NATIONAL OPEN UNIVERSITY OF NIGERIA

SCHOOL OF SCIENCE AND TECHNOLOGY

COURSE CODE: BIO 305

COURSE TITLE: MOLECULAR BIOLOGY

BIO 305 MOLECULAR BIOLOGY

TABLE OF CONTENTS **PAGE**

MOLECULAR BIOLOGY

INTRODUCTION

Molecular biology is concerned with understanding the mechanisms responsible for the transmission and expression of the genetic information that ultimately governs cell structure and function. Understanding the molecular biology of cells is fundamental to all biological sciences. Because of its growing number of practical applications in Agriculture, Biotechnology, and Medicine.

All cells share a number of basic properties (a kind of underlying unity of cell biology) that are particularly apparent at the molecular level. Such unity has allowed scientists to choose simple organisms like bacteria as models for many fundamental experiments, with the expectation that similar molecular mechanisms are operative in organisms as diverse as *E. coli* and humans.

UNIT 1: MOLECULAR BIOLOGY – AN OVERVIEW

CONTENTS

1.0 INTRODUCTION

Molecular biology is concerned with understanding the mechanisms responsible for the transmission and expression of the genetic information that ultimately governs cell structure and function. Understanding the molecular biology of cells is fundamental to all biological sciences, because of its growing number of practical applications in agriculture, biotechnology and medicine.

2.0 OBJECTIVES

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

1. Define molecular biology.
2. Outline its relationship with other biological sciences.
3. Describe the techniques of molecular biology.
4. Identify the significance of molecular biology.

3.0 MAIN BODY

3.1 Meaning of Molecular Biology

The term molecular biology was first used in 1945 by William Astbury to refer to the study of the chemical and physical structure of biological molecules. Contemporary molecular biology is the study of life at the molecular level; and it is principally concerned with the understanding and interactions between the various systems of a cell, including the interactions between the different types of DNA, RNA and protein biosynthesis and how these reactions are regulated – the mechanisms responsible for the transmission and expression of the genetic information that ultimately governs cell structure and function i.e. the study of the process of replication, transcription and translation of genetic

material. The central dogma of molecular biology where the genetic material (DNA) is transcribed into RNA and then translated into protein provides the starting point for understanding the field.

3.2 Relationship With Other Biological Sciences

Molecular signs of evolution show that evolutionary relationships among species are reflected in their DNA and proteins – in their genes and gene products. Researchers in molecular biology use techniques and ideas from other areas of biology and chemistry particularly genetics – the transfer of biological information from cell to cell, from parents to offsprings and thus from generation to generation and to the effects on organisms – and biochemistry – the study of the chemical substances and vital processes occurring in living organisms particularly the role, the function and the structure of biomolecules.

Much of the work in molecular biology is quantitative and require ideas from computer science in bioinformatics and computational biology. Molecular genetics – the study of gene structure and function is a prominent sub-field of molecular biology.

Many other areas of biology focus on molecules, either directly studying their interactions in their own right such as in cell biology and developmental biology or indirectly where the techniques of molecular biology are used to infer historical attributes of populations or species as in fields of evolutionary biology such as population genetics and phylogenetics.

3.3 Techniques of Molecular Biology

To understand different biochemical events of prokaryotic and eukaryotic cells at molecular level and be able to characterise, isolate and manipulate molecular components of cells and organisms, a wide array of bio-physico-chemical techniques are used in molecular biology. These include:

3.3.1 Expression of Cloned Genes

Molecular cloning enables the determination of the nucleotide sequences of genes and also provide new approaches to obtaining large amounts of proteins for structural and functional characterisation. In expression colony, DNA coding for a protein of interest is cloned using polymerase chain reaction or restriction enzymes into a plasmid or phase vector known as expression vector.

The plasmid can be inserted into either bacterial or animal cells; DNA coding for a protein of interest is now inside a cell, and the protein can now be expressed. A variety of systems, such as inducible promoters and specific cell-signalling factors, are available to help express the protein of interest at high levels. Large quantities of a protein can then be extracted from the bacterial or eukaryotic cell. The protein can be tested for enzymatic activity under a variety of situations, the protein may be crystallised so its tertiary structure can be studied, or in the pharmaceutical industry, the activity of new drugs against the protein can be studied.

3.3.2 Polymerase Chain Reaction

Molecular cloning allows individual DNA fragments to be propagated in bacteria and isolated in large amounts. The

polymerase chain reaction is a versatile technique for copying DNA, and allows a single DNA sequence to be copied repeatedly, or altered in predetermined ways; essentially it is used for repeated replication of a defined segment of DNA. Single DNA molecules can thus be amplified to yield readily detectable quantities of DNA that can be isolated and quantitatively measured.

3.3.3 Gel Electrophoresis

Gel electrophoresis is one of the principal tools of molecular biology. It is a common method in which molecules (DNA, RNA and proteins) are separated based on the rates of their migration in an electric field. A gel, usually formed from agarose or polyacrylamide, is placed between two buffer compartments containing electrodes. The sample is then pipetted into preformed slots in the gel, and the electric field is turned on; the gel acts like a sieve, selectively retarding the movement of larger molecules. Smaller molecules therefore move through the gel more rapidly, allowing a mixture of nucleic acids to be separated on the basis of size.

3.3.4 Nucleic Acid Hybridisation

Cloning enabled the isolation and characterisation of individual genes. However, understanding the role of genes within cells, requires analysis of the intracellular organisation and expression of individual genes and their encoded proteins. Nucleic acid hybridisation is a method for detecting and analysing sequences of homologous DNA. This enables the mapping of genes, to chromosomes, the analysis of gene expression, and the

localisation of proteins to subcellular organelles. In this way, it is possible to study genetic differences between organisms or individuals.

Hybridisation can be achieved by southern or northern blotting.

Southern blotting is a method for probing for the presence of a specific DNA sequence within a DNA sample and it enables a researcher to determine not only whether a particular sequence is present within a sample of DNA, but how many such sequences there are; and the size of the restriction fragments that contain these sequences.

Messenger RNA can also be subjected to hybridisation analysis, in an analogous process known as Northern blotting used to study the expression patterns of a specific type of RNA molecule and is essentially a combination of denaturing RNA gel electrophoresis and a blot. In this process RNA is separated based an size and is their transferred to a membrane that is then probed with a labelled complement of a sequence of interest. It is used to determine whether a particular gene is made into MRNA, how much of that MRNA is present, and whether the abundance of that specific MRNA changes at different stages of development or in response to certain regulatory signals; that control gene expression.

3.3.5 Restriction Fragment Length Polymorphism Analysis

DNA fragments that result from cutting a particular piece of DNA wit a specific restriction enzyme give a characteristic pattern of bands upon gel electrophoresis. Fact band corresponds to a DNA restriction fragment of a certain length. Such difference are

called restriction fragment length polymorphisms (RLFPs) serving as genetic marker for a particular location the genome. A given FLP marker frequently occurs in numerous variants in a population and is inherited in a mendelian fashion. Genetic markers are used for making linkage maps. RLFP analysis is important in the diagnosis of genetic disorders and in forensic applications.

4.0 CONCLUSION

Molecular biology seeks to understand the molecular basis of life. Relating the structure of specific molecules of biological importance – such as proteins, enzymes and nucleic acids – to their functional roles in cells and organisms.

5.0 SUMMARY

Researchers in molecule biology use specific techniques native to molecule biology, but combine these with techniques and ideas from genetics and biochemistry to characterise, isolate and manipulate the molecule components of cells and organisms.

6.0 TUTOR-MARKED ASSIGNMENTS

1. Briefly define molecular biology
2. Outline the relationship of molecular biology to other biological sciences.
3. What is nucleic acid hybridisation.

7.0 REFERENCES/FURTHER READINGS

Robert, K.M.; Daryl, K.G.; Peter, A.M. and Victor, W.R. (2003). *Harper's Illustrated Biochemistry*, Twenty-Sixth Edition. McGraw Hill Publishers India.

Raven, P. H. and Johnson, G. B. (1985). *Biology.* Second Edition. Times Mirror/Mosby College Publishing.

Campbell, N. A. (1996). *Biology.* Fourth Edition. The Benjamin/Cummings Publishing Company Inc. California.

Bruce, A., Dennis, B., Lewis,J., Martin, R., Keith, R. and James, D.W. (1994). *Molecular Biology of the Cell.* Third Edition. Garland Publishing , Inc. New York

MICROBIAL GENETICS

CONTENTS

1.0 INTRODUCTION

It studies the genetics of micro-organisms, which involves the study of the genotype of microbial species and also the expression system in the form of phenotypes. It also involves the study of genetic processes taking place in these micro organisms.

2.0 OBJECTIVES

At the end of this unit you should be able to:

1. Recognise the importance of *E. coli* as an experimental model.
2. Recognise yeasts as models for studies of eukaryotic cells.

3.0 MAIN BODY

3.1 *Escherichia coli as Experimental Model*

Because of their comparative simplicity, prokaryotic cells (bacteria) are ideal models for studying many fundamental aspects of biochemistry and molecular biology. The most thoroughly studied species of bacteria is *E. coli* which is the most favoured organism for investigation of the basic mechanisms of molecular genetics. Most of the present concepts of molecular biology – DNA replication, the genetic code, gene expression and protein synthesis – are derived from studies of *E. coli*.

E. coli is useful to molecular biologists because of its relative simplicity and the ease of its propagation and study in the laboratory. For example the genome of *E. coli* consists of approximately 4.6 million base pairs and encodes about 4000 different proteins. While the human genome is more complex with approximately 3 billion base pairs and encodes about 100,000

different proteins. The small size of *E.coli* genome provides advantages for genetic analysis and the sequence of the entire *E.coli* genome has been determined.

Molecular genetic experiments are further facilitated by the rapid growth of *E.coli* under well defined laboratory conditions. *E. coli* can divide every 20-60 minutes, depending on culture conditions and a clonal population of *E.coli* all cells derived by division of a single cell of origin – can be isolated as a colony grown on agar – containing medium. Bacterial colonies contain many cells, and selecting and analysing genetic variants of *E.coli* strain is easy and rapid. This generally contributes to the success of experiments in molecular genetics.

E.coli can divide rapidly in nutrient mixtures like glucose, salts, amino acids, vitamins and nucleic acid precursors. However, *E. coli* can also grow in much simpler media consisting of only salts as source of nitrogen (such as ammonia) and a source of carbon and energy (such as glucose). But in such simple medium, the bacteria grow a little slowly (a division time of about 40 minutes) because they must synthesise all their own amino acids, nucleotides and other organic compounds.

The ability of *E. coli* to carry out these biosynthetic reactions in simple defined media has made them extremely useful in elucidating the biochemical pathways involved. Thus, the rapid growth and simple nutritional requirements of *E. coli* have greatly facilitated fundamental experiments in both molecular biology and biochemistry.

Although bacteria are models for studies of cell properties, they cannot be used to study aspects of cell structure and function that are unique to eukaryotes. Yeasts, the simplest eukaryotes,

have a number of experimental advantages similar to those of *E. coli* and have provided a model for studies of many aspects of eukaryotic cell biology.

The genome of the most studied yeasts, *Saccharomyces cervisae*, consists of 12 million base pairs of DNA and contains about 6000 genes; and is about 3 times larger than that of *E. coli* it is much more manageable than the genomes of more complex eukaryotes, such as humans. Yeasts can be readily grown in the laboratory and can be studied by many of the same molecular genetic approaches that have proved successful with *E. coli*. Although yeasts do not replicate as rapidly as bacteria, they still divide as frequently as every 2 hours and can easily be grown as colonies from a single cell. Yeasts can be used for a variety of genetic manipulations similar to those that can be performed using bacteria. Yeast mutants have been important in understanding many fundamental processes in eukaryotes, including DNA replication, transcription, RNA processing, protein sorting and regulation of cell division.

4.0 CONCLUSION

The evolution of present day cells from a common ancestor has important implications for cell and molecular biology as an experimental science. Because the fundamental properties of all cells have been conserved during evolution, the basic principles learned from experiments performed with one type of cell are generally applicable to other cells.

5.0 SUMMARY

Because of their genetic simplicity and ease of study, bacteria such as *E. coli* are particularly useful for investigation of fundamental aspects of biochemistry and molecular biology. Yeasts, as the simplest eukaryotic cells, yeasts are an important model for studying various aspects of eukaryotic cell biology.

6.0 TUTOR-MARKED ASSIGNMENTS

1. Give two reasons why *E. coli* is useful in molecular biology.
2. Yeasts have been used as models for the study of many aspects of the biology of eukaryotic cells. Why are they not suitable for analysis of animal cell movements?

8.0 REFERENCES/FURTHER READINGS

Robert, K.M.; Daryl, K.G.; Peter, A.M. and Victor, W.R. (2003). *Harper's Illustrated Biochemistry*, Twenty-Sixth Edition. McGraw Hill Publishers India.

Raven, P. H. and Johnson, G. B. (1985). *Biology.* Second Edition. Times Mirror/Mosby College Publishing.

Campbell, N. A. (1996). *Biology.* Fourth Edition. The Benjamin/Cummings Publishing Company Inc. California.

Bruce, A., Dennis, B., Lewis,J., Martin, R., Keith, R. and James, D.W. (1994). *Molecular Biology of the Cell.* Third Edition. Garland Publishing , Inc. New York

GENES AND CHROMOSOMES

CONTENTS

1.0 INTRODUCTION

Order implies information, instructions are required to arrange parts or processes in a non-random way. Biological instructions are encoded in the DNA which is a substance of genes; the units of inheritance that transmit information from parents to offspring to ensure continuity of life.

2.0 OBJECTIVES

At the end of this unit you should be able to:

1. Define the gene.
2. Define chromosomes.
3. Describe the structure of chromosomes.
4. Recognise the importance of genes and chromosomes in heredity.

3.0 MAIN BODY

3.1 Genes

The classical principles of genetics were deduced by Gregor Mendel in 1865 on the basis of the results of breeding experiments with peas. Mendel studied the inheritance of a number of well-defined traits such as plant height and was able to deduce general rules for their transmission. In all cases, Mendel could correctly interpret the observed patterns of inheritance by assuming that each trait is determined by a pair of inherited “factors”, which are now called genes.

The term gene (*Gr genos* descent) was coined by W. Johannsen in 1909 to refer to the hereditary “factors” of Mendel. The gene can be defined as a functional unit of heredity which occupies a specific place (locus) on a chromosome, is capable of

reproducing itself exactly at each cell division and directs the formation of an enzyme or other protein.

Cytologic and genetic studies show that genes are the fundamental units of inheritance regarded as indivisible units of the chromosomes on which they are located like “beads on a string”.

Genes as the chief functional genetic unit, determine the basic architecture of every cell, the nature and life of the cell, the specific protein synthesis, the enzyme formation and the self-reproduction of the cell.

Genes are molecular patterns that can maintain their identities for many generations, can be self-duplicated in each generation and can control cell processes by allowing their specificities to be occupied. Genes can mutate, can be assorted, can be shuffled in different combinations, therefore, genes are regarded as the basis for modern interpretation of evolution.

On operational or functional basis as a cistron, a muton or a recon as determined or encountered by different approaches.

3.1.1 A Cistron: A gene can be referred to as a unit of function called a cistron which is the smallest functional region on a chromosome. This idea replaced the unitary concept of the physical gene with the concept of an operational gene composed of one or more functional components (mutant alleles) that embraces an array of mutant sites, that behaves in a Mendelian fashion. The term cistron was coined by Seymour Benzer and is regarded as a portion of DNA specifying a single polypeptide chain. Each cistron is responsible for coding one messenger RNA molecule, which in turn participate in the formation of a polypeptide chain. It

has been discovered that hundreds of units of mutation (mutons) and recombination (recons) exist within each cistron. Cistrons, therefore, occupy a much greater chromosomal length than mutons or recons.

3.1.2 Muton: There are many positions or sites within a cistron where mutations can occur. A muton is the smallest genetic unit or length of DNA that can mutate i.e. a change in muton could result in mutation to produce a phenotypic effect. A muton may consist of a single nucleotide or many nucleotides.

3.1.3 Recon: Sometimes crossing over or recombination occur in a cistron to provide another sub-divisional concept of the cistron, the recon. The recon is the smallest genetic unit that can undergo crossing over (exchange of genetic material), or recombination i.e. the smallest unit within the DNA capable of being independently involved in recombination.

3.2 CHROMOSOMES

Thomas Morgan after looking at Hugo De Vries theories of mutations, bred the fruit fly *Drosophila melanogaster* for a year with no mutations, but a fly appeared with white eye instead of red eyes, within another year. 40 different kinds of mutations had been noticed. In trying to explain how organisms inherit characteristics Morgan discovered that the various mutations of the flies were associated with 4 pairs of chromosome possessed by *Drosophila*.

Thomas Morgan's work proved the chromosomal theory of inheritance which states that chromosome are the elements that

transmits inheritable characteristics, also discovered that chromosome are the carriers of genes which cause the expression of individual characteristics.

Chromosomes are filamentous rod-like or thread-like gene bearing bodies found in the nucleus during cell division. Each nucleus contain information coded in the form of DNA and organised into groups called genes. Genes are arranged on the chromosome and each gene contain enough information for the production of one protein which can have some effects on the individual chromosome vary widely between different organisms. The chromosome molecule may be circular or linear, typically eukaryotic cells (cells with nuclei) have large linear chromosomes and prokaryotic cells (cells without defined nuclei) have circular chromosomes. Chromosomes are the essential unit for cellular division and must be replicated, divided and passed successfully to their daughter cells so as to ensure the genetic diversity and survival of their progeny.

3.1.2 Chromosome Structure

Chromosome within a cell occur in matched pairs called homologous xomes, joined at the centre by a centromse. Each chromosome contains many genes, and each gene is located at a particular site on the chromosome, known as the locus. Like chromosome, genes typically occur in pairs. A gene found on one chromosome in a pair usually has the same locus as another gene in the other chromosome of the pair, and these two genes are called alleles. Alleles are alternate forms of the same gene.

In organisms that use sexual reproduction, offspring inherit one-half of their genes from each parent and then mix the two sets

of genes together. This produces new combinations of genes; so that each individual is unique but still possesses the same genes as its parents.

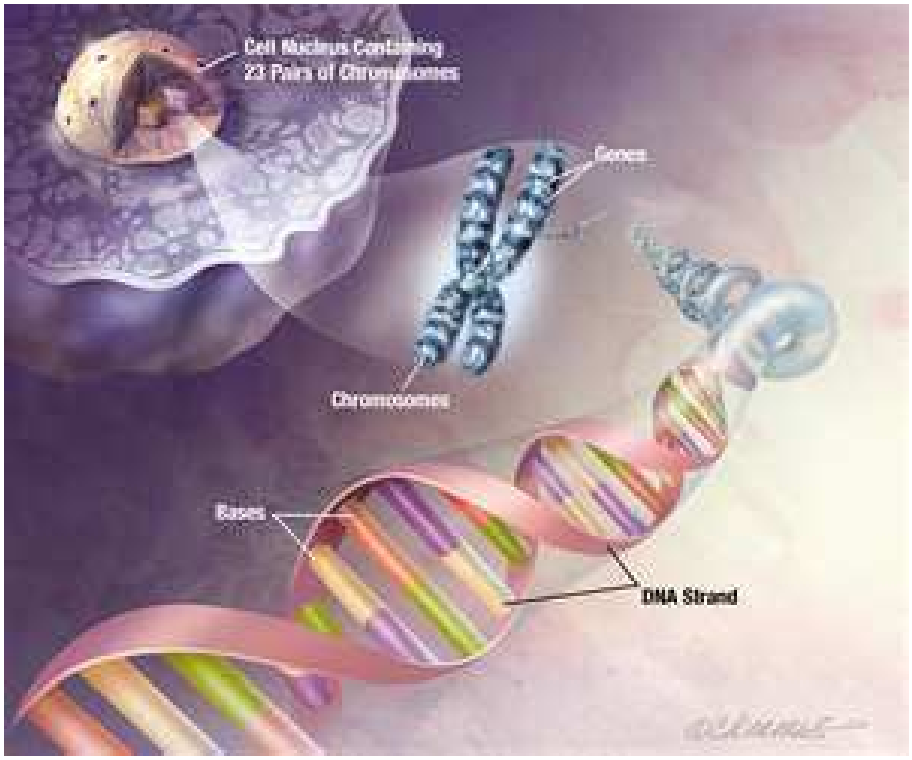


Figure 1. Chromosome structure

3.2.1 Chromosome Number

In the cells of most organisms that produce sexually, chromosomes occur in pairs; one chromosome is inherited from the female parent; and one is inherited from the male parent. The two chromosomes of each pair contain genes that correspond to the same inherited characteristics. Each pair of chromosomes is different from every other pair of chromosome in the same cell. The number of chromosome pairs in an organism varies depending on the species. The number of chromosome characteristic of a particular organism is known as the diploid number. Dogs, for example, have 39 pairs of chromosomes and a diploid number of 78 while tomato plants have 12 pairs of chromosomes and a diploid number of 24.

Gametes or sex cells (eggs and sperm) contain only half the number of chromosomes found in the other cells of an organism. This reduced number of chromosomes in the gametes is known as the haploid number. During fertilisation the gametes unite to form a cell known as a zygote containing the diploid number of chromosomes characteristics of the species.

3.2.3 Sex Chromosomes

Most organisms have complete sets of matching chromosomal pairs, known as autosomes. In mammals, birds and some other organisms, one pair of chromosomes is not identical. Known as the sex chromosomes, this pair plays a dominant role in determining the sex of an organism. Females have two copies of the X chromosome while males have one Y chromosome and one X chromosome. Both males and females inherit one sex chromosome from the mother (always an X chromosome) and one sex chromosome from the father (an X chromosome in female offspring and a Y chromosome in male offspring). The presence of the Y chromosome determines that a zygote will develop into a male.

3.2.4 Human Chromosomes and Genetic Disorders

Humans have 23 pairs of chromosomes, with a diploid number of 46 numbered according to their size. The largest is chromosome 1 and the smallest is chromosome 23. Physical and chemical meiosis (formation of gametes) can damage chromosomes or alter their number in a cell, to give rise to embryos with more of less genetic material, sometimes resulting in developmental disabilities or health problems. In a process called

non disjunction, paired members of chromosomes fail to separate from one another during meiosis. Non disjunction can lead to a condition known as Down Syndrome, in which a person inherits three copies of chromosome 21. Another condition that may result from non disjunction is Turner Syndrome, a disorder in which a female inherits only a single X chromosome.

Breakage of a chromosome can lead to four types of changes in chromosome structure. A deletion occurs when a chromosome fragment lacking a centromere is lost during cell division. In some cases the fragment may join to the homologous chromosome to produce a duplication. It may reattach to the original chromosome but in a reverse orientation, producing an inversion; the fragment can join a non homologous chromosome a rearrangement called translocation.

4.0 CONCLUSION

Every chromosome in a cell contains many genes and each gene is located at a particular site or locus, on the chromosome. Chromosomes vary in size and shape and usually occur in matched pairs called homologoues. The number of homologous chromosomes in a cell depends upon the organism.

5.0 SUMMARY

Genes are present in all living cells. They are copied in their entirety every time a cell divides so that each new cell gets a complete set. They determine virtually all traits that living organisms possess. Chromosome within a cell occurs in matched pairs. Each chromosome contains many genes. Like

chromosomes, genes typically occur in pairs, the two genes are called alleles. Alleles are alternate forms of the same gene.

6.0 TUTOR MARKED ASSIGNMENTS

1. What is a gene's locus?
2. Mendel assumed that each trait is determined by?
3. Briefly define (a) Gene (b) Cistron (c) Recon.
4. Briefly state the chromosomal theory of inheritance.
5. What are chromosomes?
6. What is chromosomal non disjunction?
7. Chromosomes within a cell occur in matched pair called ?

7.0 REFERENCES/FURTHER READINGS

Geoffrey, M. C. (2000). *The Cell. A Molecular Approach*, Second Edition. ASM Press Washington D.C.

Nancy, L. P.; Larry, S. U. and William, S. (2003). *Bioinquiry: making Connections in Biology*. John Wiley & Sons Inc.

NUCLEIC ACIDS

CONTENTS

1.0 INTRODUCTION

The principal genetic informational materials or molecules of living organisms are chemically called nucleic acids. Both DNA and RNA are polymers of complex molecules containing carbon, oxygen, hydrogen, nitrogen and phosphorus. The molecules of nucleic acids (polymers) are composed of monomers called nucleotides joined by covalent bonds.

2.0 OBJECTIVES

At the end of this unit you should be able to:

1. Describe the chemical composition of nucleic acids.
2. Identify the types of nucleic acids.
3. Describe the structure of DNA and RNA.
4. Highlight the roles of DNA and RNA.
5. Differentiate DNA from RNA.

3.0 MAIN BODY

3.1 Nucleic Acids

Nucleic acids are extremely complex molecules produced by living cells and viruses, to pass on hereditary characteristics from one generation to the next, and to trigger the manufacture of specific proteins. The name nucleic acids comes from their initial isolation from the nucleic of living cells. However, certain nucleic acids are found not in the cell nucleus but in cell cytoplasm.

Nucleic acid molecules are very large chains of repeating nucleotide units linked in many sequences. Thus, nucleic acids are high polymers with very high molecular weights. A nucleotide is a molecular unit or a nucleic acid molecule that consists of 3 subunits:

- a phosphate group.
- a pentose sugar (ribose or deoxyribose).
- a nitrogen base (purine or pyrimidine).

A nucleoside is a compound consisting of 2 subunits – a pentose sugar and a nitrogen base. It is a precursor of a nucleotide. A summary of nucleic acid formation is:

- Pentose sugar + Nitrogen base = Nucleoside
- Nucleoside + Phosphate group = Nucleotide
- Nucleotide + Nucleotide + Nucleotide = Nucleic acid.

3.2 TYPE OF NUCLEIC ACIDS

There are two types of nucleic acids – DNA (Dexyribonucleic acid) and RNA (Ribonucleic acid); both are chemical relatives that are universally present in all living cells and they form the chemical basis of life. Both DNA and RNA contain the purines – Adenine (A) and Guanine (G) and the pyrimidine cytosine (C). The second kind of pyrimidine in DNA is Thymine (T) where as it is Uracil (U) in RNA. Therefore a unique pyrimidine distinguishes DNA from RNA.

3.3 DNA

The DNA is a polymer made up of repeating units of mononucleotides carrying the genetic material of all cellular organisms and most viruses. DNA carries the information needed to direct protein synthesis ad replication. Protein synthesis is the production of the proteins needed by the cell or virus for its activities and development. Replication is the process of which DNA copies itself for each descendant cell or virus, passing on the information needed for protein synthesis. In most cellular

organisms, DNA is organised on chromosomes located in the nucleus of the cell.

3.4 DNA STRUCTURE

The DNA is a spiral ladder with the nucleotides forming the side pieces and the steps composed of a combination of purine and pyrimidine which join the deoxyribose sugars in the side pieces to hold them together. The purines are of two types Adenine (A) and Guanine (G), the pyrimidines too are of two types Cytosine (C) and Thymine (T). In forming the steps of the ladder G may join C, or C may join G, and T may join A, or A may join T, usually by hydrogen bond. With the pairings G–C, C–G, A–T or T–A occurring throughout the length of the DNA molecule. The combination of the sugar molecule, phosphate group and a nitrogenous base completes the basic structure of a nucleotide. With purine linked to a pyrimidine precisely adenine (A) always pairing with thymine (T) and guanine (G) pairing with cytosine (C). A mirror image of the nucleotide is added to produce a double nucleotide chain which will twist to produce the α – helix.

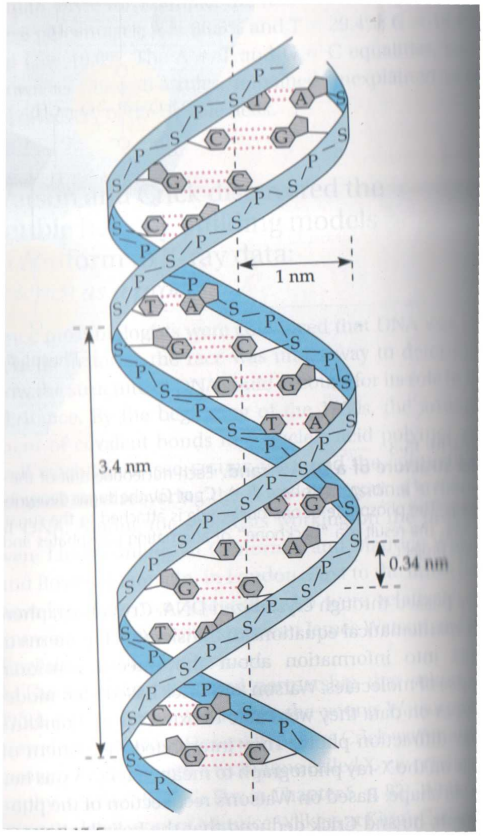


Figure 2. DNA Structure

3.5 RNA

Ribonucleic acid (RNA) - In cellular organisms is the molecule that directs the middle steps of protein production and the genetic material of certain viruses. In cellular organisms the DNA, carries the information that determines protein structure. But DNA cannot act alone and relies upon RNA to transfer this crucial information (translate) during protein synthesis – production of the proteins needed by the cell for its activities and development. In RNA viruses, the RNA directs two processes – protein synthesis (production of the virus's protein coat) and replication (the process by which RNA copies itself).

3.6 RNA STRUCTURE

The structure of RNA is similar to that of DNA and it is composed of a single string of ribonucleotides, each of which is composed of

- a pentose sugar (ribose sugar)
- a phosphate group
- a nitrogenous base (one of the two bases – adenine, guanine, uracil and cytosine)

These components are joined together in the same manner as in DNA molecule. But RNA differs chemically from DNA by being single stranded, having a D-ribose sugar instead of Deoxyribose sugar and having uracil as nitrogenous base instead of thymine. These nitrogenous bases can occur in any sequence.

3.7 TYPES OF RNA

There are three types of RNA classified based on their molecular size. The smallest type of RNA is called transfer – RNA (tRNA) which carries amino acids to the ribosomes for incorporation into a protein. Each amino acid has different classes of tRNA that read the codes of mRNA, therefore involved in protein synthesis. The tRNA receives information from mRNA, through pairing of their bases and accordingly selects particular amino acids and pass to the ribosome.

The second type of RNA is the ribosomal – RNA (rRNA), this is larger than tRNA and composes the ribosomes in the cytoplasm, the specialised structures that are the sites of protein synthesis. Transfer – RNA are the most abundant type of RNA and they coordinate the sequential coupling of tRNA molecules to the series of mRNA codons.

The largest type of RNA is the messenger – RNA (mRNA). Messenger – RNA is a strand of RNA that is complementary to the DNA sequence for a gene and carries the genetic blueprint copied from the sequence of bases in a cell's DNA. This blue print specifies the sequence of amino acids in a protein. All the types of RNA are formed as needed, using specific sections of the cell's DNA as template.

4.0 CONCLUSION

Two classes of nucleic acids are the deoxyribonucleic acids (DNA) and ribonucleic acids (RNA). All living cells contain the genetic material DNA, that determines the shape, the form and the function of the offspring. While the RNA takes part in the actual synthesis of the proteins a cell produces, the structure of which is specified by the DNA.

5.0 SUMMARY

The DNA occurs almost exclusively in the chromosomes and to a small extent in the mitochondria and chloroplasts. RNA occurs mostly in the cytoplasm, nucleolus, ribosomes and to some extent in the chromosomes. DNA is the sole genetic material that migrates intact from generation to generation; through the reproductive units (gametes). The DNA is responsible for the development of specific characters in the successive generations. It is also the controlling agent of all the vital activities of the cell and is responsible for all biosynthetic processes including protein synthesis. Therefore the DNA holds and controls all the secrets of life of the cell. The RNA is under the instructions of DNA and acts

as a messenger carrying information from the DNA to the ribosomes for synthesis of proteins.

6.0 TUTOR-MARKED ASSIGNMENTS

1. What are the 3 major components of a nucleotide?
2. What are the major differences between DNA and RNA?
3. Briefly outline the major functions of the 3 types of RNA.

7.0 REFERENCES/FURTHER READINGS

Geoffrey, M. C. (2000). *The Cell. A Molecular Approach*, Second Edition. ASM Press Washington D.C.

Nancy, L. P.; Larry, S. U. and William, S. (2003). *Bioinquiry: making Connections in Biology*. John Wiley & Sons Inc.

Verma, P. S. and Agarwal, V. K. (2008). *Cell Biology, Genetics, Molecular Biology, Evolution and Ecology.* S. Chand and Company Limited, New Delhi.

DNA REPLICATION

CONTENTS

1.0 INTRODUCTION

DNA is the genetic material that makes up the genes; it contains all the information needed for the cell's growth, operation, and division into two similar cells. During replication, an exact copy of the DNA is made; with the existing DNA being used as a template for the synthesis of new DNA strands in the cell nucleus.

2.0 OBJECTIVES

At the end of this unit you should be able to:

1. Show how the Watson-Crick Model accounts for replication.
2. Know the semi-conservative nature of DNA replication.

3.0 MAIN BODY

DNA as the sole genetic material of living organisms must be able to replicate itself exactly if information is to be transferred from parents to offspring and from generation to generation. In most cellular organisms, replication of a DNA molecule takes place in the cell nucleus and occurs just before the cell divides.

During replication the parent double helix DNA molecule uncoils when the hydrogen bonds between the nitrogenous bases are broken, and as a result the double helix DNA begins to unzip and unwind. The unzipping creates two separate parent strands of DNA. Each parent strand becomes the template (pattern) for the creation of a daughter strand.

The unzipping exposes chemical bonds on the purines and pyrimidines, the nucleoplasm is a reservoir of free nucleotides from which each A on the parent strand attracts a T nucleotide, each C attracts a G nucleotide and so on. When the nucleotides are lined

up they join together to form a polynucleolide chain, the DNA polymerase helps the nucleotides link up; by bonding the phosphate group of nucleotide to the sugar molecule of the adjacent nucleotide during the side rail of the new DNA molecule.

After each daughter strand bonds to the parent strands, the molecules twist again into a double helix, forming two identical strands. Each new DNA molecule retains half of the original DNA material and pairing of bases occur; this type of replication is semi-conservative and complementary.

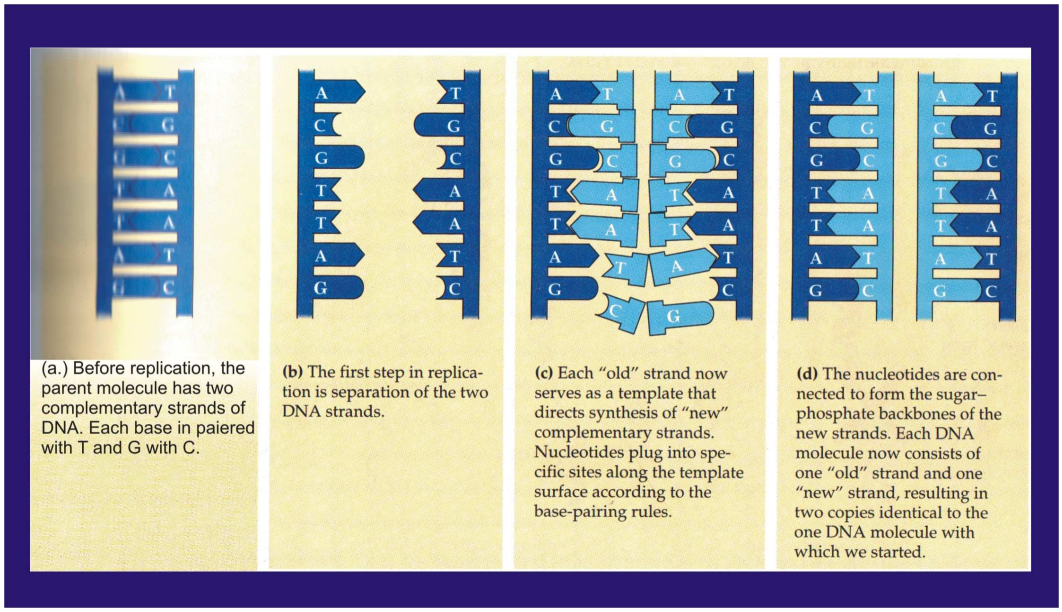


Figure 3. DNA Replication

4.0 CONCLUSION

The double helix models of DNA molecule of Watson and Crick embodies a built-in template system for self-replication. Because of the specificity of base pairing, the sequence of base along one chain automatically determines the base sequence along the other. Thus, each chain of the double helix can serve as a template for the synthesis of the other.

5.0 SUMMARY

The Watson-Crick Model of DNA explains how genetic replication occurs. The process produces two complete double-chained molecules, each identical in base sequence to the original double-chained molecule.

6.0 TUTOR-MARKED ASSIGNMENT

1. Define DNA replication.
2. What is a template?

8.0 REFERENCES/FURTHER READINGS

Robert, K.M.; Daryl, K.G.; Peter, A.M. and Victor, W.R. (2003). *Harper's Illustrated Biochemistry*, Twenty-Sixth Edition. McGraw Hill Publishers India.

Raven, P. H. and Johnson, G. B. (1985). *Biology.* Second Edition. Times Mirror/Mosby College Publishing.

Campbell, N. A. (1996). *Biology.* Fourth Edition. The Benjamin/Cummings Publishing Company Inc. California.

Bruce, A., Dennis, B., Lewis,J., Martin, R., Keith, R. and James, D.W. (1994). *Molecular Biology of the Cell.* Third Edition. Garland Publishing , Inc. New York

DNA TRANSCRIPTION

CONTENTS

1.0 INTRODUCTION

Genes are the instructions for making specific proteins. But a gene does not build a protein directly. The bridge between genetic information and protein synthesis is ribonucleic acid (RNA). The synthesis (creation) of RNA from a DNA template occurs during transcription.

2.0 OBJECTIVES

At the end of this unit you should be able to:

1. Explain the process of transcription.
2. Recognise the importance of transcription in protein synthesis.

3.0 MAIN BODY

3.1 Transcription

The DNA molecule represents information required for the building of a phenotype structure i.e. the DNA is a molecule carrying a message that when passed in some way to a ‘manufacturing site’ and translated controls the formation of the individual structural units necessary to the complete individual.

The DNA does not play any part in the manufacturing process but merely give instructions as to what shall be processed and how this should be done. The DNA carries its message in some form of code, this being represented by the sequence of the bases in the poly-nucleotide chain. Each message or sequence being equivalent to the gene.

This DNA with its genetic code is located in the nucleus of cells and within chromatin and chromosomes. While the protein synthesising machinery (ribosomes) of the cell is located in the

cytoplasm, so the information must therefore be transported from the nucleus to the cytoplasm. This usually happens was the code along the DNA molecule is copied by a strand of mRNA. The copying of codon sequences from DNA to mRNA is called transcription. Transcription results in an RNA complement that includes uracil (U) in all instances where thymine (T) would have occurred in a DNA complement.

If one codon on the DNA molecule is AAA the complementary condon on a strand of mRNA would be UUU, TAT would transcribe as AUA. Then the mRNA with a faithful reverse copy of the genetic code, separates from the DNA template. The mRNA then passes through minute pores in the nuclear membrane and into the cytoplasm.

Unlike replication, transcription does not progress along the entire length of a chromosome. Instead, certain parts of the chromosome are transcribed. The whole process is divided into the following stages: Pre-initiation, initiation, promoter clearance, elongation and termination.

3.2 Pre-Initiation

The first step in transcription is binding of RNA polymerase to a DNA molecule. Binding occurs at particular sites, the promotes, which are specific sequences of 20 to 200 bases at which several interactions occur or regions of DNA which promote transcription. A special promoter region has been identified in eukaryotic organisms. It is a short DNA sequence known as a TATA box, because it is enriched with the nitrogenous bases thymine (T) and adenine (A) found 25-30 base pairs upstream from the short site of transcription. TAT box orient the RNA

polymerase enzyme, so that synthesis proceeds from left to right. It is also the region at which the double helix opens to form the open promoter complex which is the binding site for a transcription factor known as TATA bining protein (TBP) constituting the pre-initiation complex which is a highly stable complex and an active intermediate in chain initiation. It is in this complex a local unwinding or melting of the DNA helix occurs, which is necessary for pairing of the incoming ribonucleotides.

3.3 Initiation

Once an open-promoter complex has been formed, RNA polymerase is ready to initiate RNA synthesis. RNA polymerase contains two nucleotides binding sites, called the initiation site and the elongation.

In eukaryotes, RNA polymerase does not directly recognise the core promoter sequences. Instead, a collection of proteins called transcription which are proteins needed to initiate transcription but are not part of the RNA polymerase mediate the binding of RNA polymerase and the initiation of transcription. Only after certain transcription factors are attached to the promoter does the RNA polymerase bind to it. The completed assembly of transcription factors and RNA polymerase bind to the promoter, forming a transcription initiation complex. Once active RNA polymerase is bound to a promoter region, the enzyme begins to separate the two DNA strands at the initiation site, and transcription is underway.

3.4 Promoter Clearance

After the first bond is synthesised, the RNA polymerase must clear the promoter. During promoter clearance there is tendency for the RNA transcript to be released to produce truncated transcripts in a process known as abortive initiation. Abortive initiation continues to occur resulting in the transcription elongation complex.

3.5 Elongation

One strand of the DNA, the template strand (non coding strand), is used as a template for RNA synthesis. As transcription proceeds, RNA polymerase transverses the template strand from 3′→5′ direction and uses base pairing complementarity with the DNA template to create an RNA copy. Transcription proceeds in the 5′→3′ direction. This produces an RNA molecule from 5′→3′, an exact copy of the coding strand (with thymines replacing uracils and ribose sugar replacing deoxyribose in the sugar phosphate backbone). During transcription multiple RNA polymerase can be involved on a single DNA template and multiple rounds of transcriptions, so many mRNA molecules can be rapidly produced from a single copy of a gene.

Elongation also involves a proof reading mechanism that can replace incorrectly incorporated bases. This may correspond with short pauses during transcription that allow appropriate RNA editing factors to bind.

3.6 Termination of Transcription

Transcription proceeds until the RNA polymerase reaches a termination site on the DNA. The sequence of nitrogenous bases

that marks this site signals RNA polymerase to stop adding nucleotides to the RNA strand and release the RNA molecule.

4.0 CONCLUSION

Transcription is a process of creating an equivalent RNA copy of a sequence of DNA. During transcription a DNA sequence is read by RNA polymerase, which produces a complementary RNA strand.

5.0 SUMMARY

Transcription involves the production of a special kind of RNA known as messenger RNA (mRNA). The process can be summarised in these simple steps. DNA unwinds or unzips as the hydrogen bonds break, the free nucleotides of the RNA pair with complementary DNA bases, RNA sugar-phosphate backbone forms aided by RNA polymerase, the hydrogen bonds of the untwisted RNA and DNA ladder break, freeing the new RNA.

6.0 TUTOR-MARKED ASSIGNMENTS

1. What is transcription?
2. What is transcription unit?
3. What are transcription factors?

8.0 REFERENCES/FURTHER READINGS

Robert, K.M.; Daryl, K.G.; Peter, A.M. and Victor, W.R. (2003). *Harper's Illustrated Biochemistry*, Twenty-Sixth Edition. McGraw Hill Publishers India.

Raven, P. H. and Johnson, G. B. (1985). *Biology.* Second Edition. Times Mirror/Mosby College Publishing.

Campbell, N. A. (1996). *Biology.* Fourth Edition. The Benjamin/Cummings Publishing Company Inc. California.

GENE EXPRESSION

CONTENTS

1.0 INTRODUCTION

In multicellular organism every cell in the body has identical genetic information, individual cells have different structural and functional characteristics. Gene expression is the most fundamental level at which genotype gives rise to the phenotype. The genetic code stored in DNA in form of nucleotide sequence is interpreted by gene expression, and the properties of the expression products give rise to the organism's phenotype.

2.0 OBJECTIVES

At the end of this unit you should be able to:

1. Explain how gene expression controls the process of development in multicellular organisms.

3.0 MAIN BODY

3.1 Control of Gene Expression

Gene expression is a process by which genes coded information is converted into the structures operating in a cell. The process is used by all known life – eukaryotes, prokaryotes and viruses – to generate the macromolecular machinery for life. Gene expression gives the cell control over structure and function and is the basis for cellular differentiation, morphogenesis and the versatility and adaptability of any organism.

3.2 Gene Expression in Bacteria

Bacterial cells are genetically simpler than eukaryotic cells, with just one chromosome and only about 3,000 genes. Bacteria can be grown rapidly in large numbers under controlled conditions in the laboratory, they have been especially useful for studying the

regulation of gene expression. Francois Jacob and Jacques Monod in 1960, formulated a powerful model of the control of gene expression in bacterial cells, based on their investigation of enzyme synthesis in *E. coli*. The Jacob-Monod model proposes that three parts of the chromosome are involved in controlling transcription of the structural genes. The regulator gene, the operator region and the promoter region.

The regulator gene, controls indirectly the activity of the structural genes. The regulator gene is located near the structural genes and encodes the information for the synthesis of a repressor protein. When the repressor binds to the operator, it blocks the promoter's binding sites for RNA polymerase and thus prevents transcription of the structural genes. In this system, the genes specifying particular enzymes are inactive until turned on by an inducer substance. In a negative control system such as the lac operon, a repressor protein binds to the operator and turns off transcription. When the repressor protein is inactive, the operator is turned on and transcription and translation automatically occur. In a positive control system, proteins called transcription factors bind to the promoter and activate transcription.

3.3 Hormonal Control of Gene Expression

In higher plants and animals signals in various glands and/ or secretory cells somehow stimulate target tissue or target cells to undergo dramatic changes in their metabolic patterns. These changes frequently include altered pattern of differentiation that are generally dependent on altered patterns of gene expression. Peptide hormones such as insulin and steroid hormones such as estrogen, progestrone, testosterone (in animals like mammals) and

ecdysone (in insects). In higher animals, hormones are synthesized in specialized secretory cells called endocrine cells and are released into the blood stream. The peptide hormones do not normally enter cells because of their relative large size. Their effects are mediated by receptor proteins located in target-cell membranes and by the intracellular levels of secondary messenger called cyclic AMP (cAMP). The cAMP activates a protein kinase which activates many specific enzymes. The steroid hormones on the other hand, are small molecules that readily enter cells through the plasma membrane. Once inside the appropriate target cells, the steroid hormones become attached to specific receptor proteins which are present only in the cytoplasm of target cells. The hormone-receptor protein complexes activate the transcription of specific and correct genes by binding to specific DNA sequences present in the *cis*-acting regulatory regions of the genes.

4.0 Conclusion

In genetics, gene expression is the most fundamental level at which genotype gives rise to the phenotype. The genetic code stored in the DNA in form of nucleotide sequence is interpreted by gene expression , and the properties of the expression products give rise to the organism's phenotype.

5.0 Summary

Gene expression is the process by which coded information in from a gene is used in the synthesis of functional gene products, which are often proteins. The process of gene expression is used

by all known life- eukaryotes, prokaryotes and viruses – to generate the macromolecular machinery for life.

6.0 Tutor-Marked Assignments

1. Write short notes on (i) gene expression and (ii) steroid hormones.

7.0 REFERENCES/FURTHER READINGS

Robert, K.M.; Daryl, K.G.; Peter, A.M. and Victor, W.R. (2003). *Harper's Illustrated Biochemistry*, Twenty-Sixth Edition. McGraw Hill Publishers India.

http://en.wikipedia.org/wiki/geneexpression.

Raven, P. H. and Johnson, G. B. (1985). *Biology.* Second Edition. Times Mirror/Mosby College Publishing.

Campbell, N. A. (1996). *Biology.* Fourth Edition. The Benjamin/Cummings Publishing Company Inc. California.

GENETIC CODE

CONTENTS

1.0 INTRODUCTION

The essential question of gene expression is how does the order of nucleotides in a DNA molecule encode the information that specifies the order of amino acids in a protein, i.e. the correspondence between nucleotide triplets and amino acids in proteins. The letters A,G,T, and C correspond to the nucleotides found in DNA. They are organised into three-letter code words called codons, and the collection of these codons makes up the genetic code.

2.0 OBJECTIVES

At the end of this unit you should be able to:

1. Know the nature of the genetic code.
2. Outline the characteristics of the genetic code.
3. Know the condon, anti-condon and nonse condon.

3.0 MAIN BODY

3.1 The Nature of the Genetic Code

The genetic code is the set of rules by which information encoded in genetic material (DNA or mRNA sequences) is translated into proteins (amino acids) by living cells.

In 1961 Francis Crick and his colleagues reasoned that the genetic must likely consist of a series of blocks of information, each block corresponding to an amino acid in the encoded protein. They further hypothesised that the information in one block was probably a sequence of three nucleotides specifying a particular amino acid, they arrived at the number three because a two nucleotide block will not yield enough different combinations to

code for the 20 different kinds of amino acids that commonly occur in proteins.

Within genes that encode proteins the nucleotide sequences of DNA is usually in increment of three (3) consecutive nucleotide without penetration between the increment. Each block of three (3) nucleotide code of one amino acid. These 3 nucleotide blocks are called codons. Translation occurs on the ribosome, first the initial portion of mRNA transcribed in a gene binds to an rNRA molecule interwoven in the ribosome, the mRNA lies on the ribosomes in such a way that only the 3 nucleotide portion of the mRNA molecule – the codon is exposed at the polypeptide making site as each bit of the mRNA message is exposed in turn. A molecule of tRNA in the complementary 3 nucleotide sequences or anticodon binds to the mRNA, because the tRNA molecule carries a particular amino acid, that amino acid and no other is added to the polypeptide chain in that position.

Protein synthesis occurs as a series of tRNA molecules bond one after another to the exposed portion of mRNA molecule as it moves through the ribosomes, each of this tRNA molecule has attached to it an amino aid and the amino acid it brings to the ribosome is added one after another to the end of a growing polypeptide chain. The anticodon of a tRNA is 3 nucleotide long; the base sequences of the tRNA anticodons are complementary to the associated sequences of mRNA. Since there are 4 different kinds of nucleotides in mRNA (C, G, A, U) there are 4^3 or 64 different 3 letter code words or codons possible. The list of different mRNA codons specific for each of the 20 amino acids is called the Genetic code. The genetic code is the same in all organisms with only a few exceptions. A particular codon such as

AGA corresponds to the same amino acid (Arginine) in bacteria as in humans.

Note: That 3 out of the 64 codons – UAA, UAG and UGA do not correspond to triplets that are recognised by any activating enzyme. These 3 codons, called nonsense codons they serve as chain terminators or as stop signals in the mRNA message, marking the end of a polypeptide i.e. they specify where the polymerisation of amino acids into a protein is to stop.

The codon AUG both codes for methionine and serves as an initiation site. The first AUG in an mRNA's coding region is where translation into protein begins, marking the beginning of a polypeptide amino acid sequence. The ribosome uses the first AUG that it encounters in the mRNA message to signal the start of its translation.

Table 1 RNA codon table

Nonpolar polar basic acidic (stop codon)

1st base	2nd base U		2nd base C		2nd base A		2nd base G	
U	UUU	(Phe/F) Phenylalanine	UCU	(Ser/S) Serine	UAU	(Tyr/Y) Tyrosine	UGU	(Cys/C) Cysteine
	UUC	(Phe/F) Phenylalanine	UCC	(Ser/S) Serine	UAC	(Tyr/Y) Tyrosine	UGC	(Cys/C) Cysteine
	UUA	(Leu/L) Leucine	UCA	(Ser/S) Serine	UAA	Ochre (*Stop*)	UGA	Opal (*Stop*)
	UUG	(Leu/L) Leucine	UCG	(Ser/S) Serine	UAG	Amber (*Stop*)	UGG	(Trp/W) Tryptophan
C	CUU	(Leu/L) Leucine	CCU	(Pro/P) Proline	CAU	(His/H) Histidine	CGU	(Arg/R) Arginine
	CUC	(Leu/L) Leucine	CCC	(Pro/P) Proline	CAC	(His/H) Histidine	CGC	(Arg/R) Arginine
	CUA	(Leu/L) Leucine	CCA	(Pro/P) Proline	CAA	(Gln/Q) Glutamine	CGA	(Arg/R) Arginine
	CUG	(Leu/L) Leucine	CCG	(Pro/P) Proline	CAG	(Gln/Q) Glutamine	CGG	(Arg/R) Arginine
A	AUU	(Ile/I) Isoleucine	ACU	(Thr/T) Threonine	AAU	(Asn/N) Asparagine	AGU	(Ser/S) Serine
	AUC	(Ile/I) Isoleucine	ACC	(Thr/T) Threonine	AAC	(Asn/N) Asparagine	AGC	(Ser/S) Serine
	AUA	(Ile/I) Isoleucine	ACA	(Thr/T) Threonine	AAA	(Lys/K) Lysine	AGA	(Arg/R) Arginine
	AUG[A]	(Met/M) Methionine	ACG	(Thr/T) Threonine	AAG	(Lys/K) Lysine	AGG	(Arg/R) Arginine
G	GUU	(Val/V) Valine	GCU	(Ala/A) Alanine	GAU	(Asp/D) Aspartic acid	GGU	(Gly/G) Glycine
	GUC	(Val/V) Valine	GCC	(Ala/A) Alanine	GAC	(Asp/D) Aspartic acid	GGC	(Gly/G) Glycine
	GUA	(Val/V) Valine	GCA	(Ala/A) Alanine	GAA	(Glu/E) Glutamic acid	GGA	(Gly/G) Glycine
	GUG	(Val/V) Valine	GCG	(Ala/A) Alanine	GAG	(Glu/E) Glutamic acid	GGG	(Gly/G) Glycine

[A] The codon AUG both codes for methionine and serves as an initiation site: the first AUG in an mRNA's coding region is where translation into protein begins. (Nakamoto, 2009).

3.2 Characteristics of the Genetic Code

1. The code is a triplet codon.
2. The code is non-overlapping i.e. in translating mRNA molecule the codons do not overlap but are sequentially arranged.

3. The code is commaless i.e. no punctuation and once the reading is commenced at a specific codon, there is no punctuation between codons, and the message is read in a continuing sequence of nucleotide triplets until a translation stop codon is reached.
4. The genetic code is unambiguous, with a particular codon always coding for the same amino acid.
5. The code is universal ranging from bacteria to man.
6. Some codes act as start codons (AUG).
7. Some act as stop codons (UAA, UAG, UGA).
8. The code has polarity i.e. it is always read in a fixed direction the 5′ 3′ direction.
9. Degenerate: The code is degenerate i.e. more than one condon may specify the same amino acid.

4.0 CONCLUSION

The (genome) full complement of genetic information that an organism inherits from its parents is inscribed in DNA. The portion of the genome that codes for a protein or RNA is referred to as a gene. Those genes that code for proteins are composed of tri-nucleotide units called cocoons each coding to a single amino acid. The genetic code represents the order of the nucleotide sequences in DNA or RNA that form the basis of heredity through their role in protein synthesis.

5.0 SUMMARY

Genetic instructions from DNA are written in three nucleotide units called codons. There are 64 codons in the genetic

code, 61 of the codons code for amino acids and 3 of the 64 cocons function as stop signals.

6.0 TUTOR-MARKED ASSIGNMENTS

1. What do you understand by ‘genetic code’?
2. Outline any four characteristics of the genetic code.
3. Briefly define the following: codon, anticodon and nonsense codon.

7.0 REFERENCES/FURTHER READINGS

Robert, K.M.; Daryl, K.G.; Peter, A.M. and Victor, W.R. (2003). *Harper’s Illustrated Biochemistry*, Twenty-Sixth Edition. McGraw Hill Publishers India.

http://en.wikipedia.org/wiki/geneticcode.

Raven, P. H. and Johnson, G. B. (1985). *Biology.* Second Edition. Times Mirror/Mosby College Publishing.

Campbell, N. A. (1996). *Biology.* Fourth Edition. The Benjamin/Cummings Publishing Company Inc. California.

Nakamoto T (2009). "Evolution and the universality of the mechanism of initiation of protein synthesis". *Gene* **432** (1-2): 1–6. doi:10.1016/j.gene.2008.11.001. PMID 19056476.

Pamela K. Mulligan; King, Robert C.; Stansfield, William D. (2006). *A dictionary of genetics*. Oxford [Oxfordshire]: Oxford University Press. pp. 608. ISBN 0-19-530761-5.

PROTEIN SYNTHESIS

CONTENTS

1.0 INTRODUCTION

DNA with its correct mechanism of replication, serves to carry genetic information from cell to cell and from generation to generation. The information is translated into proteins that determine the phenotype. Protein synthesis involves how the information present in the sequences of bases (triplet codons) of the mRNA is translated into a sequence of amino acids in proteins.

2.0 OBJECTIVES

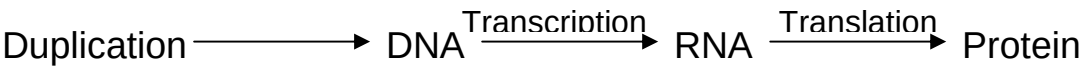
At the end of this unit you should be able to:

1. Describe the mechanism of protein synthesis.
2. Describe the central dogma.
3. Identify the roles of transcription and translation in the flow of genetic information.

3.0 MAIN BODY

3.1 The Central Dogma

Genes are the instructions for making specific proteins. But a gene does not build a protein directly. The bridge between genetic information and protein synthesis is RNA. The process of synthesis of protein involves one of the central dogma of molecular biology; which postulates that genetic information flows from nucleic acids to protein. The first step of the central dogma is known as transcription and does not involve a change of code since DNA and mRNA are complementary. The second step involves a change of code from nucleotide sequences to amino acid sequences and is called translation illustrated as follows:



3.2 Mechanism of Protein Synthesis

Protein synthesis is a very complex biochemical transformation performed by cells resulting in the formation of a polypeptide chain. The mechanism of protein synthesis can be divided into the following 3 main steps: chain initiation, chain elongation and chain termination. All three steps require protein factors (about 200 different proteins) mostly enzymes that aid in mRNA, tRNA and ribosomes in the translation process. Chain initiation and elongation require energy usually provided by GTP (guanosine triphosphate), a molecule closely related to ATP.

3.2.1 Chain Initiation

The initiation stage brings together mRNA a tRNA bearing the first amino acid of the polypeptide chain, and the two subunits of a ribosome. First a small ribosomal subunit binds to both mRNA and a special initiator tRNA. The small ribosomal subunit attaches to the end of the mRNA (loading site). Down stream from the loading site is the initiation codon, AUG, where translation actually begins. The initiator tRNA which carries the amino acid methionine, attaches to the initiation codon. The union of mRNA, imitator tRNA and small ribosomal subunit is followed by the attachment of a large ribosomal subunit to form a functional ribosome. Proteins called imitation factors are required to bring these components together. The cell also spends energy in the form of one GTP to form the initiator complex. At the completion of the initiation process, the initiator tRNA sits in the P site of the ribosome, and the vacant A site is ready for the next tRNA.

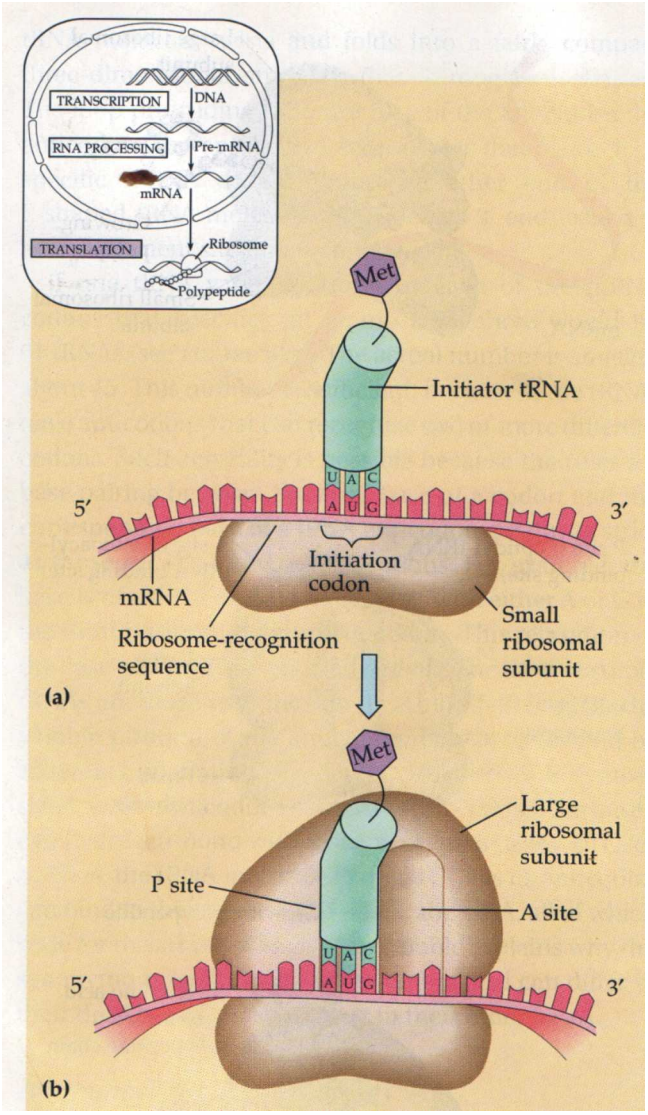


Figure 4. Chain Initiation

3.2.2 Chain Elongation

The elongation stage amino acids are added one by one to the initial amino acid joined by peptide bond. Each addition involves the participation of several proteins called elongation factors. The whole process occurs in a thee-step cycle.

3.2.2.1 Codon Recognition

The mRNA codon in the A site of the ribosome forms hydrogen bonds with the anticodon of an incoming molecule of tRNA carrying its approrpite amino acid. An elongation factor ushers the tRNA into the A site. This step requires the hydrolysis of a phosphate bond from GTP.

3.2.2.2 Peptide Bond Formation

A component of the large ribosomal subunit catolyzes the formation of a peptide bond between the polypeptide extending from the P site and the newly arrived amino acid in the A site. In this step, the polypeptide separates from the tRNA to which it was bound and is transferred to the amino acid carried by the tRNA in the A site.

3.2.2.3 Translation

The tRNA in the P site dissociates from the ribosome. The tRNA in the A site, now attached to the growing polypeptide, is translocated to the P site. As the tRNA changes sites, its anticodon remains hydrogen-bonded to the mRNA codon, allowing the mRNA and tRNA molecules to move as a unit. This movement brings the next codon to be translated into the A site. The

translocation step requires energy, which is provided by hydrolysis of a GTP molecule.

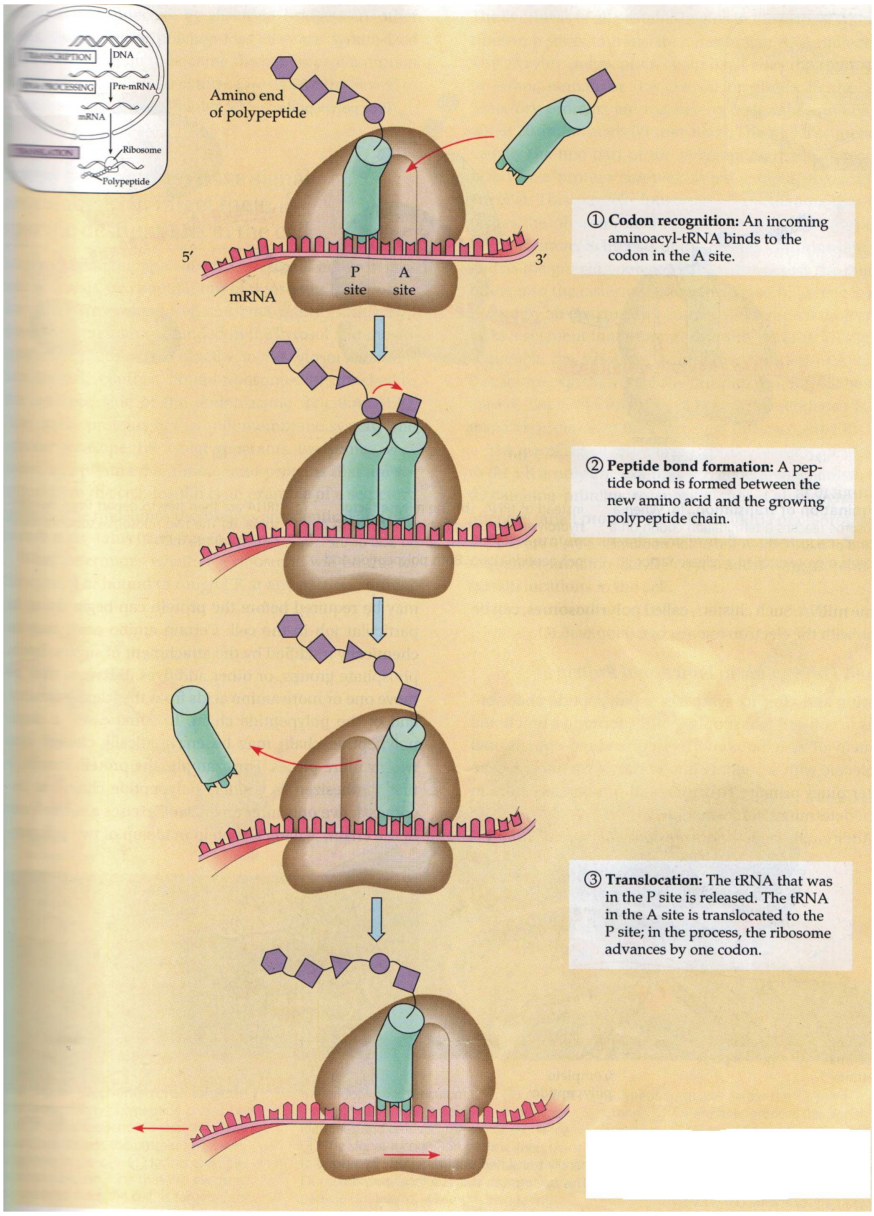


Figure 5. The elongation cycle translation

3.2.3 Chain Termination

The final stage of translation is termination. Elongation continues until a termination codon reaches the A site of the ribosome. Nonsense base triplets – UAA, UAG and UGA – do not code for amino acids but instead act as signals to stop translation.

A protein called a release factor binds directly to the termination codon in the A site. The release factor causes the ribosome to add a water molecule instead of an amino acid to the polypeptide chain. This reaction hydrolyses the completed polypeptide from the tRNA that is in the P site, thereby freeing the polypeptide from the ribosome. The ribosome then separates into its small and large subunits.

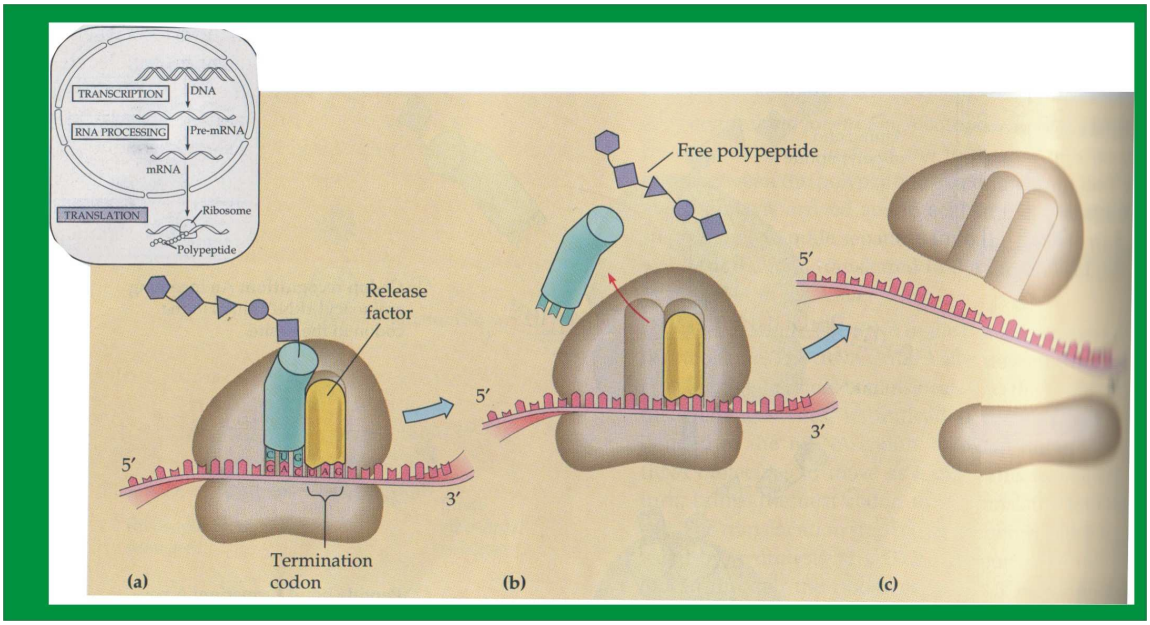


Figure 6. Termination of translation

3.3 From Polypeptide to Functional Protein

During and after its synthesis, a polypeptide chain begins to coil and fold spontaneously, forming a functional protein of specific conformation: a three-dimensional molecule with secondary and tertiary structures. A gene determines primary structure and primary structure in turn determines conformation.

4.0 CONCLUSION

Amino acids are incorporated into proteins in a unique order, specified by a gene. The order of nucleotides in a gene specifies

the amino acid sequence of a protein through translation in which messenger RNA acts as a template for protein synthesis.

5.0 SUMMARY

Cells are governed by a molecular chain of commands that flows from DNA $\rightarrow$ RNA $\rightarrow$ protein. DNA transcribes RNA of the copying of the information contained in DNA by the RNA. The RNA the translates the information and assemble protein on the ribosomes.

6.0 TUTOR-MARKED ASSIGNMENTS

1. Write short notes on (i) Central Dogman (ii) Transcription (iii) Translation.
2. What is initiation complex?
3. Which process in protein require hydrolysis of GTP.

7.0 REFERENCES/FURTHER READINGS

Raven, P. H. and Johnson, G. B. (1985). *Biology.* Second Edition. Times Mirror/Mosby College Publishing.

Campbell, N. A. (1996). *Biology.* Fourth Edition. The Benjamin/Cummings Publishing Company Inc. California.

Robert, K.M.; Daryl, K.G.; Peter, A.M. and Victor, W.R. (2003). *Harper's Illustrated Biochemistry*, Twenty-Sixth Edition. McGraw Hill Publishers India.

DNA SEQUENCING

CONTENTS

1.0 INTRODUCTION

Once an interesting piece of DNA has been isolated or identified, there is need to determine if the sequence of nucleotides in the fragment is related to known genes and to determine what kind of protein it might produce. DNA sequencing refers to sequencing methods for determining the order of the nucleotides bases – adenine, guanine, cytosine and thymine – in a molecule of DNA.

2.0 OBJECTIVES

At the end of this unit you should be able to know:

1. What DNA sequencing is.
2. The methods of DNA sequencing.

3.0 MAIN BODY

3.1 DNA SEQUENCING

DNA sequencing makes it possible to determine the precise order, or sequence of nucleotide bases within a fragment of DNA. In DNA sequencing, many copies of a single-stranded DNA fragment that will be used to synthesise a new DNA strand, are created. Through DNA sequencing a gene can be characterised in terms of a linear sequence of AGCT bases that in turn, can be used to predict the amino acid sequence of the corresponding protein using the genetic code. There are three methods for determining DNA sequences.

3.2 Chain-Termination Method (Sanger Method)

This method requires a single-stranded DNA template, a DNA primer, a DNA polymerase, radioactively labelled nucleotides

and modified nucleotides that terminate DNA strand elongation. The DNA sample is divided into four separate sequencing reactions, containing all four of the stranded deoxynucleotides and the DNA polymerase. To each reaction is added only one of the four dideoxynucleotides which are the chain-terminating nucleotides, that terminate DNA strand extension and result in DNA fragments of varying lengths.

The newly synthesised and labelled DNA fragments are heat denatured, and separated by size, by gel electrophoresis on a denaturing polyacrylamide-urea gel with each of the four reactions run in one of four individual lanes (lanes A, T, G, C) the DNA bands are then visualised by autoradiography or UV light and the DNA sequence can be directly read off the x-ray film.

3.3 Maxam and Gilbert's Chemical (Degradation Method)

The method is based on chemical modification of DNA and subsequent clearage at specific bases; it requires potassium labelling at one end and purification of the RNA fragment to be sequenced. The chemical treatment with **** RNAs generates breaks at every nucleotide base. Thus a series of labelled fragments is generated from the potassium labelled end to the first cut site in each molecule. The fragments in the four reactions are arranged side by side in gel electrophoresis for size separation. The fragments are visualised when the gel is exposed to hydrolysis enzymes for autoradiography, yielding a series of cubes each corresponding to a RNA fragment from which the sequence may be determined.

4.0 CONCLUSION

DNA sequencing allows the isolation of individual fragments of DNA in quantities suitable for detailed characterisation and the determination of nucleotide sequence.

METABOLIC PATHWAYS

CONTENTS

1.0 INTRODUCTION

Metabolic pathways are a series of chemical reactions occurring within a cell. In each pathway or chemical reaction, a principal chemical is modified by a series of chemical reactions. Enzymes catalyse these reactions, and often require dietary minerals, vitamins and other cofactors in order to function properly. Because of the many chemicals (metabolites) that may be involved metabolic pathways can be elaborate. In addition, numerous distinct pathways co-exist within a cell; collectively called the metabolic network. Metabolic pathways are important to the maintenance of homeostatis within an organism.

2.0 OBJECTIVES

At the end of this unit you should be able to:

1. Define metabolism, catabolism and anabolism.
2. Describe the stages involved in the break down of glucose.
3. Outline the processes involved in glycolysis.

3.0 MAIN BODY

3.1 Meaning of Metabolism

Metabolism: The sum of the enzyme-mediated chemical reactions occurring within a cell or a whole organism. These chemical reactions are divided into two – catabolism and anabolism.

Catabolism (catabolic pathway) is a metabolic pathway that releases energy by breaking down complex molecules into simpler compounds.

Anabolism (anabolic pathway) is a metabolic pathway that involves the synthesis or assembling of organic molecules and new protoplasm.

Before the energy stored in lipids, proteins and carbohydrates can be used by the cell to do work, the molecules must be broken down in a series of chemical reactions and the energy produced used to synthesise adenosine triphosphate (ATP) – an adenine – containing nucleoside triphosphate that releases free energy when its phosphate bonds are hydrolyzed. ATP is regarded as the universal energy currency of living organisms. The energy is used to drive endergonic reactions in cells.

The complete degradation of an energy-rich compound such as glucose to carbondioxide and water involves many enzymatically controlled reactions. The complete breakdown of glucose involves four stages:

- Glycolysis (Stage I)
- Fermentation or oxidation of pyruvic acid to acetyl-coenzyme A (Stage II).
- The Kreb's Critic Acid Cycle (Stage III).
- Oxidative phosphorylation (Stage IV).

3.2 Glycolysis

The first series of reactions in the degradation of glucose is termed glycolysis and the most important features of glycolysis are:

- Each molecule of glucose (C_6 H_{12} O_6) is broken down to two molecules of pyruvic acid (C_3 H_4 O_3).

- Two molecules of ATP are used to initiate the process, later four new ATP molecules are synthesised with a net gain of two molecules of ATP and two molecules of NADH (an energy or electron carrier molecule).
- ATP molecules are produced by substrate level phosphorylation and NAD^+ (nicotinamide adenine dinucleotide) is reduced to NADH by oxidation of the food.
- The net energy yield from glycolysis, per glucose molecule is 2 ATP and 2 NADH (Stage I).
- No molecular oxygen is used, glycolysis can occur with (aerobic metabolism) or without oxygen (anerobic metabolism).
- The reactions of glycolysis occur within the cytosol of the cell in the cytoplasm outside the mitochondria. While the krebs cycle enzymes and intermediates are dissolved in the fluid within the mitochondria.

3.3 Fermentation

The fate the pyruvic acid produced during glycolysis depends on oxygen supply. Fermentation enables a cell to continue the reactions of glycolysis in the absence of oxygen.

- In the absence of O_2, the pyruvic acid may be reduced by NADH to CO_2 and ethyl alcohol or lactic acid, in a process called fermentation. In the process the glycolytic pathway leads to the production of alcohol or lactic acid and it enables the cell to continue synthesising ATP molecules by the breakdown of nutrients under anerobic conditions.

- Under aerobic conditions the pyruvic acid can be further oxidised, with the accompanying synthesis of ATP. The process begins when pyruvic acid moves from the cytoplasm into the inner compartment of the mitochondrion to form two each of acetyl-COA, CO_2 and NADH (Stage II).
- The acetyl-CoA formed is fed into a complex circular series of reactions.

3.4 The Krebs Citric Acid Cycle

The cycle functions as a metabolic “furnace” that oxidises organic fuel derived from pyruvate, the product of glycolysis.

- In the course of the cycle, two carbon atoms are lost as CO_2, a molecule of ATP is synthesised, and eight electrons and eight hydrogens are picked up by carrier compounds, forming three molecules of NADH and one of $FADH_2$.
- The cycle generates 1 ATP per turn by substrate phosphorylation.
- Since one molecule of glucose gives rise to two molecules of acetyl-CoA, two turns of the cycle occur for each molecule of glucose oxidised.

3.5 Oxidative Phosphorylation (Electron Transport Chain)

A series of oxidation – reduction reactions that passes electrons from higher energy levels to lower energy levels. As a result, ADP is phosphorylated to form ATP. The exergonic reaction of hydrogen with oxygen to form water releases a large amount of energy in the form of heat and light. In cellular respiration, an electron transport chain breaks the “fall” of electrons in this

reaction into a series of small steps and stores some of the released energy in a form that can be used to make ATP.

The electron transport chain molecules are found in the inner membrane of the mitochondria. The transfer of electrons along the electron-transport chain results in the pumping of H^+ ions from the inner compartment of the mitochondria to the outer compartment. As a result, the H^+ concentration increases in the outer compartment and an electrostatic and osmotic concentration gradient is built up across the inner membrane. Special enzyme complexes, called ATP synthetases act as H^+ ion channels in the inner mitochondrial membrane. As the H^+ ions move down the electrochemical gradient through the complex, energy is released and can be used to synthesise ATP. The total number of new ATP molecules produced by the complete metabolic breakdown of glucose is usually 36, two from glycolysis, two from krebs cycle and 32 from electron-transport phosphorylation in the mitochondria. This yield represent about 39% of the energy of glucose, the rest of the energy is released as heat.

4.0 CONCLUSION

Cellular metabolism is characterised by metabolic pathways or sequences of enzyme-catalysed reactions in which the product of one reaction serves as the reactant of the next. Sometimes pathways are linear, but can be circular, or interconnect with each other, they can even be branched, in which the product of one reaction can go in either directions depending on the needs of the cell at any particular time.

5.0 SUMMARY

A metabolic pathway involves step-by-step modification of an initial molecule to form another product. The resulting product can be used-either immediately, as the end product of a metabolic pathway, to initiate another metabolic pathway or be stored by the cell. Metabolic pathways usually provide precursors for cell components and provide energy for synthetic and other energy requiring processes.

6.0 TUTOR MARKED ASSIGNMENT

1. What are the two major functions of metabolic pathways?
2. What are the four stages involved in the breakdown of glucose?
3. What is the function of the krebs citric acid cycle?

7.0 REFERENCES/FURTHER READINGS

McFadden, C. H. and Keeton, T. W. (1995). *Biology: An Exploration of Life.* W. W. Norton and Company Inc., New York.

Verma, P. S. and Agarwal, V. K. (2008). *Cell Biology, Genetics, Molecular Biology, Evolution and Ecology.* S. Chand and Company Limited, New Delhi.

Hans, G. S. (1995). *General Microbiology.* Cambridge University Press.