

Biophysics : Aspects of Amino Acids Sequence in Proteins and Nucleotide Sequence in Nucleic Acids

Khadka Bahadur Chhetri

Department of Physics, Prithvi Narayan Campus, Pokhara

kckhadak@yahoo.com

Abstract: Protein is the polypeptide chain of amino-acid sequence. Proteins of all species, from bacteria to humans, are made up from the same set of 20 standard amino acids. In order to carry out their function they must take a particular shape which is known as fold. All the enzymes hormones and antibodies are also proteins. To treat certain toxic-microorganism or invader we need certain antigen-antibody complex in the organisms. Just as amino-acid sequence forms the proteins, the polynucleotide sequence forms the nucleic acids. The gene is a part of DNA macromolecule responsible for the synthesis of protein chains. There are 20 amino-acids responsible for the formation of protein and 4 nucleotides responsible for the formation of DNA (RNA). Therefore, we can say that protein text is written in 20-letter and the DNA (RNA) text is written in 4-letter language. The information contained in genes in DNA is transferred to mRNA during transcription.

Keywords: polypeptide chain, globular and fibrous proteins, neurotransmitters, chemical kinematics, misfolded proteins, conformational change, activation energy, entropy, active centers, transcription, enzymatic catalysis, haptens.

1. INTRODUCTION

Proteins are invariable participants in all life processes which control the metabolic and bio-energetic processes. Proteins are involved in the storage, transfer, transformation, recoding and reception of chemical signals in macromolecules, molecules and ions of living systems. DNA replication, transcription of genetic information from DNA to mRNA and transfer of the information encoded in the mRNA are the main functions of protein.

2. PROTEINS

Protein is the polypeptide chain of amino-acid sequence. Proteins of all species, from bacteria to humans, are made up from the same set of 20 standard amino acids. Nineteen of these amino acids are with a primary amino group (-NH₂-) and a carboxylic acid (carboxyl; -COOH-) group attached to a central carbon atom. On the side of the central carbon; a hydrogen atom and a variable side-chain or 'R' group are attached. This 'R' group determines the type of amino-acid. The bonding between amino group (-NH₂-) of one amino-acid and a carboxylic

acid (carboxyl; -COOH-) of another amino-acid is called peptide bonding. The one exception to this general structure is proline, which has a secondary amino group and is really an α -imino acid. The names of the amino acids are often abbreviated to three letters, for example, proline is abbreviated to Pro.

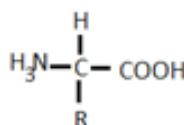


Fig.1 Free amino-acid

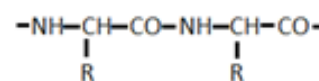


Fig.2 Peptide bonding

From the structural aspects, proteins can be grouped into globular and fibrous proteins. Fibrous proteins are soluble in water because they have high amount of polar molecules on the surface which are hydrophilic in nature. Globular proteins are insoluble in water because they have globular or hydrophilic molecules in excess amount at their surface.

Some fibrous proteins are:

- i. Myosins: Proteins of contractile muscle
- ii. Collagens: found in bones, teeth, skin, connective tissues...etc.
- iii. Fibrin: found in blood clot

- iv. Elastins: found in ligaments, blood vessels, neck of grazing animals...etc.
- v. Keratins: found in nail, hair, wool...etc.

Some globular proteins are:

- i. Albumins: found in blood and white part of egg.
- ii. Globulins: found in blood and other enzymes and hormones responsible for defense mechanism.

On the basis of biological function proteins are classified as below:

- i. Contractile proteins: present in contractile muscles, moving filament, etc. Ex: Myosin, elastins...etc.
- ii. Hormones: Such as growth hormones, insulin and others.
- iii. Membrane proteins: proteins present in mitochondria, protoplasm membrane...etc.
- iv. Enzymes: these are body catalysts which help in mechanochemical and electrochemical bioenergetic phenomena.
- v. Structural proteins: found in hair, nail, feather, skins, etc. Ex: keratins, collagens...etc.
- vi. Storage proteins: these are nutrients such as casein in milk, albumin in egg, spleen (stores iron) and seed proteins...etc.
- vii. Transport proteins: help on transportation and transmission. Ex: hemoglobin, serum albumin etc.
- viii. Neurotransmitters: These transmit neutrons. Ex: endorphins and eukephalines.

3. PROTEIN FOLDING

Proteins are invariably involved in all life processes. In order to carry out their function they must take a particular shape which is known as fold. Hence, before they do their function the proteins assemblage in a particular shape. This self assembling is known as folding. The protein folding represents an energy minimization process and the final structure is that which produces allover minimum energy for the protein. The folding proceeds with one step at a time. Each step is determined only by the favorable kinematics (energy change for stability) for that step. The final structure of the protein

is simply the resultant of the individual steps. The free energy difference between the folded and unfolded state of the protein i.e. transition energy is:

$$\Delta E = RT^2 \frac{d}{dT} \ln \frac{[D]}{[N]};$$

Where [D]= concentration of deformed and denatured proteins and [N]=concentration of native proteins.

Protein folding is basically the process of arrangement of various loops of amino acids in specific patterns. The different arrangement may result in the different types of proteins. Sometimes, the arrangement is in a wrong way which is termed as misfolding. The misfolded proteins are like the poisons that can harm around it. The concept of protein folding is something like “unboiling of an egg”. When an egg is boiled the proteins (albumin) unfold in hot state and when it is cooled the proteins inside it misfold and result a form which is the boiled egg.

Different diseases due to misfolding of the proteins are:

- i. Alzheimer’s disease: It is the conversion of amyloid precursor into soluble protein called Ab which aggregates in long filaments to form b. amyloid. Before Ab aggregation same type of Amyloid Precursor is present in all but why Ab aggregation takes place in some people’s brain but not in other’s is still remaining as a mystery.
- ii. Madcow disease: It is due to the conversion of benign form of protein containing four α -helices into scrapie form containing two α -helices and four β -sheets. Under single amino-acid substitution benign form can convert into scrapie form.
- iii. FAP (Families Amyloidic Polyneuropathic): It is hereditary disease in which peripheral nerves and other organs are damaged due to deposit of amyloidal type.
- iv. Sickle-cell anemia: It is due to single point mutation on the 6th position of β -chain of hemoglobin.
- v. Many cancers are also resulted due to protein misfolding.

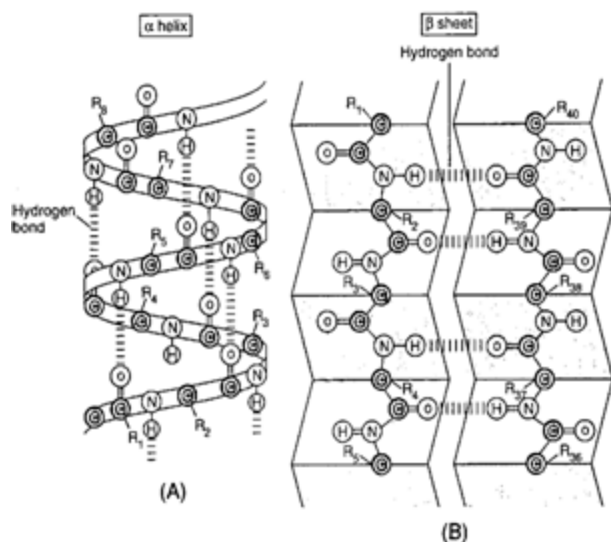


Fig. 3: The α helix and β sheet secondary structures of protein molecules. The R side chains bonded to α -carbons are numbered

The gene is a part of DNA macromolecule responsible for the synthesis of protein chain. The genetic codes are triplets i.e. one coding includes three nucleotides out of four. There are twenty amino acids responsible for formation of protein chains. Among many combinations of genetic codes first half are responsible for synthesis of primary structured proteins but the second half of the genetic code are responsible for folding into polypeptide tertiary structure from its primary structure, which is still a challenging problem. This task of predicting the tertiary structure of protein from its primary amino-acid sequence is taken as protein folding problem. This is taken as problem because the query, why some sequences fold into ordered segments such as α -helix and β -sheets while others fold into turns and loops, and these secondary structural elements fold into final structures are still unsolved issues.

4. ENZYMES

All the enzymes are proteins that act as catalysts for biochemical reactions. Enzymes accelerate the rate of reaction without being changed themselves. Some basic properties of enzymes are:

- i. Enzymes are the high molecular weight proteins.
- ii. Enzymes are globular because they are found in aqueous environment and hence the hydrophobic molecules remain inside and

hydrophilic i.e. polar molecules remain on the surface.

- iii. The affinity between enzyme and substrate is very high.
- iv. The enzymes help to change others but do not change themselves during reaction process.
- v. The enzymes may go conformational change either before or after the reaction but not at the time of reaction.
- vi. The enzymatic catalysis obeys stoichiometry law of interaction strictly i.e. one protein globule interacts with one substrate at one time.

5. CHEMICAL KINETICS ON ENZYMATIC CATALYSIS

Enzymes serve as catalysts in all bio-chemical reactions to accelerate the rate of reaction but their amount is always conserved i.e. they change substrate without changing themselves.

The reaction rate constant is given by the formula:

$K = A \exp \{-E^{++}/RT\}$; where, E^{++} = the energy of activation i.e. the threshold energy just to start the reaction. The exponential relation indicates that reaction rate increases with the decrease of activation energy. The catalysts help to lower the activation energy and increase the reaction rate.

For a reversible reaction: $A \leftrightarrow B$; the reaction rates are given by: $\vec{v} = k_1[A]$ and $\vec{v} = k_{-1}[B]$; where $[A]$ and $[B]$ represent concentrations k_1 and k_{-1} are forward and reverse reaction rate constants.

At equilibrium, and the equilibrium constant is

$$K = \frac{[B]}{[A]} \Big|_{eq} = \frac{k_1}{k_{-1}} \dots \dots \dots (1)$$

Again from, $K = \exp \left\{ \frac{-\Delta G}{RT} \right\}$; we have,

$$\Delta G = \Delta H - T\Delta S = RT \ln K = -RT \ln \frac{[B]}{[A]} \dots \dots \dots (2)$$

Where, $\Delta H = \text{enthalpy change}$ and $\Delta S = \text{entropy change}$. The condition $\Delta G = 0$ refers to the phase transition.

The reaction is possible only if free energy is decreased. The condition is necessary but not sufficient because sometimes E^{++} may be very high and reaction rate may be vanishingly small. In such

a case we have to reduce E⁺⁺ by using catalysts.

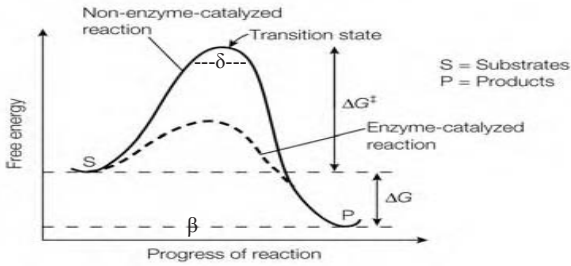


Fig.4: free energy profile for biochemical reaction

Above figure shows a profile of free energy for a certain biochemical reaction. Let ΔG be the free energy difference between S (substrate) to P (product). Let G^{++} be the free energy of activation from $S \rightarrow P$. Let us consider an interval δ at the reaction coordinate β of transition state. Let C' be the concentration of the complex system in the interval δ . Reaction rate is defined as the number of complex systems (coalesce of substrate and enzyme) that pass over energy barrier per unit time.

And, therefore, $\delta f = C' \bar{v}$ (3)

Where, $\bar{v} = \frac{\int_0^\infty v \exp\{-\frac{1}{2}mv^2\} dv}{\int_0^\infty \exp\{-\frac{1}{2}mv^2\} dv} = \sqrt{\frac{K_B T}{2\pi m}}$ (4)

and f = frequency of reaction (products/time)

At the same time, $k = \frac{f}{C_a C_b} = \frac{C'}{C_a C_b} \sqrt{\frac{K_B T}{2\pi m}} \cdot \frac{1}{\delta}$
 $\Rightarrow k = K \sqrt{\frac{K_B T}{2\pi m}} \cdot \frac{1}{\delta}$; Where $K = \frac{C'}{C_a C_b} = \frac{Z_{ac}}{Z_a Z_b}$ (5)

Where, C' = concentration of reaction complex
 C_a, C_b = concentration of reactants a,b
 Z_{ac} = partition function of the activated complex
 Z_a, Z_b = partition function of the reactants a, b
 Again, $Z = \sum_i \exp\{-\frac{E_i}{K_B T}\} = \exp\{-\frac{G}{RT}\}$

For a translational process i.e. with respect to translational degrees of freedom, the partition function is;

$$Z_{tr} = \frac{V}{h^3} (2\pi m K_B T)^{\frac{3}{2}} = \frac{\delta^3}{h^3} (2\pi m K_B T)^{\frac{3}{2}}$$

For the activation complex,

$$Z_{ac} = Z'_{ac} \frac{\delta}{h} (2\pi m K_B T)^{\frac{1}{2}}$$
 (6)

[Since, the activation complex has only one translational d.f. along the direction of reaction.]
 Where, Z'_{ac} = partition function for all d.f. except translational d.f.

From v and v_i

$$k = \frac{Z_{ac}}{Z_a Z_b} \sqrt{\frac{K_B T}{2\pi m}} \cdot \frac{1}{\delta} = \frac{Z'_{ac}}{Z_a Z_b} \sqrt{\frac{K_B T}{2\pi m}} \cdot \frac{1}{\delta} \cdot \frac{\delta}{h} (2\pi m K_B T)^{\frac{1}{2}} = \frac{K_B T}{h} \frac{Z'_{ac}}{Z_a Z_b}$$

$$= \frac{K_B T}{h} \frac{\exp\{-\frac{G}{RT}\}}{(\exp\{-\frac{G_a}{RT}\} \exp\{-\frac{G_b}{RT}\})}$$

$$= \frac{K_B T}{h} \exp\{-\frac{(G' - G_a - G_b)}{RT}\}$$

$$= \frac{K_B T}{h} \exp\{-\frac{G^{++}}{RT}\}; \text{ where, } G^{++} = (G' - G_a - G_b)$$

$$\therefore k = \frac{K_B T}{h} \exp\{-\frac{H^{++}}{RT}\} \cdot \exp\{-\frac{S^{++}}{R}\}$$

But the activation complex which has reached the activation barrier cannot yield if the process is non-adiabatic. Therefore, let us introduce a parameter such that;

$$k = \kappa \frac{K_B T}{h} \exp\{-\frac{H^{++}}{RT}\} \cdot \exp\{-\frac{S^{++}}{R}\}$$

For adiabatic process, $\kappa = 1$

And for non-adiabatic process, $\kappa \ll 1$.

For $\kappa = 1$

$$\ln k = \ln \frac{K_B T}{h} + \left\{ \frac{-H^{++}}{R} \right\} \cdot \frac{1}{T} + \left\{ \frac{S^{++}}{R} \right\}$$

For an adiabatic process $\ln \frac{K_B T}{h}$ slightly depends on T and hence can be neglected.

$$\therefore \ln k = \left\{ \frac{-H^{++}}{R} \right\} \cdot \frac{1}{T} + \left\{ \frac{S^{++}}{R} \right\}$$

The non-linear dependence of k on T^{-1} indicates the complexity of the process i.e. the process becomes cooperative. In such a case the free energy of activation is not constant and depends on number of system that have reacted and on absolute temperature T. On cooperative system, degree of cooperation increases with increase in concentration so reaction rate is slow. But the reaction will proceed if the energy barrier is surmounted.

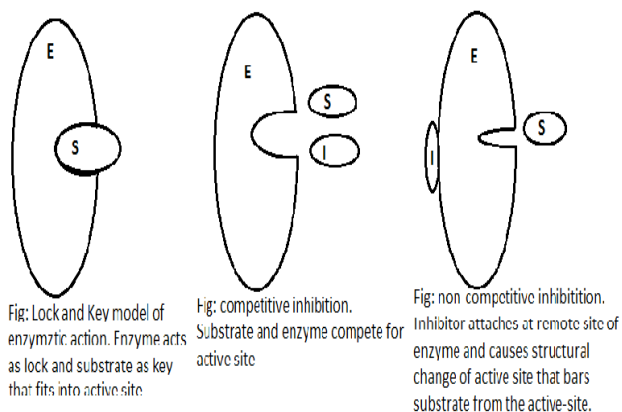


Fig: Lock and Key model of enzymtic action. Enzyme acts as lock and substrate as key that fits into active site

Fig: competitive inhibition. Substrate and enzyme compete for active site

Fig: non competitive inhibition. Inhibitor attaches at remote site of enzyme and causes structural change of active site that bars substrate from the active-site.

Fig.5: Lock and key model of enzymtic action

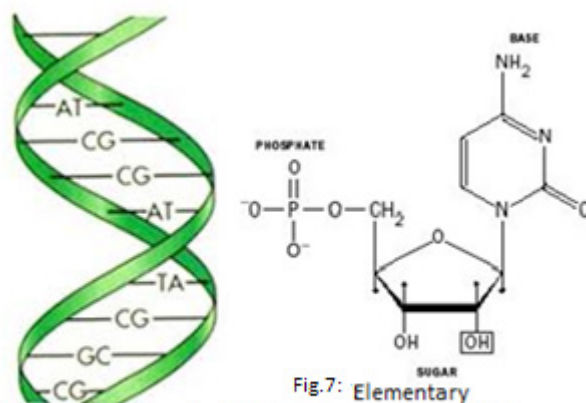


Fig.6: Double Strands of DNA
Fig.7: Elementary Composition of DNA Strand

6. NUCLEIC ACIDS

Just as amino-acid sequence forms he proteins, the polynucleotide sequence forms the nucleic acids. There are two types of nucleic acids.

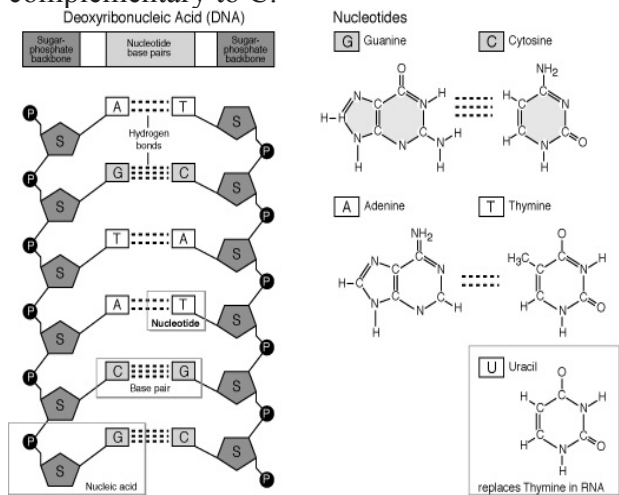
i. Deoxyribonucleic acid; found in chromatin

ii. Ribonucleic acid; found mainly in cytoplasm

Nucleic acid(CHONP)→Pentose sugar+Phosphoric acid + nitrogenous bases.

The nitrogenous bases of DNA are: Adenine (A) always pairing with thymine (T) and vice-versa and guanine(G) always pairing with cytosine(C) and vice-versa.

The nitrogenous bases of RNA are similar as above except the thymine in place of uracil. The two chains of DNA in the double helix are complementary to each other i.e. T is complementary to A and G is complementary to C.



The gene is a part of DNA macromolecule responsible for the synthesis of protein chains. The genetic information is contained in primary structure of DNA. In viruses of the RNA type, the genetic information is coded in RNA molecules.

The correspondence between the nucleotide sequence in DNA and RNA, and the amino-acid sequence in protein chain depends on following aspects:

- i. Correspondence between the information content of DNA and that of protein
- ii. Quantitative relationship between nucleotides and amino-acids
- iii. Physical meaning of genetic code
- iv. Evolutionary origin of genetic code.

There are 20 amino-acids responsible for the formation of protein and 4 nucleotides responsible for the formation of DNA (RNA). Therefore, we can say that protein text is written in 20-letter and the DNA (RNA) text is written in 4-letter language. Hence, from simple considerations the number of nucleotides that code for one amino-acid cannot be less than three. Otherwise, suppose if that was two i.e. <3 , then the number of combination of four nucleotide in group of 2 is $4^2 = 16 < 20$ i.e. less than the number of amino-acids. If the nucleotides are considered in groups of three than there are $4^3 = 64$ different combinations which is very greater than 20. Gammow suggested there may be more than one triplet for each amino-acid and several triplets are without function which reduces the excess combinations. Therefore, genetic code should be

triplet degenerate code.

Gammow further suggested that protein chain is built up on the DNA double helix and each amino-acid being accommodated in the groove between the four nucleotides. This groove has a rhombic form, two nucleotides being to one chain and the other two to the other chain. Each nucleotide of the first chain forms a Watson-Crick pair with a nucleotide on the second chain. Considering right-handed and left-handed forms of rhombus to be identical then there will be only 20 different rhombuses.

Examples:

AAA: codes for Lys = Lysine

UUU: codes for Phe = Phenyl alanine

GGG & GGU: code for Gly = Glycine

CCC & CCU : code for Pro = Proline

UAG, UGA & UAA: specify no amino-acid and are called stops.

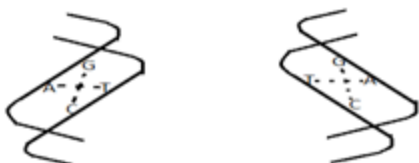


Fig.8: right handed and left handed rhombus are same

The genetic coding is analogous and hence universal for all organisms from virus to humans i.e. for all organisms. The universality of the code is directly proved by the multiplication of bacteria, viruses... etc in the cells of the organisms. For the RNA or DNA virus to be multiplied the virus uses the nucleotides and hence amino-acids of the cells and synthesizes its own proteins. Hence, the genetic coding of the host cells and the virus should be same. This gives proof for the universality of genetic code.

Another proof is obtained in cell free extracts from the cells of frogs and guinea pig which showed the identical results. In cell free extraction the poly-U stimulates to synthesize Glycine and so on in the cells of mammals, algae and other animals.

The potential reading frame (methods of encoding a polynucleotide chain) for any mRNA sequence is different depending on which nucleotide is read first. For instance, let us take a nucleotide sequence: UUUAGGUACCU...

Reading frame-1: UUU AGG UAC CU...
(Phe)

Reading frame-2: U UUA GGU ACC U...
(Gly)

Reading frame-3: UU UAG GUA CCU ...
(Pro)

7. DECIPHERING OF GENETIC CODES

The term deciphering of genetic code means understanding of a code. This can be studied through 'cell-free system' experiments. The term cell-free system refers to the system which does not contain DNA and mRNA, the codon systems but contains ribosomes, tRNAs, energy source ATP and the necessary enzymes.

When synthetic polynucleotide is injected to the cell-free system the protein synthesis occurs even if there are no DNA and mRNA. For example, when synthetic poly-C chain is injected in the cell-free system then it stimulates for the formation of Proline, poly-G stimulates for the formation of Lysine and poly-U stimulates for the formation of Phenyl-alanine.

The cell-free system can be prepared from degraded cells of E. coli. The ribosomes and other solutions of such degraded cells can be separated by centrifugation. To this isolated mixture, an ATP generating system i.e. a source of energy is added. Now, the synthetic polynucleotide was injected into the mixture. Corresponding polypeptide of the synthetic polynucleotide gets synthesized which is studied and can be understood the fact what genetic triplet codes are responsible for the formation on protein biosynthesis.

Nirenberg and Mathieu's experiment found that copolymer of AU synthesizes Phe, Leu, Asn, Ile and Lys. But, which triplet codon gives which amino-acid was not clear. In another experiment Nirenberg replaced polynucleotide with tri-nucleotides of known structure. In this case, trinucleotide-tRNA-aminoacid complexes are formed rather than synthesis of polypeptide. All the 64 triplets were thus studied and the amino-acids to which they became bound were identified. This helped to identify the amino acids corresponding to different triplet codons.

Khorana conducted another method of deciphering genetic codes. He synthesized oligodeoxyribonucleotides which are recurring triplets of known structure. Such oligomer was then used as a template for in-vitro synthesis of DNA-like polymer in a system containing nucleoside tri-phosphate of A,T,C,G; DNA-polymerase and different ions. This yields the synthesis of DNA-like double helix. Then each of chains serves as a template for synthesis of mRNA with the aid of RNA-polymerase enzyme. Thus produced double helix chain mRNA is introduced into a cell-free system and the amino-acids binding to different triplet codons are determined by Nirenberg's method.

For example: Suppose we use $(TTC)_n$ oligomer for synthesis of invitro-DNA. Then there are two helices of complementary triplet codons which in-turn produces mRNA of two helices of complementary triplet codons on inserting both helices in cell-free system different amino-acids form as below:

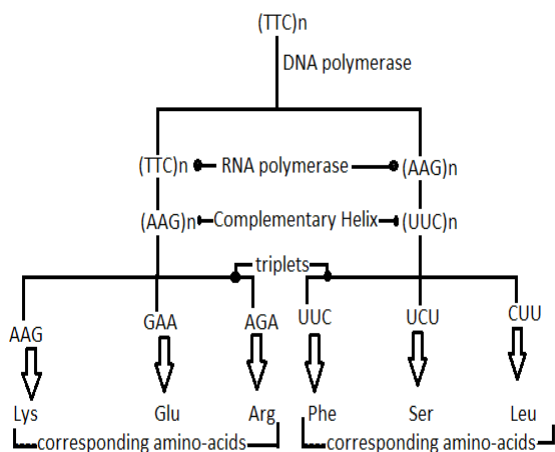


Fig.9: $(TTC)_n$ oligomer and product aminoacids

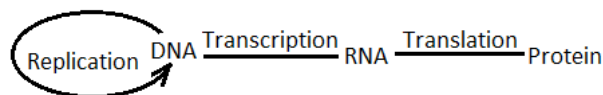
Since, $(CCC)_n$, $(GGG)_n$, $(AAA)_n$ and $(UUU)_n$ have already been easily determined by simple method of Nirenberg and Mathieu, therefore, 10 such different experiments should be carried for determination of amino-acids for remaining 60 codons

8. PROTEIN BIOSYNTHESIS

The existence of fixed primary structure of the protein chain gives evidence that there is an encoded program in the cell for building up of such structure.

The genetic codes are responsible for encoding such programs.

The information contained in genes in DNA is transferred to mRNA during transcription. Protein synthesis takes place exactly on mRNA. The central dogma theory states that the genetic information flows from DNA to RNA to proteins, but genetic information can never be transmitted from protein to nucleic-acids.



The reverse transcription is the transmission of genetic information from RNA to DNA.

i.e. $RNA' \rightarrow DNA \rightarrow RNA \rightarrow Protein$

The messenger RNA (mRNA) transmits the genetic information from chromosomes to the ribosomes on which biosynthesis is accomplished. For biosynthesis to be accomplished amino acids, appropriate thermodynamics and favorable kinematics are required. Ribosomes are nucleoproteins composed of proteins and ribosomal RNA (rRNA). The ribosomes provide appropriate interaction between mRNA and tRNA. tRNA carries amino acids.

The ribosome attaches and moves along the mRNA chain reading text (program) –codon after codon. The mRNA-ribosome complex binds tRNAs linked to amino acids. Each tRNA carries an amino-acid and interacts in complementary way with the codon of mRNA by its anti-codon. On the ribosome successive attachments of two aminoacyl-tRNA complexes to the n th and $(n+1)$ th codons of mRNA occurs. Therefore, a peptide bond is formed between two amino acids of tRNA. Now, the n th tRNA split off and detaches. The $(n+1)$ th tRNA with growing chain of peptide bonds moves by one step. Thus, following this process protein is developed continuously with the formation of other peptide bonds.

Several ribosomes rather than one ribosome move along a single mRNA chain and each of the ribosomes develops its own protein chain growing in it. Thus, a number of identical protein chains are synthesized on a single mRNA chain.

The synthesis of a protein chain in a mRNA-ribosome system (polysome) is called translation i.e. the conversion of nucleotide text to the amino acid text, following the dictations of genetic codes in codon-amino acid dictionary. The protein type formed depends on genetic codes of DNA because the transcribed mRNA and tRNAs carrying amino-acids binds with complementary codon-anticodon bond formation in ribosome, which acts just as enzyme.

In brief: DNA → mRNA → tRNA

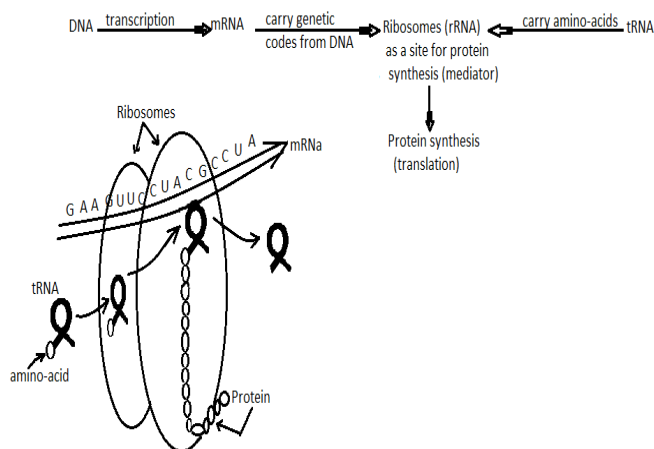


Fig.10: Protein Synthesis

9. DNA REPLICATION

DNA has self replicating property. Chromosomes separate through separation of centromeres during cell division. But, the replication of DNA itself occurs altogether earlier in the cycle of cell growth and division. The DNA replication is necessary in appropriate way so that genetic information present in cells can pass to daughter cell formed by cell division. Mutation can be caused due to crossing on chromosomes of gonad cells.

DNA replication is a process in which cell copies it's DNA before dividing. The DNA is copied by an enzyme called DNA polymerase which acts on single stranded DNA synthesizing a new strand complementary to original strand. There are three modes of DNA replication.

i. Conservative replication:

Original DNA in double helix form remains intact and a new double helix is formed.

ii. Semi-conservative replication:

Each copied DNA contains one strand derived from the parent DNA and one newly synthesized strand.

iii. Dissipative replication:

The original material is distributed among four chains of two daughter double stranded DNA.

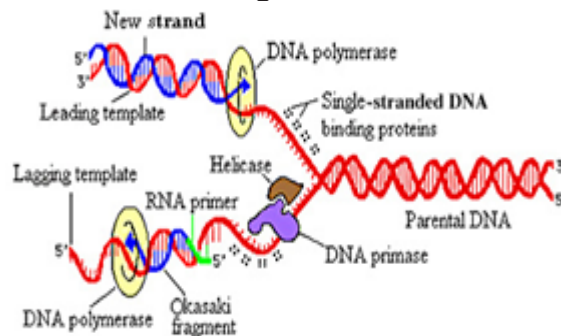


Fig.11: collaboration of proteins at the replication fork

The enzymatic action helps on unwinding of the double helix. Binding of protein helps on the DNA strands to separate templates. Corresponding new DNA strands start to form antiparallely. Extra nucleotides accumulate and align to form new polynucleic acid strand.

10. BODY IMMUNITY SYSTEM

The protection of an organism from foreign biopolymers and thereby from infectious microorganisms is accomplished through cellular and humeral immunity. The word immunity means the state of resistance or insusceptibility exhibited by the host to toxic molecules and foreign invaders.

The cellular immune response is mediated by T. lymphocytes. The T. lymphocytes are so called because their maturation from stem cells takes place in the thymus. The cytotoxic T. lymphocytes (CTL) are responsible for recognition and killing invaders. These cells are therefore called killer T-cells. The other T. lymphocytes provide essential helps for B-lymphocytes and therefore are called helper T. lymphocytes.

The humeral immune response is realized on production of soluble proteins called antibodies or immunoglobulin by B. lymphocytes. A human has five principal classes of immunoglobulin denoted as I_gG, I_gM, I_gA, I_gD and I_gE. The immunoglobulin I_gG functions as an antibody.

The B. lymphocytes are so called because their maturation takes place on bone- marrows. The foreign biopolymers which we call antigens are recognized by certain lymphocytes called recognizing lymphocytes and as a response corresponding antibodies are produced by B. lymphocytes. There are also remembering lymphocytes which remember the antigen and easily recognize the antigen which are already introduced.

The antibodies are capable of combining specifically with the antigens. Antibodies have similar structure to that of antigens except the primary structure in variable or V-region. This V-region is responsible for antigen-antibody interaction. There are two active centers of an antibody each containing one V-region.

Antigens are primarily proteins and polysaccharides. To treat certain toxic-microorganism or invader we need certain antigen-antibody complex in the organisms. So, we inject externally prepared antigen and hence antigen-antibody complex is formed in the organism.

Landsteiner gave a method for synthesis of an artificial antigen. He suggested that the R-phenylazide protein can be used as an antigen for stimulating the generation of respective antibodies. The factor that determines the antigenic specificity is just the radical R and the protein plays secondary role. The R-phenylazide protein; an artificial antigen is the combination of diazotized aromatic compounds with Tyr. Protein.

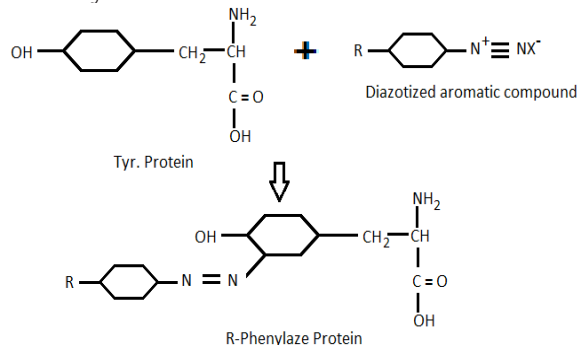


Fig.12: synthesis of artificial antigen

The antibody produced in response to R-P (P stands for protein) reacts with R-P' (P' stands for another protein) but not with R'-P (R' stands for another radical). If we add the R-antigen-R-

antibody system to the sample containing the same group R, the antigen-antibody reaction is retarded and when the concentration is further increased the reaction ceases to proceed. Small molecules do not generate antibodies in the organism are not antigen. They interact, however, with the earlier generated antibodies thereby forming soluble compounds. This is called haptens action and the small molecules mentioned are haptens or determinant groups.

The reaction between antibody and antigen takes place through weak interactions. Natural antigens contain several haptens responsible for forming soluble compounds. The presence of active centers on antibody and the haptens of antigen interact and form antigen-antibody complex.

The presence of the active centers on antigens and antibodies was proved by an experiment carried on rabbit. COOH is used as R in R-phenylazide and subjected to iodine. Then it was seen that the activity was completely suppressed. However, when iodination is carried out in the presence of haptens (determinant groups), the activity was retained.

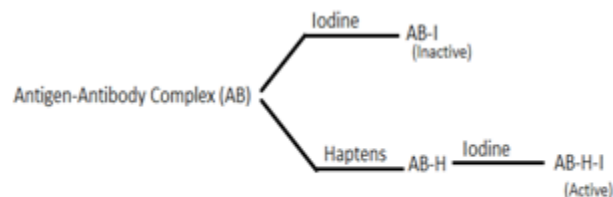


Fig.13: Hapten action

During formation of antigen-antibody complex an equilibrium condition will occur once and the equilibrium constant for this phenomenon is given by: $K = \frac{[AB] - [H]}{[AB][H]}$

Where; [AB] = concentration of antibody
And [H] = concentration of hapten or determinant group of antigen

For theoretical calculation: $K = \exp\left\{\frac{-\Delta H - T\Delta S}{RT}\right\}$

Where; ΔH = change in enthalpy, and ΔS = change in entropy

Since, ΔS should be positive as per second law of thermodynamics this may be accounted for liberation of hydrating water molecules during the formation of complexes.

Antibodies are globular proteins. Their denaturation loses their combination power with antigens. For stability of globule electrostatic salt bridge is formed in a aqueous environment which increases the entropy of the system. The formation of salt bridge leads to the release of oriented water molecules. This aqueous environment stabilizes salt bridge. The haptens present on antigens help on their renaturation.

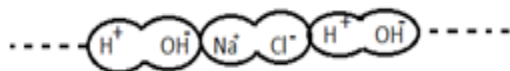


Fig.14: salt bridge in aqueous environment

11. CONCLUSION:

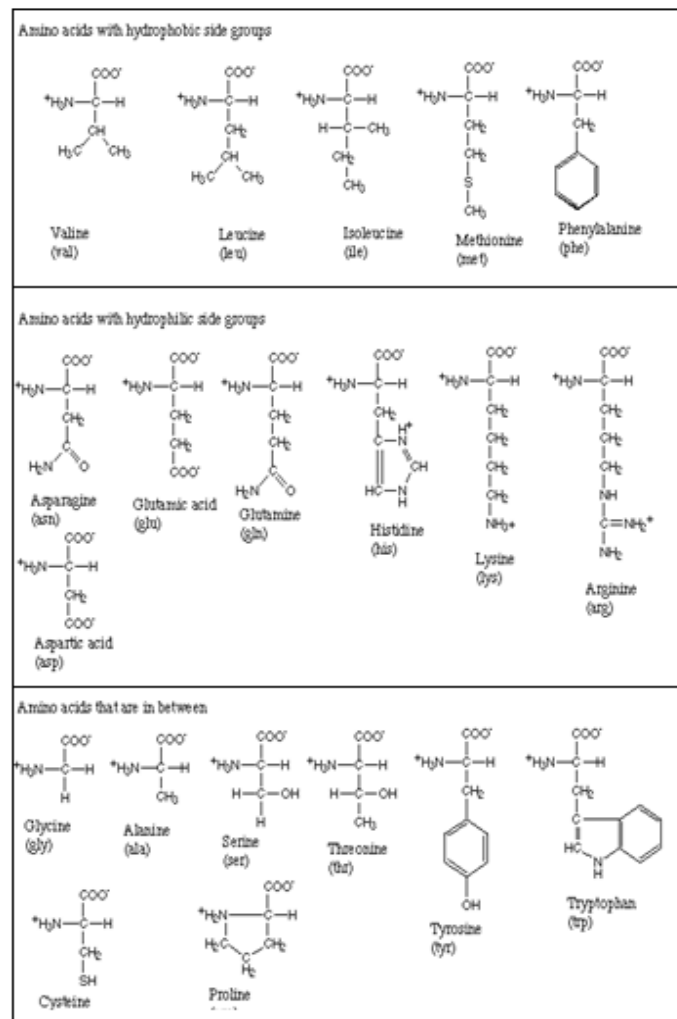
Proteins are indispensable parts in biology. But the conformational kinematics of it and different unsolved mystery behind it are on the responsibility of Physics. Now, physics is not confined on its conventional aspect of theories only but is also centered on different aspects of other sciences like biology. The Biophysics, Bioengineering, Genetics, etc, are applied branch of physics and engineering on biology which are becoming emerging sciences in this era of time.

TABLES

Table-1: Genetic codons and their corresponding amino-acids

		Second Letter							
		U		C		A		G	
1st letter	A	UUU	Phe	UCU	Ser	UAU	Tyr	UGU	Cys
		UUC		UCC		UAC		UGC	
		UUA	Leu	UCA		UAA	Stop	UGA	Stop
		UUG		UCG		UAG	Stop	UGG	Trp
C	CUU	Leu	CCU	Pro	CAU	His	CGU	Arg	
	CUC		CCC		CAC		CGC		
	CUA		CCA		CAA	Gln	CGA		
	CUG		CCG		CAG		CGG		
A	AUU	Ile	ACU	Thr	AAU	Asn	AGU	Ser	
	AUC		ACC		AAC		AGC		
	AUA		ACA		AAA	Lys	AGA	Arg	
	AUG	Met	ACG		AAG		AGG		
G	GUU	Val	GCU	Ala	GAU	Asp	GGU	Gly	
	GUC		GCC		GAC		GGC		
	GUA		GCA		GAA	Glu	GGA		
	GUG		GCG		GAG		GGG		

Table-2: Different 20 amino-acids with their chemical structures



REFERENCES:

- [1] David Hames and Nigel Hooper, Biochemistry, Taylor and Francis Group, New York.
- [2] DM Vasudevan and Sreekumari S, Biochemistry, Jaypee Brothers, New Delhi
- [3] Hands out given by Prof. Dr. S. K. Aryal at C.D.P, T.U.
- [4] H. F. Gilbert, Biochemistry, McGraw-Hill, New York
- [5] Narayanan P., Biophysics, New Age Int. Publisher, New Delhi
- [6] P.K. Banerjee, Biophysics, S. Chand, New Delhi.
- [7] U. Satyanarayana and U. Chakrapani, Biochemistry, Books and Allied (P) Ltd., Kolkatta
- [8] Volkenshtein M.V., Biophysics, Mir Publisher, Moscow.