (a solution of mercuric and mercurous ions in nitric and nitrous acids) to form red complexes.

3.4.4 Reaction of Tryptophan with Formaldehyde

Tryptophan reacts with formaldehyde reagent in the presence of concentrated H_2SO_4, to form a bluish- purple compound appearing as ring at the interface of the liquids. This reaction is characteristic of the indole- ring of tryptophan.

3.4.5 Sakaguchi Reaction

Arginine, an amino acid containing guanidinium group, reacts with α-naphthol and an oxidizing agent, such as bromine water, to give a red colour compound.

4.0 CONCLUSION

This unit has examined how chemical reactions of amino acids are basically reactions of specific functional groups present in amino acids.

5.0 SUMMARY

In this unit, you have learnt that:

- understanding the reactions of amino acids will help in the detection, quantitative measurement and understanding of biological functions of amino acids and polypeptides
- α-carboxyl group of amino acids can react with ammonia and primary amines to yield unsubstituted and substituted amides respectively
- among many reactions of α–amino group of amino acids, the Ninhydrin test is very useful in spectrophotometric measurement of amino acids
- various functional groups on the side chains of amino acids also give characteristic reaction with different reagents.

6.0 TUTOR-MARKED ASSIGNMENT

1. Provide the reasons for each of the following:

a. It is essential to know how amino acids react.
b. Proline gives yellow colour with Ninhydrin instead of purple.

2. Explain the principle behind Ninhydrin test.
3. Copy and complete the table below

Chemical Reaction of amino acid	The functional group being screened for
---	indole- ring
Millon's reaction	---
---	α-amino group
---	Aromatic nucleus
Sakaguchi reaction	---

7.0 REFERENCES/FURTHER READING

Thomas, M. Devlin. (Ed) (2006). *Textbook of Biochemistry with Clinical Correlations*. Canada: Wiley-Liss.

David, L. Nelson & Michael M. Cox (2005). *Lehninger Principles of Biochemistry*. New York: W. H. Freeman and Company.

UNIT 5 PH, PK_A AND BUFFERING CAPACITY OF AMINO ACIDS

CONTENTS

1.0 INTRODUCTION

Biological activities occurring in body fluids take place within certain range of pH. This unit focuses on the concept of pH and how it is maintained by buffers.

2.0 OBJECTIVES

By the end of this unit, you should be able to:

- define pH and use its mathematical expression to perform calculation
- use dissociation of weak acid to explain pKa of amino acids
- define buffer solution and describe how buffers perform their function
- solve problems using Handerson-Hasselbalch equation.

3.0 MAIN CONTENT

3.1 Ionisation of Water

Water ionizes according to the equilibrium equation below:

$$H_2O \rightleftharpoons H^+ + OH^-$$

water hydrogen ion hydroxide ion

The equilibrium constant for the ionisation of water is given by K_{eq}

$$K_{eq} = \frac{[H^+][OH^-]}{[H_2O]}$$

……*equation-1*

At 25 ^oC, the concentration of one litre of pure water is 55.5 M (Grams of H2O in one litre divided by its gram molecular weight: 1000 gL^{-1}/18.015 $gmole^{-1}$**.**Hint: *molarity* = $\frac{gram\ concentration}{molecular\ weight}$).

So,

$$K_{eq} = \frac{[H^+][OH^-]}{55.5\ M}$$

Or

$$K_{eq}(55.5\ M) = [H^+][OH^-]$$

……*equation-2*

The product K_{eq}(55.5 M) can be taken as a new constant known as ionic product of water (k_w).

i.e. $K_{eq}(55.5\ M) = k_w = [H^+][OH^-]$.

By electrical conductivity measurements, K_{eq} of pure water at 25 ^oC is 1.8×10^{-16} M.
Now, substituting this value into equation-2 above gives

$$k_w = (55.5\ M) \times (1.8 \times 10^{-16}\ M)$$
$$= 1.0 \times 10^{-14}\ M^2$$

This means that in aqueous solution at 25 ^{o}C, the product $[H^+][OH^-]$ is always $1.0 \times 10^{-14}\ M^2$.
In pure water (neutral pH), the concentration of hydrogen ion equals the concentration of hydroxide ion ($[H^+] = [OH^-]$).
$[H^+]$ and $[OH^-]$ can be calculated thus:

$$k_w = [H^+][OH^-]$$

…….*equation-3*

$$= [H^+]^2 \quad (\text{if } [H^+] = [OH^-])$$

And so you have,

$$\sqrt{k_w} = [H^+]$$

Hence,

$$[H^+] = \sqrt{1x10^{-14}}$$
$$= 1x10^{-7}$$

From the equation-3, any of the concentration terms can be calculated if the other is known.

Example

A solution was found to have $1.5x10^{-3}$ M hydrogen ion at 25 ^{o}C. Calculate the concentration of the hydroxide ion in the solution.

Solution:

$$k_w = [H^+][OH^-]$$

$$1x10^{-14}\ M^2 = 1.5x10^{-3}\ M \times [OH^-]$$
$$[OH^-] = 1x10^{-14} / 1.5x10^{-3}$$
$$= 6.7 \times 10^{-12}\ M$$

3.2 Definition of pH and Calculation of Hydrogen ion Concentration

3.2.1 Definition of pH

The pH is a simple and convenient designation for the concentration of hydrogen ion in aqueous solution (pOH signify the hydroxide ion concentration). pH is defined as the negative logarithm to base 10 of molar hydrogen ion concentration.

This can be expressed mathematically thus:

$$pH = -\text{Log}_{10}[H^+]$$

or

$$pH = \text{Log}_{10}\frac{1}{[H^+]}$$

For a neutral solution at 25 ^oC

$$pH = - \text{Log } 1.0 \text{ x} 10^{-7}$$
$$= 7$$

Example involving calculation on pH

Calculate the pH of blood sample whose hydrogen ion concentration was found to be 0.001 Molar(i.e 1.0 x 10^{-3} M).

Solution:

$$pH = -\text{Log}_{10}[H^+]$$

$$pH = -\text{Log } 1.0 \text{ x } 10^{-3}$$
$$= 3\text{Log } 10$$
$$= 3$$

You should note that the expression pOH = -Log 1/[OH^-] is equally correct.

Again, by taking the negative logarithm of both sides of $1x10^{-14}$ = [H^+][OH^-], we get another useful equation; 14 = pH + pOH which is used to form a pH-scale.

3.2.2 pH-Scale

This is a scale of 0 to 14 values. Figure 1 below gives the pH scale showing the locations of the pH and some biological fluids.

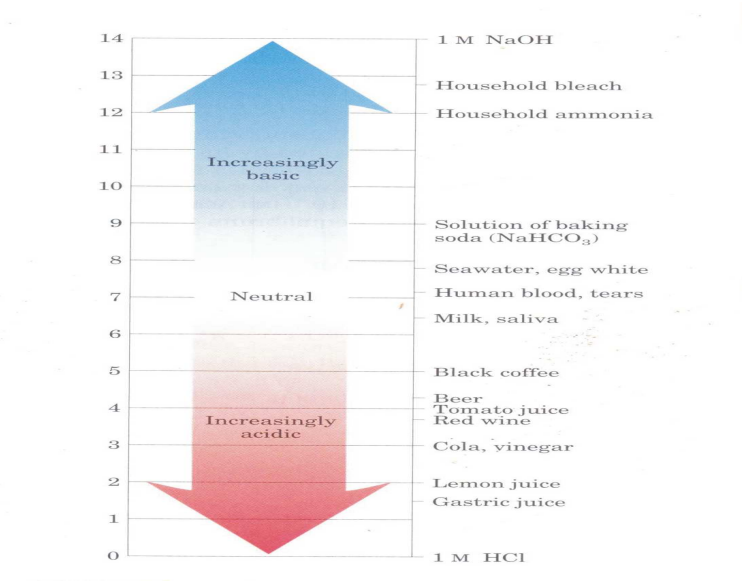


Fig. 1: The pH of some Fluids
Source: Lehninger

From the scale, solutions with values < 7 are acidic while those > 7 are basic. The pH value of 7 is taken as the neutral pH.

The pH of biological fluids are measured approximately using indicators or more accurately with the aid of pH meters.

3.3 Dissociation of Weak Acids and Definition of pK_a

Just as the dissociation of water gave us pH, the dissociation of weak acid will result into pK_a as you will see soon.

3.3.1 Dissociation of Weak Acids

A weak acid is one which dissociates partially in water. Acetic acid is a good example of a weak acid; in water it releases only one of its protons.

$$\underset{\text{Acetic acid}}{CH_3COOH} \overset{K_{eq}}{\rightleftharpoons} \underset{\text{Acetate ion}}{CH_3COO^-} + \underset{\text{proton}}{H^+}$$

The equilibrium constant in the above equation is a measure of the tendency of the acid (CH_3COOH) to lose a proton and form its conjugate base (CH_3COO^-).

Generally, for the dissociation of any weak acid (HA) the equilibrium constant is written as K_a

$$HA \underset{}{\overset{K_a}{\rightleftharpoons}} A^- + H^+$$

weak acid — conjugate base — proton

The K_a for acetic acid is 1.74×10^{-5}.

Weaker acids are known to have smaller dissociation constants as opposed to larger values for stronger acids.

3.3.2 pK_a of Weak Acids

Just as for pH, pK_a is defined as the negative logarithm to base 10 of K_a and is expressed mathematically as:

$$pK_a = - \log K_a$$

or

$$= \text{Log } 1/K_a$$

Usually, the stronger the tendency to lose a proton, the stronger the acid, the larger its K_a and the lower its pK_a. Hence, weaker acids have smaller dissociation constants (K_a) and higher pK_a values.

Amino acids are weak acids also. On a titration curve, the pK_a values of Ionisation groups exist at the point of inflection of their dissociation (see Figure 2 below).

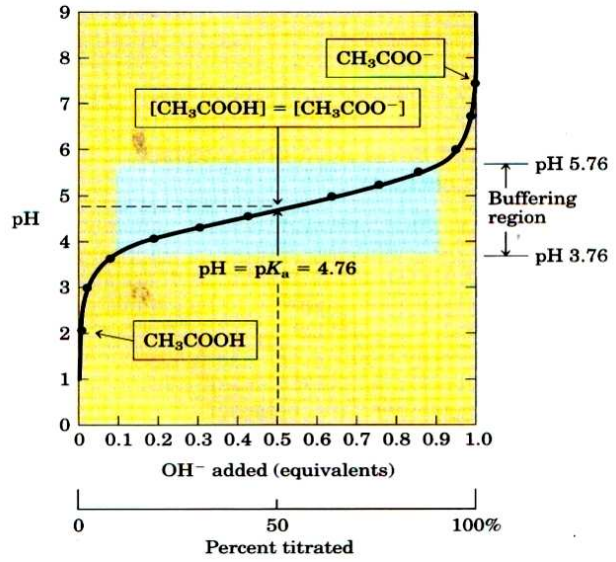


Fig. 2: Titration Curve of Acetic Acid
Source: Lehninger

3.4 Buffer Solutions

3.4.1 Definition and Examples of Buffer Solution

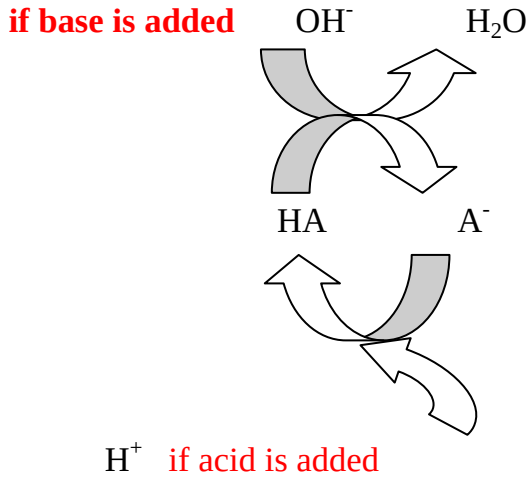
A buffer solution is a solution that resists slight changes in pH when little amount of acid or base is added to it. A buffer system consists of a weak acid and its conjugate base. An example of a buffer is a mixture of equal concentration of acetic acid and acetate ion (its conjugate base), found at the midpoint of the titration curve.

3.4.2 Roles of Biological Buffers

The buffer system in living organisms appears to be their first line of defense against any slight change in pH. For example, excessive production of acid during muscular exercise, release of toxic substances during infection caused by micro organisms or even in nutritionally deficient state (as in diabetes) could cause slight change in pH. This must be resisted as fast as possible by the actions of several buffers found in the body.

3.4.3 Buffer Action

Buffers perform their action by the interaction of the weak acid and conjugate base with OH^- and H^+ respectively. The chart below is summary illustration of buffer action:



From the chart, the following points can be noted:

i. the buffer system is reversible;
ii. if base is added, the weak acid (HA) interacts with it to give water and conjugate base;
iii. if acid is added, the conjugate base interacts with it to produce the weak acid and so resist the change in the pH.

3.4.4 Handerson-Hasselbalch Equation

In the laboratory, buffers are sometimes prepared using the Henderson-Hasselbalch equation, which is a description of the titration of all weak acids. The equation is expressed mathematically as:

$$pH = pK_a + \frac{[\text{conjugate base}]}{[\text{weak acid}]}$$

Calculation involving Handerson-Hasselbalch equation:

Calculate the pK_a of lactic acid given that at pH 4.8, the concentration of lactic acid and its conjugate base is 0.001 and 0.087 M respectively.

Solution

$$pH = pK_a + \text{Log}\frac{[\text{Lactate}]}{[\text{Lactic acid}]}$$

$$4.8 = pK_a + \text{Log}\frac{0.087}{0.001}$$

$$pK_a = 4.8 - \text{Log}\frac{0.087}{0.001}$$

$$= 2.86$$

3.4.5 Amino Acids as Biological Buffers

Among many other biological buffers, ionisable groups of some amino acid residues in proteins could act as buffers. For example, histidine has a pK_a of 6.0. Proteins containing histidine residues therefore buffer effectively near neutral pH.

4.0 CONCLUSION

Amino acids have acid-base properties that allow them to, among many functions, buffer biological fluids.

5.0 SUMMARY

In this unit, you have learnt that:

- pH is the negative logarithm of hydrogen ion concentration
- K_a and pK_a are measure of ionisation of weak acids
- buffer solution helps to maintain pH of biological systems

- handerson-Hasselbalch equation can be used when preparing buffers in the laboratory
- side chains of amino acids could act as buffers if they have ionisable group with pK_a values near physiological pH.

6.0 TUTOR-MARKED ASSIGNMENT

1a. Define pH?
b. Calculate the pH of a laboratory solution if its hydrogen ion concentration is 10^{-2} M.

2. The equilibrium constant K_a of the dissociation of an unknown substance, X, is $1.25x10^{-3}$. Calculate the pK_a of X and use your result to predict the range of pH where X can act as buffer.

3a. What is buffer solution?
b. Describe how any named buffer solution perform their action.
c. Given that the molar ratio of dilute solution of potassium acetate to acetic acid (pK_a=4.76) is 3:1, calculate the pH of the solution.

7.0 REFERENCES/FURTHER READING

Thomas, M. Devlin (Ed) (2006). *Textbook of Biochemistry with Clinical Correlations*. Canada: Wiley-Liss.

David, L. Nelson & Michael M. Cox (2005). *Lehninger Principles of Biochemistry*. New York: W. H. Freeman and Company.

Donald, Voet & Judith G. Voet (2004). *Biochemistry*(3rd Ed.). USA: John Wiley & Sons, Inc.

David, E. Metzler (no date). *Biochemistry- the Chemical Reactions of Living Cells* (2nd Ed.). Elsevier: Academic Press.

MODULE 2 PEPTIDES

UNIT 1 PEPTIDES-FORMATION AND NOMENCLATURE

CONTENTS

1.0 INTRODUCTION

Peptides constitute one of the two major classes of polymers of amino acids. Proteins constitute the other class. In this unit, you shall learn about peptides, how they are formed and their basic properties.

2.0 OBJECTIVES

By the end of this unit, you should be able to:

- define peptide
- explain how peptides are formed
- describe the nature of peptide bond
- explain how peptides are named.

3.0 MAIN CONTENT

3.1 Definition of Peptide

Peptides are biopolymers of amino acids in which amino acids are joined by peptide bonds.

3.2 Differences between Peptide and Protein

Proteins are also polymers of amino acids. The basic distinguishing feature between peptide and protein is in respect to their molecular weight. Amino acid polymers (polypeptides) with molecular weight greater than 10000 are termed proteins while those with molecular weight less than 10000 are called peptide. In many proteins, two or more polypeptides are linked via covalent (disulfide bridge) or noncovalent hydrophobic interactions.

3.3 Peptide Bond Formation and the Nomenclature of Peptide

3.3.1 Peptide Bond Formation

Peptides are formed through covalent bonding between two or more amino acids' molecules. This covalent bond, the peptide bond, is an amide linkage. The peptide bond is formed by the elimination of a water molecule from the α-carboxyl group of one amino acid and the α-amino group of another as shown in Fig. 1 below:

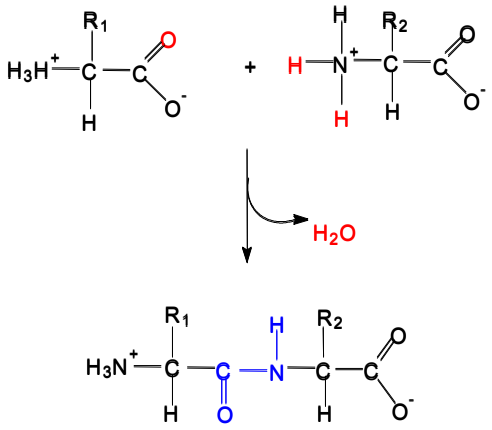


Fig. 1: Formation of a dipeptide, showing the peptide bond in blue

Source: Voet & Voet

This peptide bond formation is initiated via an attack of the lone-pair electrons of the amine on the carbonyl carbon of the carboxyl group. If such a condensation polymer consists of two amino acids joined by a peptide bond, it is referred to as dipeptide. Three amino acids can be joined by two peptide bonds to form a tripeptide. Similarly, four amino acids will form a tetrapeptide while five amino acids will produce a pentapeptide and so forth. Polymers composed of few (3-10) amino acids residues are known as oligopeptide. The term polypeptide refers to a polymer containing many amino acid residues.

Amino acid unit that exists as part of a peptide is often called residue or amino acid residue, **in recognition of** the loss of molecule of water during the polymerization. The precise number of water molecule lost is one less than the number of residues.

3.3.2 Nomenclature of Peptides

In a peptide, amino acid residue is linked to its neighbour in a head-to-tail manner, forming linear polymer. The end of the linear polymer with a free α-amino group is the N-terminus while the amino acid residue at this end is called the amino terminal (or N-terminal) residue. In the same way, the opposite end of the polymer with a free carboxyl group and its terminal amino acid are called the C-terminus and carboxyl-terminal residue respectively.

A peptide is named according to the amino acid residues in it and usually begins with the N-terminal residue. Thus, the pentapeptide in Fig. 2 below is named serylglycyltyrosylalanylleucine and abbreviated as ser-Gly-Tyr-Ala-Leu or SGYAL (using one-letter abbreviation).

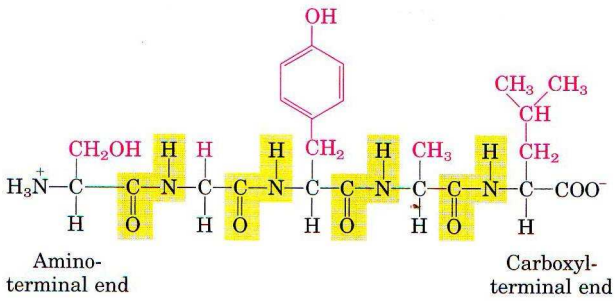


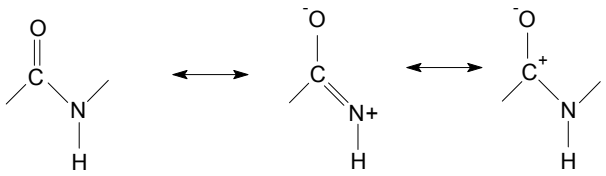
Fig. 2: A Pentapeptide showing the N- and the C- termini
Source: Lehninger

In describing the amino acid sequence of a peptide, it is customary to place the amino-terminal residue at the left and the carboxyl-terminal residue at the right. Again, residues are numbered from the N-terminus with the N-terminal residue as 1.

3.4 Properties of Peptide Bond/Group

The three-dimensional structure of protein (as you would see in module 3) is dependent upon the properties of the amide linkages between amino acid units. The peptide bond/peptide group possesses the following key properties:

i) Peptide group is a resonance hybrid of the following structures:



Apart from the second and third properties mentioned below, another consequence of resonance of peptide bond is that the oxygen atom acquires partial negative charge and the NH group partial positive charge. These opposite ends of the dipole tend to associate to form hydrogen bonds which are very important in stabilising protein structure.

ii) It is rigid and plainer. This arises from resonance interactions.

iii) It is a partial double bond. This is also due to resonance.

iv) Except in few exceptions, it assumes the trans conformation in which successive C_α atoms are on opposite sides of the peptide bond joining them.

v). It is 1.32 Å in length. This value is in between that of single bond (1.49 Å) and that of double bond (1.27 Å).

4.0 CONCLUSION

The peptide bond plays critical role to the understanding of the three-dimensional properties of proteins.

5.0 SUMMARY

In this unit, you have learnt that:

- peptides are polymers of amino acids linked by means of peptide bonds
- a peptide is named according to the amino acid residues in it
- the existence of resonance among peptide groups is basically responsible for their properties.

6.0 TUTOR-MARKED ASSIGNMENT

1. Define the term peptide? Differentiate between peptides and proteins.
2. Write the structure of the tripeptide Alanylserylvaline.
3. Explain resonance in peptide group.

7.0 REFERENCES/FURTHER READING

Thomas, M. Devlin (Ed). (2006). *Textbook of Biochemistry with Clinical Correlations*. Canada: Wiley-Liss.

David, L. Nelson & Michael, M. Cox (2005). *Lehninger Principles of Biochemistry*. New York: W. H. Freeman and Company.

David, E. Metzler (no date). *Biochemistry- the Chemical Reactions of Living Cells* (2nd Ed.). Elsevier: Academic Press.

UNIT 2 PROPERTIES, EXAMPLES AND FUNCTIONS OF BIOLOGICAL PEPTIDES

CONTENTS

1.0 INTRODUCTION

You learnt in unit one of this module that peptides are amino acid polymers formed through peptide bond. You also learnt the nature of peptide bond. In this unit, you shall learn more about biological peptides especially, their properties, examples and functions.

2.0 OBJECTIVES

By the end of this unit, you should be able to:

- state the properties of peptides
- list some common biologically active peptides
- associate biological peptides with specific functions.

3.0 MAIN CONTENT

3.1 Biologically Active Peptides

Peptides occur naturally in some biological systems even though synthetic peptides can be produced. Biologically active peptides are those with identifiable biological roles.

3.2 Properties of Peptides

Since peptides are polymers of amino acids, one should expect them to exhibit ionic and acid base properties.

3.2.1 Ionic Property

The Ionisation property of peptides is as a result of contribution from the N-terminal free α-amino group, the C-terminal α-carboxyl group and the ionisable R groups of some residues. Since the α-amino groups and the α-carboxyl groups of all non-terminal amino acids take part in peptide bonds formation, they do not contribute to the overall ionisation property.

3.2.2 Titration Curves

The acid-base property of peptides results from their ionisation property discussed above. These positive and negative groups could participate in acid-base titration. Peptides, thus, like free amino acids, have characteristic titration curves and characteristic isoelectric pH (pI) at which they assume electrical neutrality and so remain static in an electric field. You should note that when an amino acid having an ionisable R group becomes a residue in a peptide, the pK_a value of the R group can change due to a number of factors. First, is as a result of loss of charge in its α-amino and α-carboxyl groups. Second, is due to interaction with other peptide R groups, and finally, other environmental factors such as salt concentration. Because of these aforementioned reasons, it is difficult to pinpoint the pK_a values from the titration curves of both peptide and proteins as is the case for free amino acids. Figure 1 gives the titration curve of the enzyme ribonuclease A.

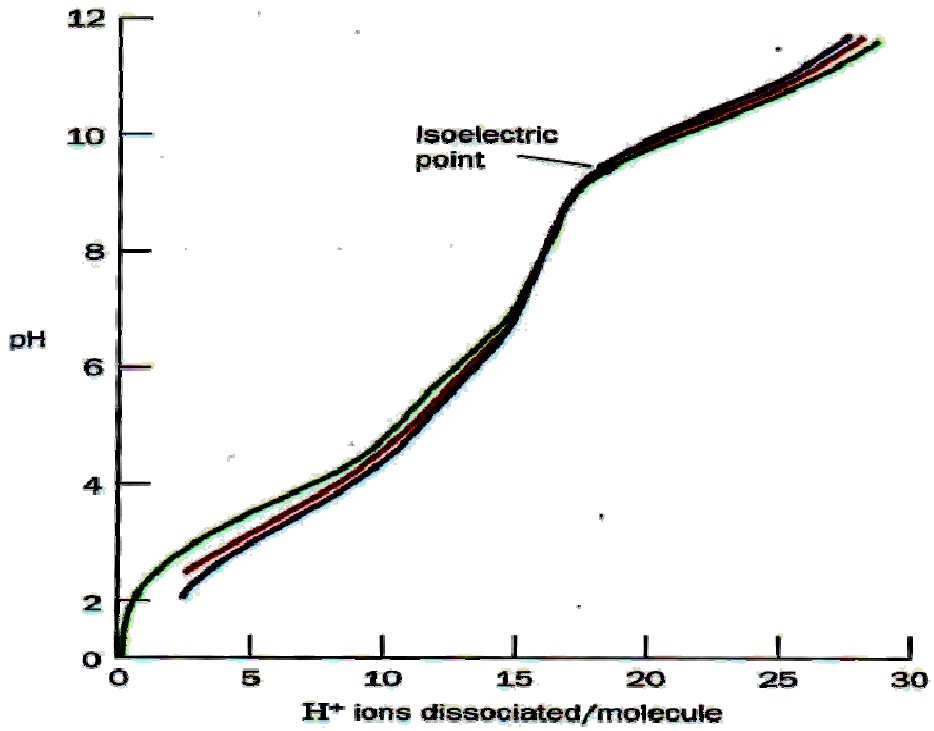


Fig. 2: Titration Curves of Ribonuclease A (the concentration of KCl is 0.01M for the blue curve, 0.03M for the red curve, and 0.15M for the green curve).

Source: Voet & Voet

3.3 Examples and Functions of Biologically Active Peptides

Peptides occur in abundance in nature. Glutathione, a tripeptide is the most abundant peptide in mammalian tissue. A number of hormones are peptides molecules. The table below gives the names of some mammalian peptides:

Hormone	Number of Residues
Oxytocin	8
Vasopressin	8
Glucagon	29
Adrenocorticotropic hormone	39

Some of the hormone regulatory factors which are secreted by the hypothalamus are peptides that govern the release of hormones by other endocrine glands.

Many of the antimicrobial agents produced by microorganisms are peptides that can contain both the D- and L-amino acid residues.

Penicillin contains a cyclic peptide as part of its structure. Gramicidins, the tyrocidins, the polymyxins, the subtilisins and the bacitracins are other peptide antibiotics.

4.0 CONCLUSION

The ionic/acid-base property may be the reason several biologically active peptides occur in nature and perform various functions.

5.0 SUMMARY

In this unit, you have learnt that:

- peptides, being polymers of amino acids, are charged species also
- peptides have characteristic titration curves that are not as distinct as those of free amino acids
- several biologically active peptides occur in nature.

6.0 TUTOR-MARKED ASSIGNMENT

1a. Discuss the contribution of various groups to the overall ionic properties of peptides.

b. Compare and contrast the acid-base properties of peptides with that of free amino acids.

2. What do you understand by the term, biologically active peptide? Name five peptides that are found in living systems

7.0 REFERENCES/FURTHER READING

Thomas, M. Devlin (Ed) (2006). *Textbook of Biochemistry with Clinical Correlations*. Canada: Wiley-Liss

David, L. Nelson & Michael, M. Cox (2005). *Lehninger Principles of Biochemistry*. New York: W. H. Freeman and Company

David, E. Metzler. *Biochemistry- the Chemical Reactions of Living Cells* (2nd Ed.). Elsevier: Academic Press.

UNIT 3 SEPARATION OF PEPTIDES I

CONTENTS

1.0 INTRODUCTION

The study of the properties and activities of peptides and proteins require that these polymers be relatively free from contaminants. In this unit and in the next, you would be learning the various separation techniques for peptides and proteins. Since both peptides and proteins are identical, the techniques for their purification are to a large extent the same. The techniques take advantage of some basic properties of these polymers, such as solubility, molecular size, ionic charge and affinity to other biological molecules. This unit focuses on separation on the basis of the first two properties while unit 4 will be devoted to the last two.

2.0 OBJECTIVES

By the end of this unit, you should be able to:

- state the principles of peptide/protein purification
- explain how peptides are separated on the basis of solubility
- explain peptide purification techniques on the basis of molecular size.

3.0 MAIN CONTENT

3.1 Preliminary Steps to Peptide/Protein Purification

Peptides and proteins occur naturally as parts of tissues or cells. The first step in protein purification procedure is usually cell disruption. The aim of this step is to release the protein content of these cells/tissues. Cell disruption is carried out using blender or pestle and mortar, and sometimes by repeated cycles of icing and thawing.

After cell disruption, the crude extract is then centrifuged to separate the cellular pellets or debris from the clear solution which contains the mixture of proteins. This clear supernatant is subsequently subjected to other purification protocols until the protein of interest is isolated.

3.2 Purification Technique on the Basis of Solubility

Early fractionation steps in protein/peptide purification utilise differences in protein/peptide solubility. At high ionic strengths, the solubilities of polypeptides decrease with the increasing salts concentration. This phenomenon is called 'salting out' and it arises because of the competition between the added salt ions and the dissolved polypeptides for molecules of solvation. The numerous solute ions are so solvated that the polypeptides are not sufficiently solvated but forced to precipitate out of the solution. Precipitation methods are popular first steps in protein purification because they can be carried out in large 'batch" scale.

The addition of certain salts in the right amount can selectively precipitate some polypeptides, while others remain in solution.

Ammonium Sulphate is the most commonly used reagent for salting out proteins. Once precipitated, the proteins are then removed or concentrated through filtration or centrifugation. Organic solvents such as acetone and ethanol are also good precipitating agents. Conversely, salting in refers to increase in protein solubility at low ionic strength with the increase in salt concentration.

3.3 Purification Technique on the Basis of Molecular Size

3.3.1 Dialysis

Dialysis is a separation procedure in which macromolecules like polypeptides are separated from smaller molecules of solvent, salts, minerals and other metabolites. This is done with the aid of semi permeable membrane such as cellophane (made of cellulose acetate), taking advantage of variation in molecular size. The solution containing the protein mixture is sealed in a dialysis bag (see figure 1below). When the sealed bag is immersed in a much larger volume of buffered solution and allowed some time to equilibrate, the membrane barrier

allows the flow of the small molecules out of the bag but not the proteins. The flow of these molecules is via osmosis and it can be enhanced by stirring. By continuously putting the bag in fresh buffer solution, it is possible to get rid of the smaller salts ions.

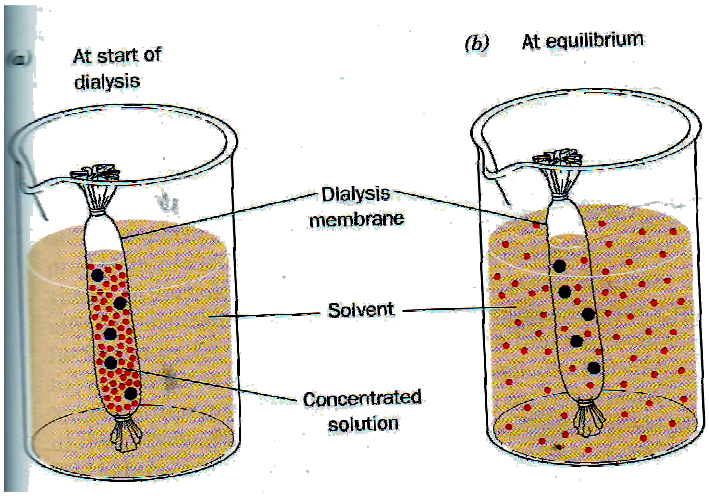


Fig.1: Diagram of Dialysis
Source: Voet & Voet

3.3.2 Ultracentrifugation

This is another separation procedure in which proteins are purified on the basis of their molecular masses or sizes. When a solution of macromolecules is subjected to an ultracentrifugal force (a force greater than 4 x 10^5 times that of gravity), the proteins accelerate rapidly to a constant velocity of sedimentation. This is expressed as the sedimentation constant, S. S is the rate per unit of centrifugal force and is given by the equation below:

$$S = \frac{v}{W^2 r}$$

Where **v** is the velocity of the protein, **W** is the angular velocity of the centrifuge and **r** is the radius of rotation, which is the distance from the center of the tube in which the protein mixture is placed to the centre of rotation.

Each protein sediment's with a characteristic sedimentation coefficient.

The unit of **s** is Svedberg (S) named after the Swedish biochemist who first developed ultracentrifuge in 1923. (1S= 10^{-13}).

It is worth noting that ultracentrifugation has been used in the determination of proteins' molecular weights by simply taking advantage of the following equation:

$$s = v/W^2r = \frac{M(1-\bar{v}\rho)}{Nf}$$

M= molecular mass in daltons
$\bar{v}$= partial specific volume in ml/gram
ρ (rho)= density (g/ml) of the solvent.
f = frictional coefficient

In order to achieve enhanced resolution of the macromolecules, improved process of sedimentation which involves conducting the process in density gradient solution of inert substances such as sucrose or CsCl has been developed. This version of ultracentrifugation is called density gradient ultracentrifugation.

3.3.3 Gel Filtration

This separation technique is also called molecular exclusion, size exclusion and molecular size chromatography. It separates molecules of proteins/peptides according to their sizes and shape. It is a column chromatography and like all chromatographic methods, it is made up of stationary and mobile phases. The protein samples dissolved in liquid solvent make up the mobile phase while the stationary phase consists of small insoluble matrix or beads of hydrated, sponge-like material known as the gel. The gel contains pores or cavities of a particular size. If a solution of proteins of various sizes is allowed to pass through this gel-containing column, the gel acts as a "molecular sieve" permitting their elution on the basis of size. Small proteins penetrate the pores of the gel and have a larger solvent volume through which to journey down the column (than larger proteins) and so are eluted last. The larger proteins transverse the column more rapidly since they are not trapped in the gel cavities. Therefore, larger polypeptides are the first to be eluted from the column as depicted in figure 2 below.

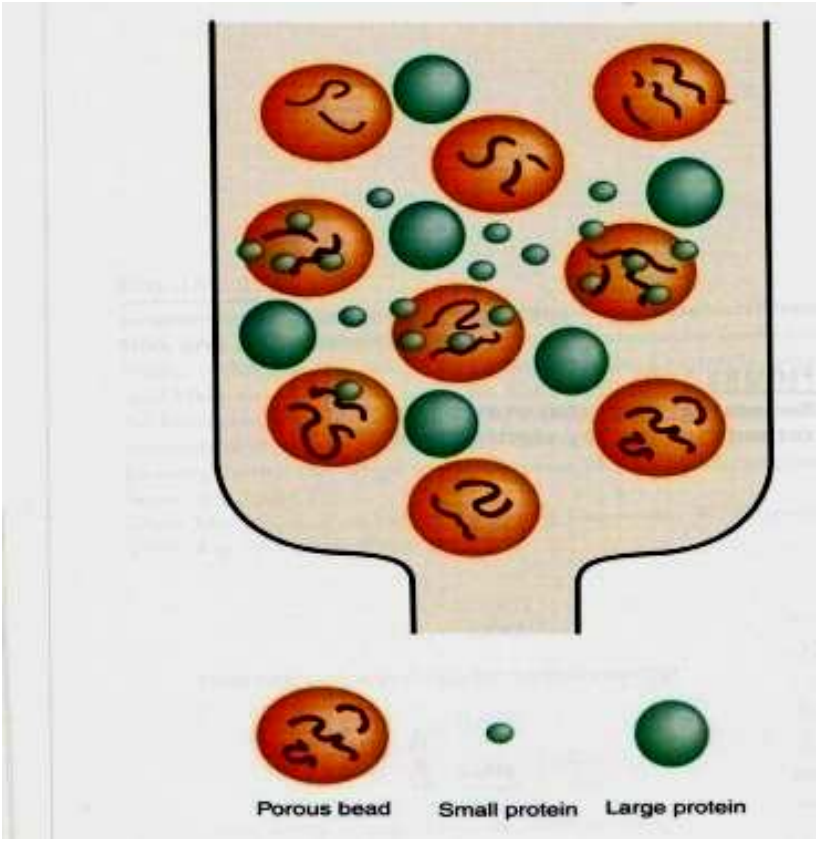


Fig. 2: Diagram of Gel Filtration
Source: Thomas Devlin

Types of gel materials

Various types of materials have been used as gel materials. The common ones include:

1. Dextran (a high molecular mass polymer of glucose produced by the bacterium *Leuconostoc mesenteroides*). Dextran is sold under the trade name Sephadex.
2. Agarose (a linear polymer of alternating D-glucose and 3.6-anhydro-L-galactose obtained from algae. It is commercially available as Sephapose and Bio-Gel A.
3. Polyacrylamide (commercially available as Bio-Gel P). Polyacrylamide gels are largely used for separation of proteins.

Determination of molecular weight of proteins using gel filtration

One important application of gel filtration is in the determination of molecular weights of proteins. This takes advantage of the linear relationship between the relative elution volume (V_e/V_o) of a protein (the ratio of volume of the solvent required to elute a protein from the column after it has first contacted the gel,$v_{e,}$ and the volume of the

column occupied by the gel beads, V_o) and the logarithm of its molecular mass. The molecular mass of an unknown peptide is obtained by carrying out its gel filtration and those of standard proteins (reference proteins of known masses). If a plot of relative elution volume (V_e/V_o) or elution volume (V_e) in ml against relative molecular mass (M_r, in Daltons) of all the proteins is carried out, it is possible to estimate the molecular mass of an unknown protein from its position on the graph.

The figure below gives such type of graph:

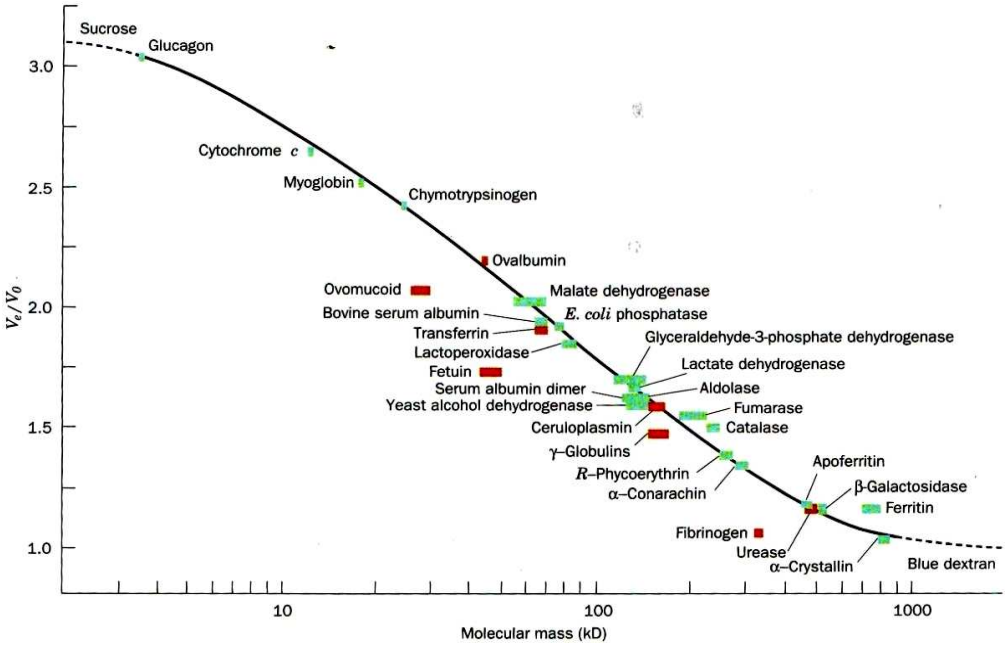


Fig. 2: Standard Curve for Calculating Molecular Mass of Unknown Protein

Source: Voet &Voet

4.0 CONCLUSION

By exploiting the differences in solubility property and size, it is possible to purify a protein/peptide from a mixture of proteins and/or peptide.

5.0 SUMMARY

In this unit, you have learnt that:

- cell disruption and centrifugation are preliminary steps in isolation and purification of proteins and peptides
- addition of salts to protein solution could lead to their precipitation and possible separation

- the use of a semi permeable membrane, the application of ultracentrifuge force and the use of molecular sieve of gels could assist in the separation of proteins in their solutions on the basis of size differences
- ultracentrifugation as well as gel filtration could be applied in the determination of protein's molecular mass.

6.0 TUTOR-MARKED ASSIGNMENT

1a. What do you understand by semi permeable membrane?
b. Explain the use of dialysis in protein purification.

2a. What is ultra centrifugal force?
b. Explain its application in purification/separation of peptides and proteins

3a. Name any two matrix materials use for column chromatography
b. Why are gels called "molecular sieve"?
c. Differentiate between mobile and stationary phases of a chromatographic technique.
d. Describe how gel filtration can be used to separate a mixture of proteins of varying sizes.

7.0 REFERENCES/FURTHER READING

Thomas, M. Devlin (ed.) (2006). *Textbook of Biochemistry with Clinical Correlations*. Canada: Wiley-Liss.

David, L. Nelson & Michael, M. Cox (2005). *Lehninger Principles of Biochemistry*. New York: W. H. Freeman and Company.

David, E. Metzler (no date). *Biochemistry- the Chemical Reactions of Living Cells* (2nd ed.). Elsevier: Academic Press.

Donald, Voet & Judith G. Voet (2004). *Biochemistry*(3rd ed.). USA: John Wiley & Sons, Inc.

UNIT 4 SEPARATION OF PEPTIDES II

CONTENTS

1.0 INTRODUCTION

This unit is a continuation of the separation techniques for peptides and proteins started in unit 3. We shall focus on methods of purification that are based on two properties-**charges** and **affinity** for biomolecules.

2.0 OBJECTIVES

By the end of this unit, you should be able to:

- define terms like electrophoretic mobility and isoelectric pH in relation to protein
- separation
- differentiate SDS-polycarylamide gel electrophoresis from isoelectric focussing
- describe the procedure of affinity chromatography
- give examples of various ion-exchangers.

3.0 MAIN CONTENT

3.1 Further Approaches to Separation of Peptides

In unit 3 of this module, you have seen that proteins can be purified and separated by simply taking advantage of differences in their solubility and in their sizes. These are not the only parameters used in purifying

polypeptides. Since proteins and peptides are charged molecules and have various functional groups that could serve as potential anchors for other molecules, these dual properties have been used to develop separation techniques for protein/peptides and even other macromolecules with such properties e.g. carbohydrates and nucleic acids.

3.2 Separation Based on Charges

Polypeptides possess various charged groups and, in solution, do have characteristic pH at which they assume electrical neutrality (net charge of zero). This pH, you know, is called the isoelectric point (pI). Both electrophoresis and isoelectric focus on separate proteins/peptides on the basis of the rates of their movement in an electric field.

3.2.1 Technique of Electrophoresis

The term electrophoresis refers to the migration of charged molecules in an electric field. Electrophoretic technique is a widely used analytical separation procedure for proteins and peptides (among many other charged biomolecules). Under the influence of electric field of few milli amperes of current, proteins migrate and separate. Those with a net negative charge migrate towards the anode while those with net positive charge move towards the cathode of the electrophoresis appliance. This migration is expressed as electrophoretic mobility (μ) and is mathematically given as: $\mu = \frac{v}{E} = \frac{q}{f}$. Where *v* is the velocity of migration, *E* is the electric field strength; q is the charge on the ion while *f* is the frictional coefficient which is a function of size, shape, state of salvation of ion and the viscosity of solvent. Each protein has its characteristic electrophoretic mobility. The above equation shows that μ (rate of migration) depends on several factors, such as:

i. the potential gradient of the voltage applied;
ii. net charge on the protein/peptide molecule- molecule with larger net charge migrate faster; and
iii. molecular size and shape (from *f* in the equation)-larger proteins migrate at a slower rate than smaller ones.

There are several types of electrophoresis. In gel electrophoresis, one of the most common electrophoretic methods, electrophoresis is carried out in polyacrylamide gel which acts as a molecular sieve. The molecular separation is based on gel filtration as well as electrophoretic mobility.

The diagram below is a type of gel electrophoresis.

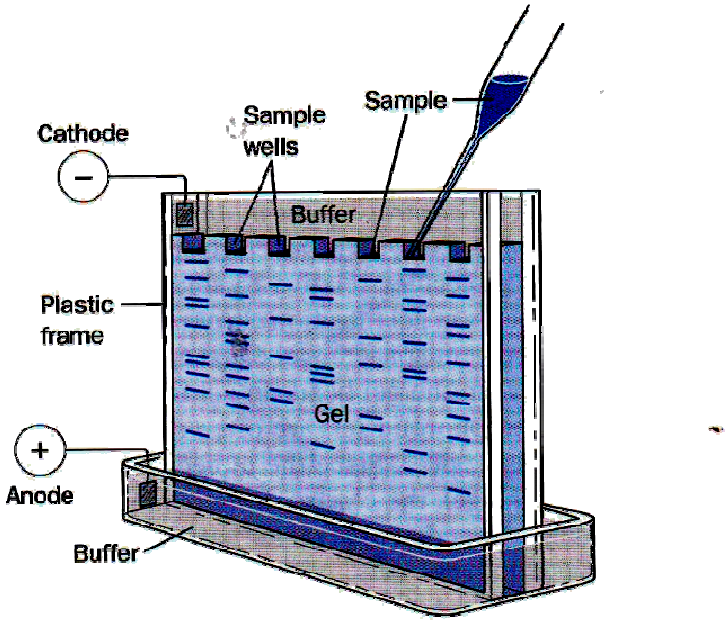


Fig. 1: A Set up for Gel Electrophoresis
Source: Voet & Voet

Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) is another gel electrophoretic method commonly employed for estimating the purity and molecular weight of proteins using the detergent sodium dodecyl sulphate (SDS).

3.2.2 Isoelectric Focusing

In isoelectric focusing, a mixture of low molecular weight polyamino-polycarboxyl acid ampholytes is used to establish pH gradient by distributing them across a gel. When a solution containing proteins is applied on the gel, each protein moves until it gets to the pH that equals its isoelectric point. Proteins with different isoelectric points, therefore, separate on the gel even if they have identical size on gel filtration.

Isoelectric focusing, therefore, has an extremely high resolution power.

To achieve even higher resolution, isoelectric focusing has been combined with SDS-PAGE to produce a technique known as two-dimensional electrophoresis.

3.3 Separation Based on Affinity to other Biological Molecule

3.3.1 Ion Exchange Chromatography

This technique depends upon electrostatic interactions between the charged groups of proteins/peptides and oppositely charged groups on ion exchanger (also called ion exchange resin). These ion exchangers consist of insoluble matrix to which charged groups have been covalently bound. There are two types of ion exchangers. The ionic groups in cation exchangers are negatively charged while those of the anion exchangers are positively charged. Table 1 below gives examples of each type.

Table 1: Examples of Ion Exchangers

S/ No	**Ion Exchanger**	**Functional Group**
	Anion Exchangers	
1	Amino Ethyl (AE)	$-OCH_2CH_2NH_3^+$
2	Diethylaminoethyl (DEAE)	$-OCH_2CH_2NH^+(CH_2CH_3)_2$
3	Quarternary Aminoethyl (QAE)	$-OCH_2CH_2N+(C_2H_5)_2CH_2CH(OH)CH_3$
	Cation Exchangers	
1	Carboxymethyl (CM)	$-OCH_2COO^-$
2	Phospho	$-PO_4H_2^-$
3	Sulphopropyl (SP)	$-CH_2CH_2CH_2SO_3^-$

Carboxymethyl-sephadex, phosphocellulose and DEAE-cellulose are ion exchangers of choice for protein /peptide separation.

The separation procedure involves passing the sample solution through a column packed with a type of ion exchanger using a buffer solution. The separation is carried out by reversible adsorption in two stages:

Stage 1: The protein mixture is allowed to interact with the resin so that the protein of interest gets absorbed unto the exchanger. Again, proteins with relatively low affinities for the ion exchanger move through the column faster than the bound protein with higher affinities.

Stage2: This involves the separation and elution of the bound polypeptide from the exchanger and is carried out by using a fresh eluting buffer of different pH so that the affinity of the bound proteins to the matrix is greatly reduced.

In ion exchange chromatography, one may choose whether to bind the protein/peptide of interest or absorb out the contaminants and allow the substance of interest to elute out of the column.

3.3.2 Affinity Chromatography

Proteins are known to have high affinity for various substances such as substrates, prosthetic groups, receptors, inhibitor and antibodies raised against them. When any of these affinity compounds (called ligands) is covalently attached (immobilised) to insoluble resin (matrix), it can be used to purify its conjugate protein by allowing a mixture containing the protein of interest to pass through a column of the immobilized ligand.

During their passage, only the protein with complimentary site to that of the immobilized ligand is retarded and hence separated from others. The desired protein can then be recovered from the column by changing the elution conditions.

The major requirements of affinity chromatography are:

(i) a biospecific ligand which can be covalently coupled to a chromatographic matrix;

(ii) that such ligands retains their biological activities.

A ligand ideal for affinity chromatography should exhibit the following characteristics:

i. specific and reversible binding;
ii. presence of chemically modifiable groups that can be used for attachment to the
matrix while it still retain its binding properties;
iii. its affinity for the binding site of the biomolecule should be within 10-4 and 10-8 M in free solution; and
iv. it should interact with the biomolecule noncovalently

Affinity chromatography can be used to purify proteins and peptide such as enzymes, antibodies, lectins, hormones and even whole cells.

4.0 CONCLUSION

Manipulating protein affinities for their purification may appear quite cumbersome but its successful application gives rise to extremely high resolution of proteins/peptides.

5.0 SUMMARY

In this unit, you have learnt that:

- polypeptides migrate differently under the influence of electric fielding. This could be used in separating these biomolecules
- isoelectric focusing exploits variation in pIs of proteins
- both ion exchange chromatography and affinity chromatography take advantage of differences in affinities of polypeptides to separate them.

6.0 TUTOR-MARKED ASSIGNMENT

1a. Define electrophoresis.
b. Enumerate the factors that could affect the rate of migration of proteins in gel electrophoresis.

2. Differentiate between electrophoresis and isoelectric focusing.

3. Write short notes on:

(a) Anions ion exchangers
(b) Affinity chromatography.

7.0 REFERENCES/FURTHER READING

Thomas, M. Devlin (ed.) (2006). *Textbook of Biochemistry with Clinical Correlations*. Canada: Wiley-Liss.

David, L. Nelson & Michael, M. Cox (2005). *Lehninger Principles of Biochemistry*. New York: W. H. Freeman and Company.

Donald, Voet & Judith G. Voet (2004). *Biochemistry* (eds.). USA: John Wiley & Sons, Inc. USA.

UNIT 5 PEPTIDE SEQUENCING

CONTENTS

1.0 INTRODUCTION

The knowledge of the orderly arrangement of amino acids in a peptide (peptide sequence) is very important in predicting its function, particularly in diseased condition. Amino acid sequence of peptides and proteins is also useful in determining evolutionary relationships among the proteins and the plants/animals that produce them. This unit is concerned with the various steps to be followed in order to arrive at a peptide/protein primary structure (i.e. its amino acid sequence).

2.0 OBJECTIVES

By the end of this unit, you should be able to:

- state the preparative steps to sequencing of peptides and proteins
- define amino acid composition and explain how it is gotten
- state reagents use for peptide sequencing and their specific roles
- describe how peptides and proteins are sequenced.

3.0 MAIN CONTENT

3.1 General Outline for Sequencing Peptides and Proteins

In 1958, Sanger sequenced the first protein and insulin, and was given the Nobel Prize in 1958 for this work. The procedure for the determination of primary structure of proteins and peptides involves the following major steps:

- establishing the number of polypeptide chain in the protein;
- cleavage of disulfide bridges available;
- separation and purification of individual polypeptide;
- determination of amino acid composition;
- fragmentation of individual polypeptide to yield small peptides that can be sequenced;
- separation and purification of the small peptides;
- determination of amino acid sequence of each of the small peptides;
- establishing the overall sequence of each polypeptide in the protein; and
- establishing the position of inter and intra disulfide-bridge, if any.

Steps 1 to 6 could be seen as the preparatory stages to the sequencing procedure. They are, however, by no means less important. In fact, failure in any of them may spell doom to the ultimate goal of sequencing a peptide or a protein. We shall now consider each of the steps.

3.2 Preliminary Steps to Sequencing

3.2.1 Establishing the Number of Polypeptide Chain in the Protein

Establishing the number of chemically distinct polypeptides available for sequencing can be achieved by identifying N-terminal and/or C-terminal residues. To do this, the peptides/protein is made to react with dansyl chloride (reagents which react specifically with primary amines) to yield dansylated polypeptide. Acid hydrolysis of the later then yields dansylamino acid that can easily be identified via chromatography.

The most useful method for the determination of N-terminal amino acid is the method of Edman degradation. The method is named after its inventor, Pehr Edman. The method labels and removes only the amino-terminal residue from a peptide leaving all other peptide bonds. In the method, phenylisothiocyanate (PITC, Edman's reagent) is made to react

with N-terminal amino groups of the peptide or protein under mild alkaline conditions to form phenylthiocarbamyl (PTC) adduct.The treatment of this adduct with anhydrous trifluoroacetic acid releases the N-terminal residue as thiazolinone derivative, leaving other peptide bonds intact. The thiazolinone–amino acid is then extracted with organic solvent, converted to a more stable phenylthiohydantoin (PTH) by treatment with aqueous acid. The PTH-amino acid is finally identified using advanced chromatographic method. Figure 1 gives the various steps of Edman degradation.

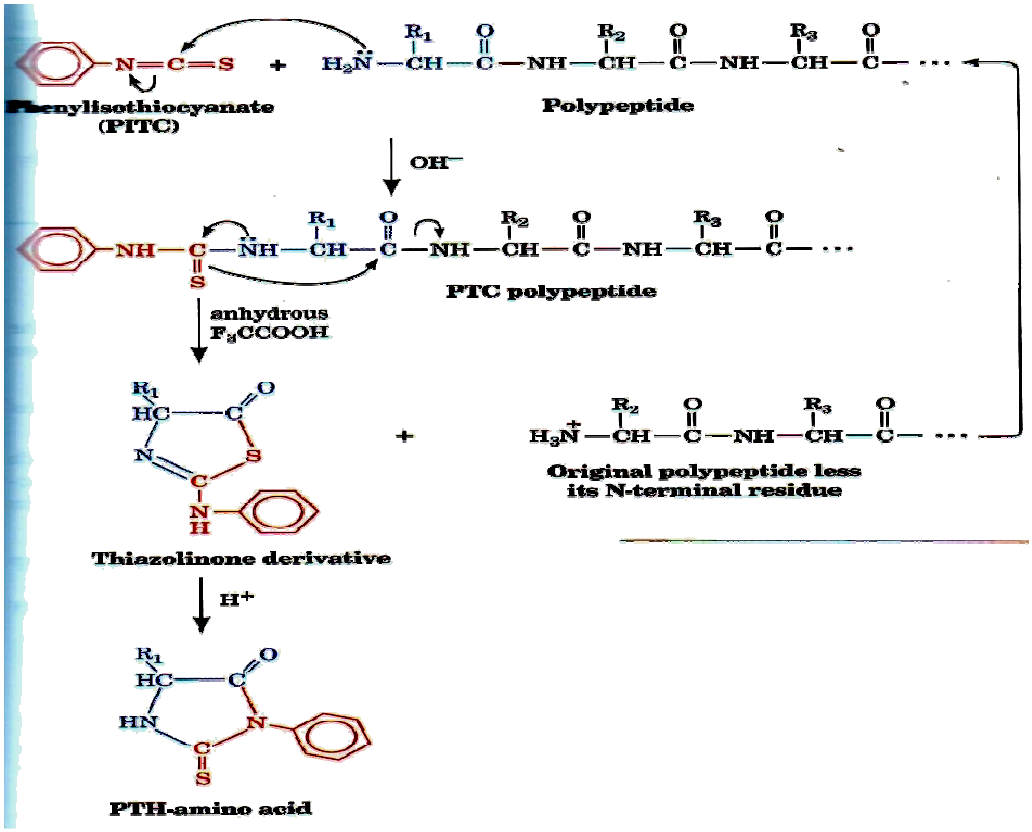


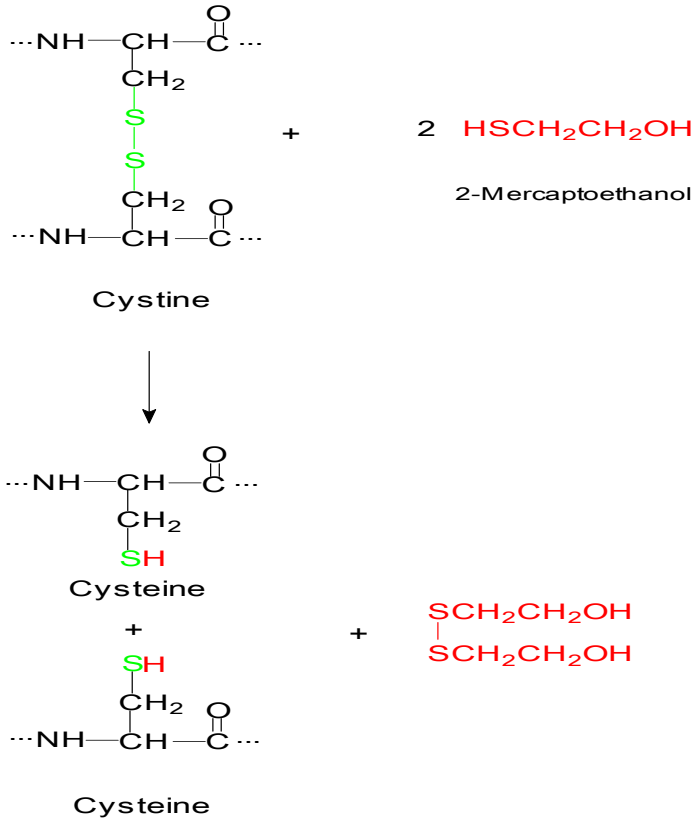
Fig. 1: Edman degradation
Source: Voet & Voet

Edman reaction has the advantage of repeating the cycle from the amino terminus inward until the entire sequence is determined. Edman degradation cycle is now carried out using an automated device known as sequenator. This device mixes reagent in a proper proportions, separate and identify the products generated and record the results.

3.2.2 Cleavage of Disulfide Bridges

Cleaving all the disulfide bonds in the protein/peptide could prevent their possible interference with the sequencing procedure. Irreversible

breakage of disulfide bonds is possible via reductive treatment with 2-mercaptoethanol as illustrated below:



Cleavage of Disulfide Bond
Source: Voet & Voet

3.2.3 Separation and Purification of Individual Polypeptide

If a peptide exists in its purified form, the next step will require the determination of its amino acid composition which would be discussed in the next section. On the other hand, polypeptide units arising from disulfide bond-cleavage must be separated and purified before their amino acid sequence can be determined. Units 3 and 4 of this module have dealt extensively with the various separation and purification procedures. The number of residues per peptide, which should also be known, can be estimated if its molecular mass is known.

3.2.4 Determination of Amino Acid Composition

The knowledge of amino acid composition is very important and its determination should precede its actual sequencing. The amino acid composition of a peptide is the number of each type of amino acid residue present in the peptide. This is determined by complete acid (6M

HCl) hydrolysis of the peptide followed by quantitative analysis of the liberated amino acids. Suffice it to mention that base (4M NaOH) and enzymatic hydrolysis is also possible. Table 1 gives the amino acid composition data of two proteins.

Table 1: Amino Acid Composition of Cytochrome c and Chymotrypsinogen

Amino acid	Number of residues per molecule of protein	
	Bovine cytochrome c	Bovine chymotrypsinogen
Ala	6	22
Arg	2	4
Asn	5	15
Asp	3	8
Cys	2	10
Gln	3	10
Glu	9	5
Gly	14	23
His	3	2
Ile	6	10
Leu	6	19
Lys	18	14
Met	2	2
Phe	4	6
Pro	4	9
Ser	1	28
Thr	8	23
Trp	1	8
Tyr	4	4
Val	3	23
Total	104	245

Source: Lehninger

3.2.5 Fragmentation of Individual Polypeptide to Yield Small Peptides that can be Sequenced

This step is necessary for proteins and large peptides because polypeptides longer than 40-100 residues cannot be directly sequenced. Fragmentation of longer polypeptides can be achieved by:

i. the use of endopeptidases e.g. Trypsin which cleaves the peptide bonds on the arboxyl terminus of Arginine and Lysine residues.
ii. the use of cyanogen bromide which specifically cleaves peptide bonds after Methionine residues.

3.2.6 Separation and Purification of the Small Peptides

After generating peptide fragments of the required size, these fragments must be separated and purified for subsequent amino acid sequence determination. Again, separation techniques discussed previously become useful for this step.

3.3 Determination of Amino Acid Sequence of Each Peptide Fragment

Each peptide can be sequenced through repeated cycles of the Edman degradation procedure discussed earlier

3.4 Deciphering the Complete Primary Structure

Once the sequencing step is completed, the remaining steps will be to establish the correct order of the fragments sequenced and the precise location of disulfide bonds in the native peptide or protein.

3.4.1 Establishing the Overall Sequence of Each Polypeptide in the Protein

Deciphering the order in which all the small peptide fragments generated in step 3.1.5 occurs in the original polypeptide(s), requires comparing the amino acid sequence of the peptide fragments produced by say, cleavage with trypsin with those generated by say cyanogen bromide, provided their specific cleavage sites overlap. Sometimes, a third or even a fourth cleavage using different cleaving reagent, is necessary before a good overlap is obtained.

3.4.2 Establishing the Position of Inter and Intra Disulfide-bridge

The identification of existing disulfide bonds starts with cleaving a fresh sample of the native protein under conditions that leave the disulfide bonds intact and the resulting fragments separated. Fragments having cysteine residues are known by examining the amino acid composition of all the fragments generated. Cysteine residues with free sufhydryl group (those not forming sulfide bonds) are identified by reacting the peptides with radioactive iodoacetic acid. The cysteine-containing

fragments are then subjected to Edman degradation from where the location of cysteine(s) is/are readily surmised.

4.0 CONCLUSION

Amino acid sequence of peptides and protein is a vital information required for the proper understanding of their tertiary/quaternary structure and function.

5.0 SUMMARY

In this unit, you have learnt that:

- sequencing peptides and proteins requires getting them first in purified form
- where disulfide bridges exist, they must be cleaved to avoid interference
- sequencing is better done if the peptides are not longer than 40-100 amino acid residues long. Chemical and enzymic approaches are available towards generating shorter fragments
- Edman degradation is a widely used sequencing procedure that identifies amino acids singly, beginning with the N-terminal residues inwards
- establishing the correct order of peptides and the location of disulfide bonds are very important in deciphering the actual amino acid sequence of the peptide or protein of interest.

6.0 TUTOR-MARKED ASSIGNMENT

1. Outline the major steps involved in the sequencing of a peptide that does not have cysteine residue in it.

2a. Mention the reagents that are necessary for sequencing peptides/proteins.

b. Describe the Edman degradation procedure.

7.0 REFERENCES/FURTHER READING

Thomas, M. Devlin (ed.) (2006). *Textbook of Biochemistry with Clinical Correlations*. Canada: Wiley-Liss.

David, L. Nelson & Michael, M. Cox (2005). *Lehninger Principles of Biochemistry*. New York: W. H. Freeman and Company.

MODULE 3 PROTEINS

UNIT1 PROTEINS NATURE, PROPERTIES, EXAMPLES AND BIOLOGICAL FUNCTIONS

CONTENTS

1.0 INTRODUCTION

Proteins constitute one of the major classes of biomolecules. They are known to perform various functions. This unit discusses the nature, properties and major functions of some biologically active proteins.

2.0 OBJECTIVES

By the end of this unit, you should be able to:

- describe the nature of protein
- state some basic properties of proteins
- give examples of some proteins
- ascribe functions to different proteins.

3.0 MAIN CONTENT

3.1 Nature of Protein

Proteins are polymers of amino acids. Some proteins exist as single polypeptide chains, while others occur as two or more polypeptides associated noncovalently. Proteins belonging to the later group are called multisubunit proteins. Proteins may occur as simple proteins, having only amino acid residues or as conjugated proteins, possessing permanent associated chemical components in addition to amino acids.

The non-amino acid part of a conjugated protein is usually referred to as its prosthetic group. The conjugated proteins could be classified on the basis of the chemical nature of their prosthetic group. For example, in lipoprotein, the prosthetic group is lipid while in metalloprotein, it is a metal. Table 1 below gives more examples of conjugated proteins and their prosthetic groups.

Table 1: Examples of Conjugated Proteins and their Prosthetic Groups

Conjugated Protein	Prosthetic Group	Example
Lipoproteins	Lipids	β-lipoproteins in blood
Glycoproteins	Carbohydrates	Immunoglobin G
Phosphoproteins	Phosphate groups	Casein of milk
Hemoproteins	Heme (iron porphyrin)	Hemoglobin
Flavoproteins	Flavin nucleotides	Succinate dehydrogenase
Metalloproteins	Iron	Ferritin
	Zinc	Alcohol dehydrogenase
	Calcium	Calmodulin
	Molybdenum	Dinitrogenase
	copper	Plastocyanin

3.2 Properties of Proteins

Proteins are known to have the following major properties:

i. Proteins have varying molecular weights**.** The molecular weights of insulin, haemoglobin and immunoglobulins are 5700, 68000 and 150000 respectively.

ii. Absorption of ultraviolet light. They strongly absorb U.V. light in the region between 200 to 240 nm. This is due to the amount of aromatic amino acid residues in them.
iii. They are charged molecules. They have characteristic isoelectric points (pI). They could be affected by change in pH. Buffers are therefore, used to stabilise their tertiary structure.
iv. They vary in size and shape. The shape of some proteins is globular while the shape of some others is fibrous. This is expected since their amino acid sequence and compositions are also quite different. For example, albumin is oval while insulin is globular.
v. They have varying solubility property. Different agents such as salts, organic solvent such as ethanol, heat, heavy metals and organic acids have been used to precipitate proteins. Protein stability, therefore, could be affected by any of these agents.

3.3 Examples of Proteins

There are diverse types of proteins occurring in and performing different functions in organisms (both plants and animals). We shall only attempt to list some commonly occurring proteins and their natural sources.

Protein	**occurrence**
Haemoglobin	Red blood cells
Transferin	Blood
Albumin	Egg
casein	milk
Luciferin	Firefly
Keratin	Hair, scales, horn, wool, nails, feather
Melanin	Skin
Myosin	Muscle
Actin	Muscle
Immunoglobulin	Blood
Rhodopsin	Eye
Collagen	Bones, tendons, and ligaments
Enzymes	Plant and animal cells

3.4 Biological Functions of Proteins

Proteins occur all over the cells of organisms and perform various biological functions. The functions of a few proteins are discussed below:

i. Melanin – this protein is found in the skin and is known to be responsible for skin colouration;

ii. **Keratin-** is a structural protein of the hair, nails, horns etc of animals;
iii. **Haemoglobin-** this is the protein pigment of blood that transports oxygen and carbon dioxide in the body of animals;
iv. **Enzymes-** this class of proteins catalyse several bio-reaction occurring in the cells of living things;
v. **Myosin and Actin- are** muscle proteins responsible for muscular contraction and relaxation.
vi. **Immunoglobulins-**these are proteins which provide the body with defence against infections;
vii. **Rhodopsin-**this protein is located in the retina of the eye. It participates in the transmission of vision.
viii. **Collagen-**provides tensile strength to bones, tendons and ligaments.

4.0 CONCLUSION

You have learnt in this unit that proteins occur in all living tissues. They exhibit various properties and perform special biological functions.

5.0 SUMMARY

In this unit, you have learnt that:

- proteins are polymers of amino acids
- proteins have several properties which are functions of their chemical structures; and
- almost every biological process is associated with one protein or the other.

6.0 TUTOR-MARKED ASSIGNMENT

1a. Name five proteins you know.
b. Differentiate between simple and conjugated proteins.

2. Outline the properties of proteins.
3. State four biological functions of proteins.

7.0 REFERENCES/FURTHER READING

Thomas, M. Devlin (ed.) (2006). *Textbook of Biochemistry with Clinical Correlations*. Canada: Wiley-Liss.

David, L. Nelson & Michael, M. Cox (2005). *Lehninger Principles of Biochemistry*. New York: W. H. Freeman and Company.

Donald, Voet & Judith G. Voet (2004). *Biochemistry* (3rd ed.).USA: John Wiley & Sons, Inc.

UNIT 2 STRUCTURAL LEVELS OF PROTEINS

CONTENTS

1.0 INTRODUCTION

Protein structure has four levels of complexity. In this unit, you will learn about this hierarchical organisation of protein structure.

2.0 OBJECTIVES

By the end of this unit, you should be able to:

- enumerate all the levels of protein structure
- define the primary structure of protein as amino acid sequence
- describe the different conformations of secondary structure of protein
- describe the three-dimensional arrangement of atoms in protein structure
- explain the interactions of subunits of polypeptides in quaternary level of protein structure.

3.0 MAIN CONTENT

3.1 Introduction to Structural Levels of Proteins

The properties of a protein are largely determined by its three-dimensional structure, which in itself is a function of its amino acid sequence and the regular folded polypeptide backbone. The four levels of protein structure include: primary structure, secondary structure, tertiary structure and quaternary structures. It is important you pay close attention to the type of chemical force that is associated with each of these levels.

3.2 Primary Structure of Proteins

The primary structure of a protein refers to the sequence of amino acid residues in it. The primary structure is defined by all covalent bonds (peptide bonds and disulfide bonds) linking the amino acid residues in a polypeptide chain. The peptide bond is vital to the integrity of the structure of any protein molecule and is the basis of primary structure.

The disulfide bridge is much involved in the cross-linking between different parts of the polypeptide chain.

3.2.1 Primary Structure of Insulin

Each protein has a distinctive amino acid sequence. For example, the primary structure of insulin consists of two polypeptide chains (A and B) linked via two disulfide bridges/cross linkages as shown in Figure 1 below.

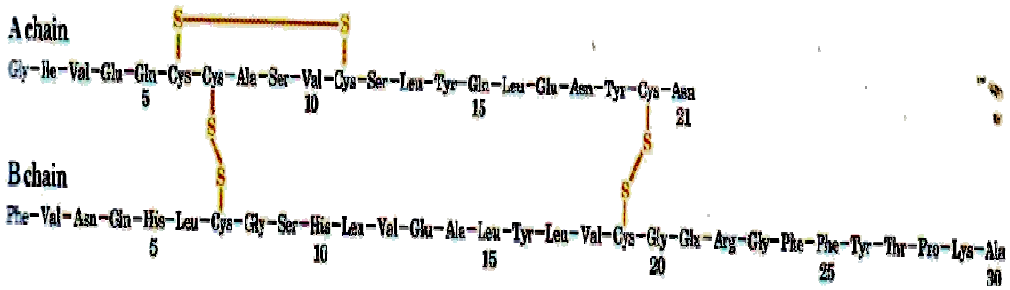


Fig. 1: Primary Structure of Insulin
Source: Voet & Voet

3.2.2 Primary Structure-Protein Function Relationship

The primary structure determines the folding pattern of a protein which in turn determines its function. A classical example in which there is a clear relationship between amino acid sequence of protein and its function is in sickle cell anaemia.

Sickle-cell Anaemia

Sickle cell anaemia is a painful life-threatening genetic disease. Those afflicted with this disease suffer repeated crisis brought about by physical exertion, leading to increase pulse rate, weakness, dizziness and difficulty in breathing.

Erythrocytes of individuals with this disease are fewer and abnormally long, thin and crescent compared to the biconcave shape of normal erythrocytes. This is because once the haemoglobin in these abnormal erythrocytes loses oxygen; it becomes insoluble and forms polymers that aggregate into tubular fibers. The blockage of the capillaries by these abnormal RBC could cause severe pain (or crisis).

The altered properties of haemoglobin **S** results from a single amino acid substitution, a Val instead of a Glu residue at position 6 in the two β chains [note the change from acidic to neutral amino acid]. This single change in the protein's primary structure affects the ability of the haemoglobin to bind and transport oxygen ultimately.

3.3 Secondary Structure of Proteins

The secondary structural level of proteins refers to common regular folding patterns of the polypeptide backbone. A few secondary structures occur widely in proteins, the most prominent ones being α helix and β conformations.

Secondary structures are dominated by the hydrogen bonds. Hydrogen bonds exist principally between the side chains of the hydrophilic groups of amino acids. The peptide backbone itself also contributes to the stability of a secondary structure. Its carbonyl groups (C=O) and its amino groups (N-H) are capable of forming hydrogen bonding with each other. As the groups are regularly spaced, it is not surprising that hydrogen bonding between them could give rise to regular structures.

3.3.1 α- Helix

This α-helix can be generated by winding the protein chain/polypeptide backbone around an imaginary axis through the helix, in such a way that the R groups are projected outward from the helical backbone, such that there are 3.6 amino acyl units per turn of the helix and an axial translation of 1.47 Å per unit as illustrated in Figure 2 below.

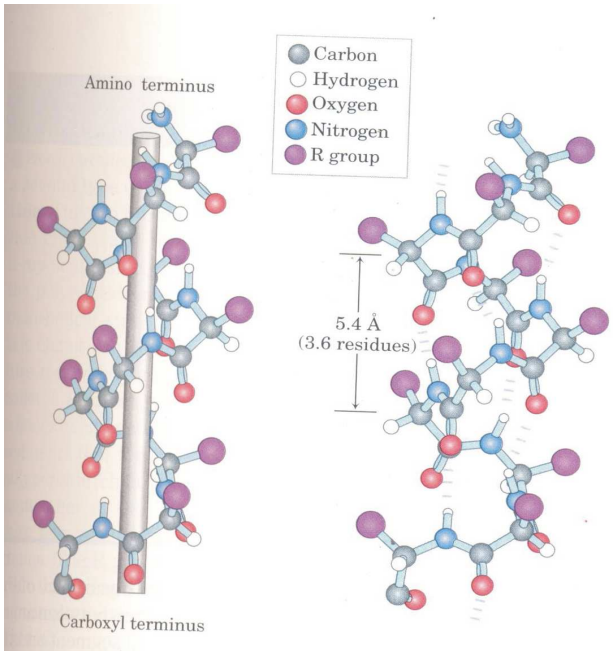


Fig. 2: **Alpha Helix**
Source: *Lehninger*

A helix may be characterised by the number, n, of polypeptide units per helical turn and by its **pitch**, *p*, the distance the helix rises along its axis per turn. This distance can also be referred to the axial distance.

3.3.2 β- Conformation/Structure

In β-conformation, another regular folded structure in naturally-occurring proteins, the polypeptide backbone is into a zigzag structure.

This zigzag polypeptide chain can be hydrogen bonded to adjacent chains to form **β-sheet**. It is also possible to have adjacent segments of a polypeptide chain forming a β-sheet via hydrogen bonding. Adjacent chains in a β-sheet can either be parallel (with the polypeptide chains having the same amino-to-carboxyl orientations) or antiparallel (with the chains having opposite orientation). The hydrogen bonding pattern in the parallel and antiparallel sheets are different. Figure 3 gives the structure of antiparallel β-pleated sheet. β –conformations exist in silk fibroin and fibroin of spider webs as β-keratin.

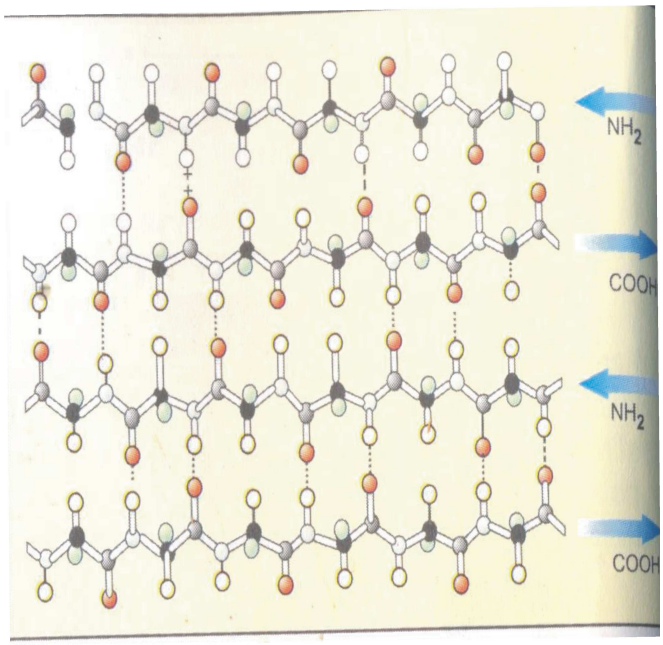


Fig. 3: Antiparallel Beta Pleated Sheet
Source: Vasudevan

β-Turn

Turns or loops, where polypeptide chains reverse direction, abound in proteins, particularly in the compact folded globular proteins. The turns connect successive runs of α-helix or β-conformations. The β-turns are more common and they connect the ends of two adjacent segments of an antiparallel β-sheet. A beta turn (β-turn) is a 180^{o} turn involving four amino acid residues, in which the carbonyl oxygen of the first residue forms a hydrogen bond with the amino-group hydrogen of the fourth residue. Gly and Pro residues occur often in β-turns.

3.3.3 Determination of Secondary Structure

X-ray diffraction has been used in demonstrating the presence of helices.

The method is based on the observation that when a parallel beam of monochromatic X-rays impinges upon a crystal or a fibre made up of a regular array of units, the light waves will be scattered by electrons of the atoms of the crystal or fibre and characteristic diffraction patterns will be obtained. X-ray diffraction patterns from naturally occurring proteins that exists in pleated sheet conformations are significantly different from those derived from naturally occurring proteins that exist predominantly in helical conformations.

3.4 Tertiary Structure of Proteins

The tertiary structure refers to the overall three-dimensional arrangement of all atoms in a protein. This includes all the bends and folds in the polypeptide chains of the protein. In tertiary structure, different segments of a protein's polypeptide chains are held by weak forces and sometimes by covalent bonds such as disulfide bridge. It is due to their tertiary structures that proteins adopt a globular shape/conformation which gives the lowest surface-to-volume ratio, thus minimising the interaction of protein with its surrounding.

3.5 Quaternary Structure of Proteins

Quaternary structure of protein refers to the arrangement of polypeptide chains in multichain protein. In other words, quaternary structural levels deal with the number of subunits in a given protein. These subunits are noncovalently bonded to each other, although there are other interactions among them. A multisubunit protein is also refered to as a multimer. Protein with two subunits is called a dimer while one with few subunits is often called an oligomer. Multimer can have identical subunits or repeating groups of nonidentical subunits. The repeating structural unit in such a multimeric protein, wether it is a single subunit or a group of subunits, is called protomer.

Quaternary Structure of Haemoglobin
A good example of an oligomeric protein is the haemoglobin. This transport protein contains four polypeptide chains and four heme prosthetic groups, in which the iron atoms are in the ferrous (Fe^{2+}) state.

The protein part, called globin, consists of two α chains (141 residues each) and two β chains (146 residues each). The subunits of haemoglobin are arranged in symmetric pairs, each pair having one α and one β subunit. Haemoglobin can, therefore, be viewed as a tetramer or a dimer of αβ protomer.

4.0 CONCLUSION

In this unit, you learnt that the four structural levels of protein have different structural complexities. In addition, you learnt that the type of bonds/interactions accounting for protein's stability at each level is also different.

5.0 SUMMARY

In this unit, you have learnt that:

- the sequence of amino acid in a polypeptide represents its primary structure. This level is defined by peptide bonds and disulfide bridges
- the function of a protein is related to its primary structure. Sickle cell anaemia is a classical example
- the secondary structures are regular folding patterns in polypeptide chains. The α-helix and the β-conformations are the most prominent secondary structures and are all characterised by hydrogen bonding
- the overall three-dimensional arrangement of all atoms in a protein, its tertiary structural level, is stabilised by weak forces
- arrangement of subunits in proteins constitutes the quaternary structures of such proteins.

6.0 TUTOR-MARKED ASSIGNMENT

1. Define the terms:

i. Primary structure of protein
ii. Secondary structure of protein
iii. Tertiary structure of protein
iv. Quaternary level of protein structure.

2. Using a good example describe how the primary structural level of a protein could be related to its function.

3a. Write short notes on the following:

i. α-Helix
ii. Multisubunits in haemoglobin

b. Calculate the number of amino acid residues in ten and a half turns made by a right handed α-helical structure. What will be the pitch?

7.0 REFERENCES/FURTHER READING

Thomas, M. Devlin (ed.) (2006). *Textbook of Biochemistry with Clinical Correlations*. Canada: Wiley-Liss.

David, L. Nelson & Michael, M. Cox (2005). *Lehninger Principles of Biochemistry*. New York: W. H. Freeman and Company.

Donald, Voet & Judith G. Voet (2004). *Biochemistry*. (3rd ed.). USA: John Wiley & Sons, INC.

Vasudevan, D.M. (2005). *Textbook of Biochemistry* (4th ed.).New Delhi, India: Jaypee.

UNIT 3 STABILITY OF PROTEINS –THE ROLES OF NONCOVALENT FORCES

CONTENTS

1.0 INTRODUCTION

Noncovalent forces play a very significant role in the stability of proteins. Whenever these forces are compromised, the native protein denatures. These forces are weak forces with strengths (1-7 $kcalmol^{-1}$) far less than those of covalent bonds ($\geq$ 50 $kcalmol^{-1}$). What is the nature of these noncovalent forces/interaction? At what level of protein structure do they contribute to protein stability? These are some questions we would answer in this unit.

2.0 OBJECTIVES

By the end of this unit, you should be able to:

- explain the process of protein folding
- enumerate the noncovalent forces that participate in protein stability
- explain the nature of the noncovalent forces
- explain how the compromise of these forces leads to denaturation of proteins' native conformations.

3.0 MAIN CONTENT

3.1 Protein Folding

The primary structure of a protein is able to fold spontaneously to its native conformation. The native conformation is the functional, folded conformations of proteins. Protein's internal residues direct this folding and is driven largely by noncovalent hydrophobic interactions among nonpolar residues. At first, folding is initiated by those short-range noncovalent interactions between a side chain and its nearest neighbours that form secondary structures in small regions of the polypeptide.

This is followed by long-range interaction between, say, two α-helices that come together to form stable supersecondary structures. The process, then, continuess until complete domains form and the entire polypeptide is folded. Figure 1 below, is a simulated folding pathway of a protein.

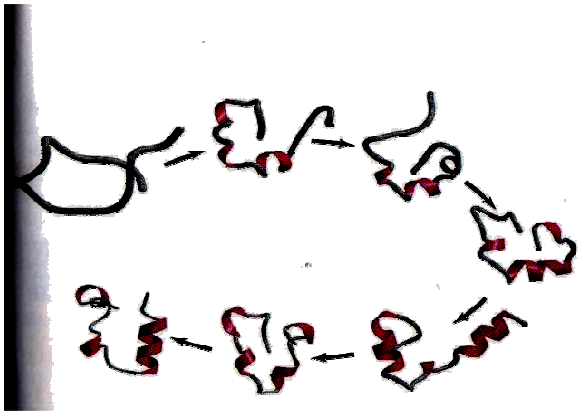


Fig. 1: Pathway of Protein Folding
Source: Lehninger

Protein folding does not proceed spontaneously all the time. They are in several instances, assisted by a class of proteins called molecular chaperones. These proteins interact with partially folded and improperly folded polypeptides, facilitating correct folding by preventing protein aggregation prior to folding.

Noncovalent forces do not only assist in the folding of polypeptide chains into unique native conformation, they also help to stabilise the native structure against denaturation- a process that involves lose of native conformation.

3.2 Forces that Stabilise Protein

The folding of polypeptides into their unique native conformation is largely governed by weak noncovalent forces, which also help to stabilise proteins. Stability here denotes tendency to maintain the native conformation. The major noncovalent forces include: hydrophobic interaction, hydrogen bonding, electrostatic forces, and van der Waals-London Dispersion Forces.

3.2.1 Hydrophobic Interaction

Hydrophobic interaction is the force that causes nonpolar substances and nonpolar portions of amphipathic molecules (molecules with both polar and nonpolar regions e.g. proteins) to minimise their contact with water.

Amphipathic substances can form, in aqueous solution, a stable structure known as micelle in which the hydrophobic portions cluster together so as to give the lowest interaction with the aqueous solvent, while the polar/hydrophilic portions of proteins are arranged to encourage or maximise contact with the water environments as shown in Figure 2 below.

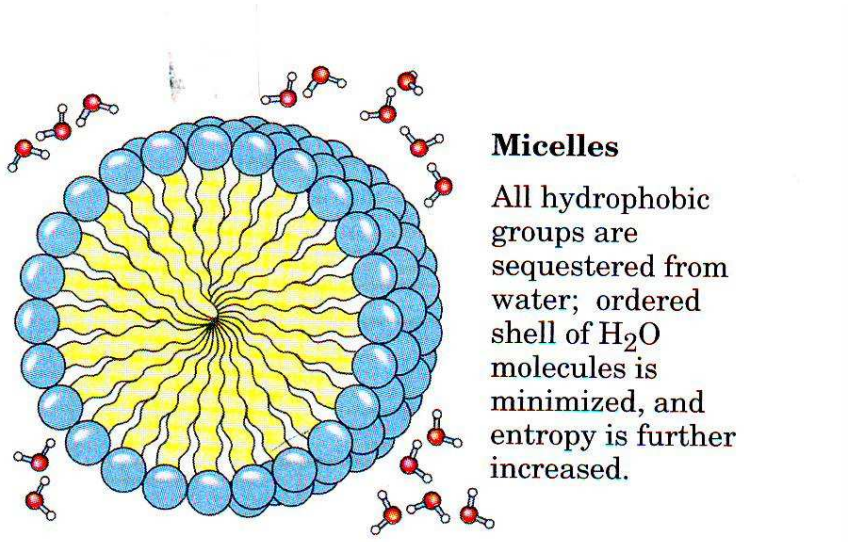


Fig. 2: **Micelle**
Source: *Lehninger*

Hydrophobic interaction is the result of the attempt by such a system to achieve greatest thermodynamic stability. Hydrophobic interactions alongside the short-range dispersion forces are major influences that cause proteins to fold into their native conformations and are among the

most powerful stabilising factors of protein structure, acting particularly at the tertiary and quaternary levels.

3.2.2 Hydrogen Bonding

Hydrogen bonding is formed when a hydrogen atom, covalently bonded to an electronegative atom, is shared with a second electronegative atom. In proteins, hydrogen bonding is found principally between side chains of hydrophilic group of amino acids. Functional groups of peptide backbone can also form hydrogen bonds. Most of the hydrogen bonds in proteins are local, meaning that they involve hydrogen bonding mates that are close together in sequence. The secondary structures of proteins are stabilised by hydrogen bonding. The α-helix and β-conformations are extensively hydrogen -bonded.

Although hydrogen bonding plays contributory role in thermodynamic stability of native conformation in proteins, it may not be a major driving force for folding. This is because peptide bonds and other hydrogen–bonding groups form hydrogen bonds, to water solvent in the denatured state. These bonds must be broken before the polypeptide can fold.

3.2.3 Electrostatic Interaction

Electrostatic interactions (ionic or salt linkages) are the strongest of all noncovalent forces. In proteins, they can occur between positive charges on ***His, Lys, Arg*** and α-amino groups and the negative charges of ***Asp, Glu***, and the α-carbonyl group. Depending on whether the interacting charges are the same or opposite, electrostatic forces could be repulsive or attractive.

Intramolecular ionic bonds are infrequently used in the stabilisation of protein structure. When they are used, it is often with great effect. For example, ionized groups stabilise interactions between proteins and other molecules (e.g. cofactors, prosthetic groups e.t.c.).

The interaction of ionized groups and solution form part of the balance of hydrophobic and hydrophilic forces. In addition, these interactions yield the dominant forces that place most charged groups of a protein on the outside of the folded structures.

3.2.4 van der Waals-London Dispersion Forces

These are the weakest of the noncovalent forces. The vander Waals force is a transient dipole. Transient dipole can interact with charged groups and with permanent dipoles. Similarly, they can induce dipole.

The interaction between a transient dipole and the dipole they induce are the most important contribution to the attractive forces between neutral atoms. They are known as London, Heitler or dispersion forces.

Dispersion forces may be among the most important in deciding the stability of the tertiary structure of protein since all the atoms of the protein are involved. Although the individual dispersion forces are of excessively short range, the overall resultant force of a very large number of such dipoles is of long-range interaction.

Such resultant force may influence the early folding-up phase process of a macromolecule by bringing its various parts close together, so that the shorter-range forces (ionic, hydrogen bonds) can take over and determine the final details of the structure.

Even though it is not easy to identify the weak dipole forces by inspection (unlike ionic and hydrogen bonds), they are, however, by no means less influential in protein structure.

3.3 Protein Denaturation

Ribosomes are the cellular organelles that are known for protein synthesis. At first, they produce polypeptide chains of amino acids residues that must fold into functional native conformation. The stability of this native conformation can be lost. Denaturation is the loss of the native secondary, tertiary, and /or quaternary structure of protein. The primary structure is not normally changed by denaturation. This means that the denaturating agents generally affect the weak interactions, leaving the covalent bonds of the polypeptide intact. Sometimes, the denatured state is associated with loss of protein's function although this is not always the case.

Denaturating Agents
Proteins are denatured by a variety of conditions and substances:

1. **Heating**- when protein is heated, its conformationally sensitive properties such as optical rotation, viscosity and UV absorption, change abruptly over a narrow range of temperature as shown in Figure 3 below.

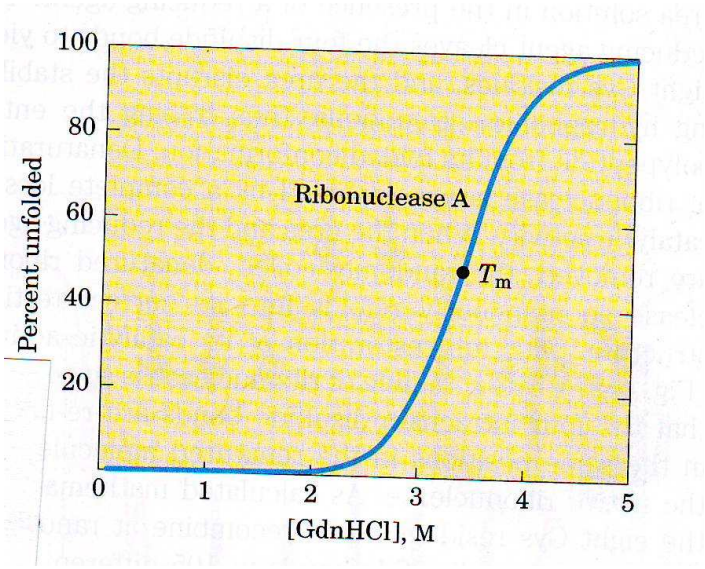


Fig. 3: Abrupt Change in Folding Pattern of Ribonulease A
Source: Lehninger

The abruptness of the change suggests that the unfolding is a cooperative process. Any partial unfolding of the structure destabilises the remaining structure, which must simultaneously collapse into the random coil. The temperature at the midpoint of this process is known as the protein's melting temperature**,** $\boldsymbol{T_m}$. Most proteins have T_m values far less than 100^{o}C.

2. **Change in pH**- variations in pH alter the ionisation states of amino acid side chains. This leads to alteration in the net charge on the protein and changes in H bonding requirements.

3. **Treatment with Detergents**- these substances affect the protein structures by tempering with hydrophobic interactions responsible for the protein's native structure.

4. **High Concentration of Organic Solvents**- organic solvents such as alcohol and acetone also interfere with the hydrophobic interactions thereby stabilising protein structure.

5. Certain solutes like urea and guanidine hydrochloride also denature proteins. Also, various salts show more variable influences on protein structure.

Unfolded proteins under denaturating agents sometimes regain their folded native conformation and biological function if the denaturant(s) is/are removed. This process is called renaturation. Figure 4 shows how

ribonuclease denatures and renatures in the presence and absence of a concentrated solution of urea in the presence of a reducing agent.

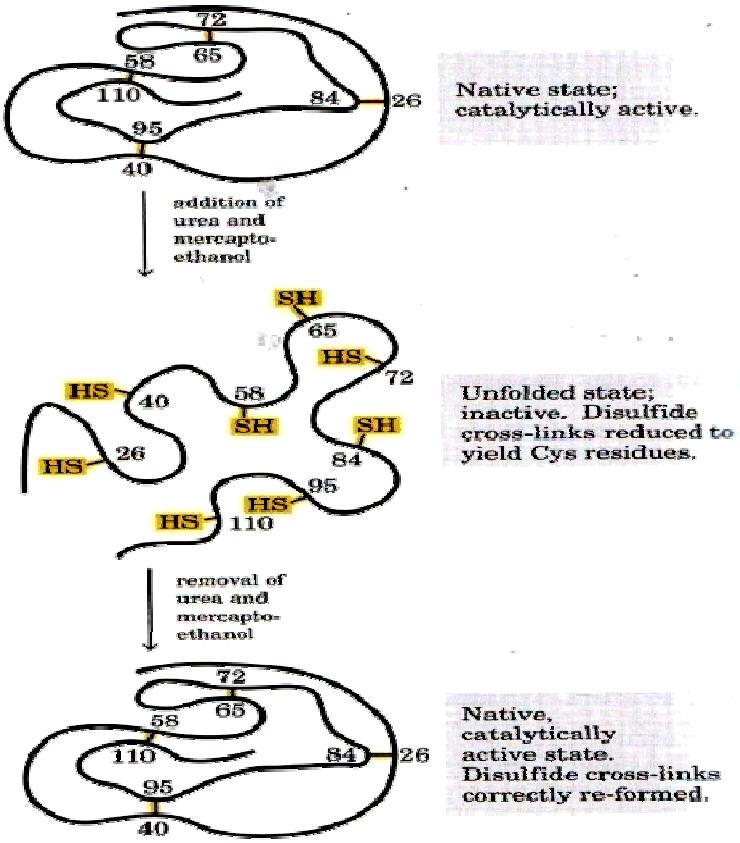


Fig. 4: Denaturation and Renaturation of Ribonuclease
Source: Lehninger

Renaturation clearly shows that the primary structure of protein determines its tertiary functional structure.

4.0 CONCLUSION

Even though individually weak, the large number of noncovalent interactions within a protein's secondary, tertiary and quaternary structural levels provides a large constraining factor that promotes protein folding. Protein stability arises from a balance among these interactions.

5.0 SUMMARY

In this unit, you have learnt that:

- protein folding begins spontaneously but the process is thereafter assisted

- protein folding and stability are governed by weak noncovalent forces
- protein native conformation can be affected by various denaturating conditions/agents.

6.0 TUTOR-MARKED ASSIGNMENT

1. Write succinctly on the concept of protein folding.
2. What do you understand by the term protein stability? Outline the forces that account for protein stability.
3. With respect to protein's structure, what would you say are the major differences between hydrogen bonding and van der Waals forces?

7.0 REFERENCES/FURTHER READING

David L. Nelson & Michael, M. Cox (2005). *Lehninger Principles of Biochemistry*. New York: W. H. Freeman and Company.

Donald, Voet & Judith G. Voet (2004). *Biochemistry*(3rd Ed.). USA: John Wiley & Sons, Inc.

Thomas, M. Devlin (ed.) (2006). *Textbook of Biochemistry with Clinical Correlations*. Canada: Wiley-Liss.

UNIT 4 CLASSIFICATION OF PROTEINS

CONTENTS

1.0 INTRODUCTION

Proteins have various shapes, structures and functions. These parameters have been used to classify proteins. This unit examines the different approaches to protein classification.

2.0 OBJECTIVES

By the end of this unit, you should be able to:

- discuss some criteria for classifying proteins
- list the major proteins and give their biological functions
- differentiate between simple and conjugate proteins
- give examples of globular and fibrous proteins.

3.0 MAIN CONTENT

3.1 Criteria for Classification of Proteins

The major criteria used in the classification of proteins include:

i. function
ii. structural complexity
iii. shape

3.2 Classification on the Basis of Function

3.2.1 Catalytic Proteins

Catalytic proteins are the class of proteins that accelerate the rate of biological reactions. This class is by far the largest, since virtually all the steps in metabolism are catalysed by enzymes that are mostly proteinuous by nature.

3.2.2 Transport Proteins

Transport proteins carry specific substances from one place to another.

Haemoglobin, for example, transports oxygen from the lungs to different parts of the body and carbon dioxide back to the lungs for exhalation.

Serum albumin, another transport protein, transports fatty acids from adipose tissues to various organs. Transmembrane proteins like Na^+/K^+ ATPase, glucose and amino acid transporters, are proteins located in the cellular membranes and thy help to transport ions and metabolites across these membranes.

3.2.3 Structural Proteins

These proteins help in the creation and maintenance of biological structures. They provide strength and protection to cell and tissues.

Examples of structural proteins include α-keratin found in hair, horns and fingernails and collagen found in bone, connective tissues, tendons, cartilage and hide. Others are elastin (in ligaments) and and fibroin (in spider’s web).

3.2.4 Regulatory Proteins

Proteins that help to regulate the ability of other proteins to carry out their physiological functions are called regulatory proteins. Insulin is a good example. It helps to regulate glucose metabolism in animals. Other hormones that are regulatory proteins include somatotropin and thyrotropin.

3.2.5 Storage Proteins

Storage proteins provide reservoir of essential nutrient. Since nitrogen is a limiting nutrient for growth, organisms have exploited proteins as a source of nitrogen in their time of need. For example, ovalbumin, the protein of egg white, provides the developing embryo with a source of nitrogen during its development within the egg. Casein of milk serves as the major nitrogen source for mammalian infants. Again, some seeds of higher plants contain proteins that act as good sources of nitrogen during germination.

3.2.6 Scaffold Proteins

The scaffold proteins provide anchorage for other proteins. They facilitate the binding of such proteins with other structures in the cell.

3.2.7 Protective and Defensive Proteins

Immunoglobulins protect the body by recognising, binding and neutralising foreign molecules. Thrombin and fibrinogen are blood-clothing proteins whose importance is in the prevention of continuous blood loss. Other proteins like phospholipase found in snake venom play a rather defensive role.

3.3 Classification on the Basis of Structural Complexity

3.3.1 Simple Proteins

Simple proteins contain only amino acids. The albumins of egg, globulins, protamines, lectins are all simple proteins.

3.3.2 Conjugate Proteins

Conjugated proteins have permanently associated chemical components, in addition to amino acids. Some subclasses of this group of proteins include: glycoproteins, lipoproteins, and metalloproteins e.t.c. Table 1 below lists some examples of conjugated proteins.

Table 1: Examples of Conjugated Proteins

Conjugated protein	Protein part	Prosthetic group
Hemoglobin	Globin	Heme
Nucleoprotein	Histones	DNA
Rhodopsin	opsin	11-cis-retinal
Succinate dehyrogenase	Protein	Riboflavin
Ferritin	apoferritin	Iron
ceruloplasmin	Apoceruloplasmin	Copper

Source: Vasudevan and Sreekumaris

3.4 Classification on the Basis of Shape

3.4.1 Globular Proteins

In terms of shape, globular proteins are spherical or oval. They are easily soluble. Examples are albumins and globulins.

3.4.2 Fibrous Proteins

Fibrous proteins are elongated or needle shaped. These proteins are largely insoluble and are very resistant to digestion. Elastin, collagen and keratin are good examples.

3.5 Classification on the Basis of Nutritional Importance

3.5.1 Complete Proteins

They contain all the essential amino acids in the proportion they are required. A good example is casein found in milk. This class of protein is nutritionally rich and is also called first class protein.

3.5.2 Incomplete Proteins

Incomplete proteins lack at least one essential amino acid and as such, cannot promote growth in growing individuals. For example protein from cereals lacks lysine.

4.0 CONCLUSION

Classification of proteins does not follow one particular criterion. It is possible for a single protein to be classified as simple as well as transport protein depending on the parameter used.

5.0 SUMMARY

In this unit, you have learnt that:

- proteins can be classified according to some criteria
- proteins could play transport, structural, protective, storage, regulatory or even defensive roles when they exist in organism
- globular proteins differ from fibrous protein not only in their shape but also in their solubility.

6.0 TUTOR-MARKED ASSIGNMENT

1a. State the criteria used for the classification of proteins?
b. Name five proteins and their functions.

2a. Write short note on the classification of proteins.
b. Using clear examples; differentiate between globular and fibrous proteins.

7.0 REFERENCES/FURTHER READING

Thomas, M. Devlin (ed.) (2006). *Textbook of Biochemistry with Clinical Correlations*. Canada: Wiley-Liss.

David, L. Nelson & Michael M. Cox (2005). *Lehninger Principles of Biochemistry*. New York: W. H. Freeman and Company.

Vasudevan, D. M. & Sreekumari, S. (2005). *Textbook of Biochemistry*.
New Delhi: Jaypee Brothers.

UNIT 5 ENZYMES, CO-ENZYMES AND VITAMINS

CONTENTS

1.0 INTRODUCTION

You have now known that proteins perform a broad range of biological functions and that there are several classes of proteins. Enzymes constitute one of these classes. Enzymes are biocatalysts required for most metabolic reactions of living organisms. This unit will focus on enzymes and co-enzymes (another class or organic substance needed by enzyme).

2.0 OBJECTIVES

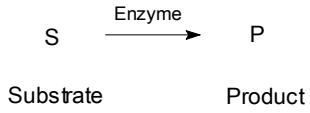
By the end of this unit, you should be able to:

- define enzyme and name some of them
- mention the major classes of enzymes
- outline the key properties of enzymes and explain how enzymes perform their work
- differentiate the roles of co-enzymes, co-factors and vitamins in relation to enzyme activity.

3.0 MAIN CONTENT

3.1 Definition of Enzymes

Enzymes are biocatalysts. They speed up the rate of metabolic reactions occurring in living cells.
The substance upon which an enzyme acts is called substrate and this substrate is usually converted to product at the end of the enzyme-catalysed reaction as represented by the scheme below:



3.2 Examples of Enzymes and the Reactions they Catalyse

There are diverse numbers of enzymes associated with different metabolic steps. Some of these enzymes, their substrates and the products of the reactions they catalyse are given in the Table 1 below:

Table 1: Enzymes, their Substrates and Products

Enzyme	Substrate	Product
Ptyalin/Salivary Amylase	Starch	Complex Sugar
Amylopsin	carbohydrates	Simple sugars
Trypsin	Proteins	Simpler peptides
Lipase	Fats and Oil	Fatty acids
Cellulase	Cellulose	Glucose
Alcohol Dehydrogenase	Alcohol	Aldehyde

3.3 Enzyme Nomenclature and Classification

3.2.1 Nomenclature of Enzymes

Early scientists have assigned arbitrary names to enzymes: trypsin and chymotrypsin, for example, are names still being used. Sometimes, enzymes are simply named by adding the suffix '-ase' to the substrate.

Thus, lactase acts on lactose, amylase on amylose e.t.c.

Generally speaking, there are two types of enzyme nomenclatures in common use: the trivial or working names and the systematic names.

The trivial names are normally and sufficiently shorter for general use, though not necessarily very systematic. They are usually assigned arbitrarily. Amylase, trypsin, and lactase are examples of trivial names.

You are likely to meet more of these trivial names than the systematic names in your further reading.

The systematic name of an enzyme, on the other hand, shows the action of the enzyme as exactly as possible, thus identifying the enzyme precisely. Systematic names are assigned in accordance with definite rules laid down by Enzyme Commission. The systematic names are the ones used in enzyme classification. Lactate: NAD^+ Oxidoreductase and Xylitol: NAD^+ 2-Oxidoreductase are the systematic names of Lactate dehydrogenase and D-Xylulose reductase respectively. The details of enzyme nomenclature are beyond the scope of this course. You are, however, encouraged to consult the textbooks cited at the end of this unit for your further reading.

3.3.2 Classification of Enzymes

The standard system for the classification of enzymes has grouped enzymes on the basis of the reactions they catalyse. There are six major reaction types/classes to which six major classes of enzymes have been identified. These include;

1. Oxidoreductases – catalyse transfer of hydrogen.
2. Transferases – catalyse transfer of groups other than hydrogen.
3. Hydrolases – these cleave bonds and add water.
4. Lyases – cleave without adding water.
5. Isomerases – intramolecular transfers.
6. Ligases – catalyse ATP-dependent condensation of two molecules.

Enzymes are assigned code numbers (also called enzyme commission numbers**)**. These code numbers, prefixed by EC, contain four elements separated by points, with the following meaning:

(i) The first number shows to which of the six main divisions (classes mentioned above) the enzyme belongs;
(ii) The second figure indicates the subclass;
(iii) The third figure gives the sub-subclass; and
(iv) The fourth figure is the serial number of the enzyme in its sub-subclass.

For example the code [EC 3.2.1.18] has been assigned to the enzyme ‘sialidase’, indicating that it is a hydrolase.

3.4 General Properties of Enzymes

The following properties are characteristic of enzymes:

- with the exception of small group of catalytic RNA molecules, all enzymes are proteins.

  Thus, all the structural levels are important to their catalytic activities.

  Like other proteins, enzymes have molecular weights ranging from about 12,000 to more 1 million. A number of enzymes require co-enzymes (non protein) to activate them.
- enzymes exhibit high catalytic power; and
- enzymes have active sites. This is a three-dimensional cleft through which enzymes interact reversibly with their substrates to form enzyme-substrate complex in the process of forming the product as illustrated below:

$$E + S \rightleftharpoons ES \longrightarrow E + P$$

The key features of the active site include:

i. An active site is a small portion of the total volume of the enzyme molecule.
ii. It is three-dimensional.
iii. It binds to substrates via relatively weak forces.
iv. Active sites are clefts or depression where water is excluded. These clefts possess
   polar residues that are essential for the enzyme activity.
v. Enzymes are specific in the type of reactions they catalyse. For example, amylase will act on amylose and not glucose. Sometimes, these specificities are group specificities and in some instances stereo-specificities.
vi. Enzymes are highly regulated through a balance between their synthesis and their degradation. For instance, at gene level their production can be controlled by induction and repression mechanisms of the coding gene.
vii. Enzymes are sensitive to temperature and are denatured at high temperature.
viii. Enzymes are also sensitive to variation in pH (acidity or alkalinity of their environment).
ix. Enzymes can be inactivated by inhibitors.

3.5 How Enzymes Work

Enzyme catalysis takes place within an area in the enzyme known as active site. Usually, catalysis begin with the substrate (the molecule to be acted upon) getting bound by the enzyme at its active site. This is followed by lowering the activation energy so that the reaction will proceed at a faster rate than it would in the absence of the enzyme.

The binding of the substrate to the enzyme could be in a form of Lock – and – key (where the active site assumed a rigid shape which accommodates only substrate with identical shape fits in) or via Induced fit approach in which the enzyme undergoes a change in conformation when the substrate binds. The induced fit theory of substrate binding at the active site is the most accepted theory. Upon binding, the substrate gets sequestered completely from solution forming the enzyme – substrate complex. This complex is very central to enzyme action.

3.5 Co-enzymes and Vitamins

Co-enzymes are low molecular weight organic substances which enzymes could require for their catalytic activities. Often, co-enzymes are derivatives of vitamins. Table 2 below lists co-enzymes and the vitamins from which they are derived.

Table 2: Some Co-enzymes, their Vitamin Source and the Reaction they Catalyse

Coenzyme	Vitamin	Reaction Mediated
Biotin	Biotin	Carboxylation
Cobalamin (B_{12})	Cobalamin (B_{12})	Alkylation
Coenzyme A	Pantothenate	Acyl transfer
Flavin coenzymes	Riboflavin (B_2)	Oxidation–reduction
Lipoic acid		Acyl transfer
Nicotinamide coenzymes	Nicotinamide	Oxidation–reduction
Pyridoxal phosphate	Pyridoxine (B_6)	Amino group transfer
Tetrahydrofolate	Folic acid	One-carbon group transfer
Thiamine pyrophosphate	Thiamine (B_1)	Carbonyl transfer

Source: Thomas Devlin

Sometimes, the additional chemical group an enzyme may require is/are one or more inorganic ions. They are hence called co-factors. A co-enzyme or co-factor that is very tightly (covalently) bound to the enzyme protein is a prosthetic group. A complete catalytically active enzyme together with its bound co-enzyme and/or metal ion is called holoenzyme. The protein part of such an enzyme is called the apoenzyme or apoprotein.

4.0 CONCLUSION

Enzymes are important class of protein known for their catalytic function. It can be said that life is virtually impossible without their activities.

5.0 SUMMARY

In this unit, you have learnt that:

- most metabolic reactions are catalysed by enzymes
- enzymes are mainly protein and have characteristic features such as their specificities and sensitivity to variation in temperature and pH
- enzymes are classified on the basis of the reaction they catalysed
- the lowering of activation energy in enzyme-catalysed reactions occurs with the formation of enzyme-substrate complex
- co-enzymes and vitamins are very essential to the functioning of enzymes.

6.0 TUTOR-MARKED ASSIGNMENT

1a. What are enzymes?
b. All enzymes are protein but not all proteins are enzyme. Discuss.

2. Outline the major classes of enzymes. Mention the reactions catalysed by any two classes.
3. Write short notes on the following:

   a. Lock and Key model of enzyme-catalysed reaction
   b. Co-enzyme
   c. Co-factor
   d. Holoenzyme

7.0 REFERENCES/FURTHER READING

Thomas, M. Devlin (ed.) (2006). *Textbook of Biochemistry with Clinical Correlations*. Canada: Wiley-Liss.

David, L. Nelson & Michael, M. Cox (2005). *Lehninger Principles of Biochemistry*. New York: W. H. Freeman and Company.