

A Review on Protein Synthesis and Genetic Code

Janeshwar E Chezhiyan*

Indira Gandhi Medical College and Research Institute, Medical school in Pondicherry, India

*Corresponding author:

Janeshwar E Chezhiyan

✉ janeshwar307@gmail.com;
cinthuriya@gmail.com

Indira Gandhi Medical College and Research Institute, Medical school in Pondicherry, India.

Tel: 6380710020

Citation: Chezhiyan JE (2021) A Review on Protein Synthesis and Genetic Code. Biochem Mol Biol Vol.7 No.1:5

Abstract

In order to function our bodies, we need to supply them with a variety of nutrients. The process of chemical digestion uses different proteins and enzymes to break down the food particles into usable nutrients ourselves can absorb. The instructions to make proteins are contained in our DNA. DNA contains genes. A gene is a continuous string of nucleotides containing a region that codes for an RNA molecule. This region begins with a promoter and ends in a Terminator. Genes also contain regulatory sequences that can be found near the promoter or at a more distant location. For some genes, the encoded RNA is used to synthesize a protein in a process called gene expression for these genes expression can be divided into two processes, transcription and translation in eukaryotic cells. The central dogma of molecular biology describes the two-step process, transcription and translation, by which the information in genes flows into proteins: DNA → RNA → protein. The pathway of protein synthesis is called Translation because the language of nucleotide sequence on mRNA is translated into the language of an amino acid sequence. The process of Translation requires a Genetic code, through which the information contained in nucleic acid sequence is expressed to produce a specific sequence of amino acids.

Keywords: Genetic code; Transcription; Translation

Received: December 04, 2021; **Accepted:** January 25, 2021; **Published:** January 30, 2021

Introduction

Transcription

Transcription occurs in the nucleus where DNA is used as a template to make messenger RNA. Then in translation, which occurs in the cytoplasm of the cell, the information contained in the messenger RNA is used to make a polypeptide during transcription, the DNA and the gene is used as a template to make a messenger RNA strand with the help of the enzyme RNA polymerase [1]. This process occurs in three stages, Initiation, Elongation, and Termination (Figures 1 and 2).

Tata box in prokaryotes: About 35 base pair upstream of the start site 5'-TGTT-GACA-3'; about 10 bp upstream there is another sequence called TATA BOX or pribnow box [5'-TATAAT-3'] [2,3].

Goldbeg – hogness box: In mammals tata box is slightly difference is known as gold beg hogness box. It is located at 25-30 position. Further upstream between 70-80 there is another sequence known as CAAT Box (Figure 3).

Initiation

Bacterial system: DNA helix partial unwinds and the RNAP binds with the promoters site on DNA with help of sigma factors is called pre initiation complex. Next nucleotide attached to the

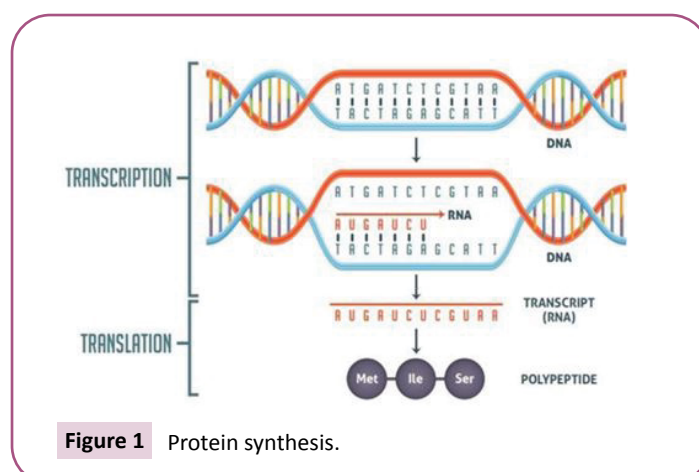


Figure 1 Protein synthesis.

RNAP a phosphodiester bond is formed. The enzymes moves next base, after the ten to twenty nucleotides is polymerized. The RNAP undergoes conformation change and moves away from promoters region this process is called promoters clearances.

Mammalian system: During initiation, the promoter region of the gene functions as a recognition site for RNA polymerase to bind. This is where the majority of gene expression is controlled by either permitting or blocking access to this site. By the RNA polymerase binding causes the DNA double helix to unwind and open.

Elongation

During elongation, the RNA polymerase slides along the template DNA strand, as the complimentary basis pair up the RNA polymerase links nucleotides to the three prime end of the growing RNA molecule (Figure 4).

Termination

Once the RNA polymerase reaches the Terminator portion of the

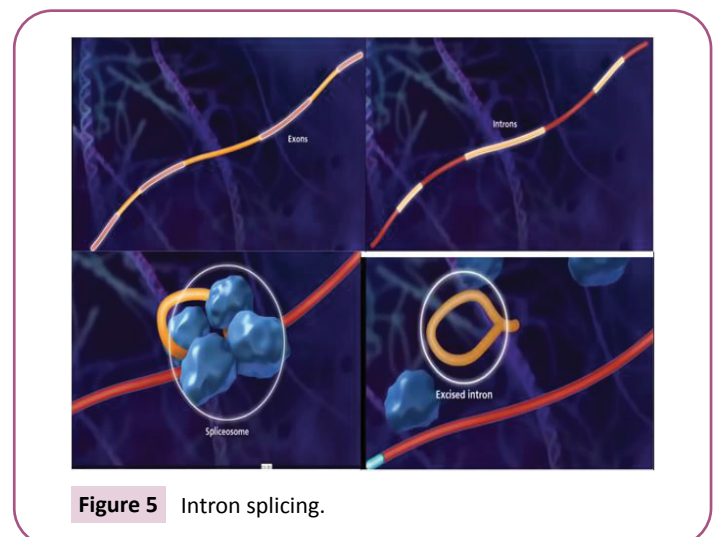
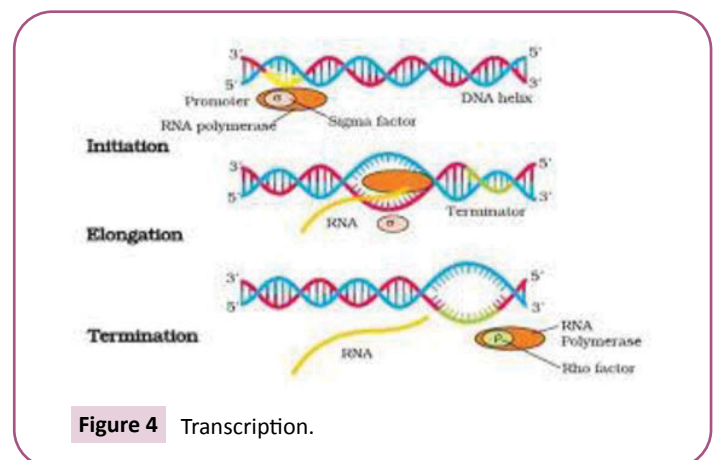
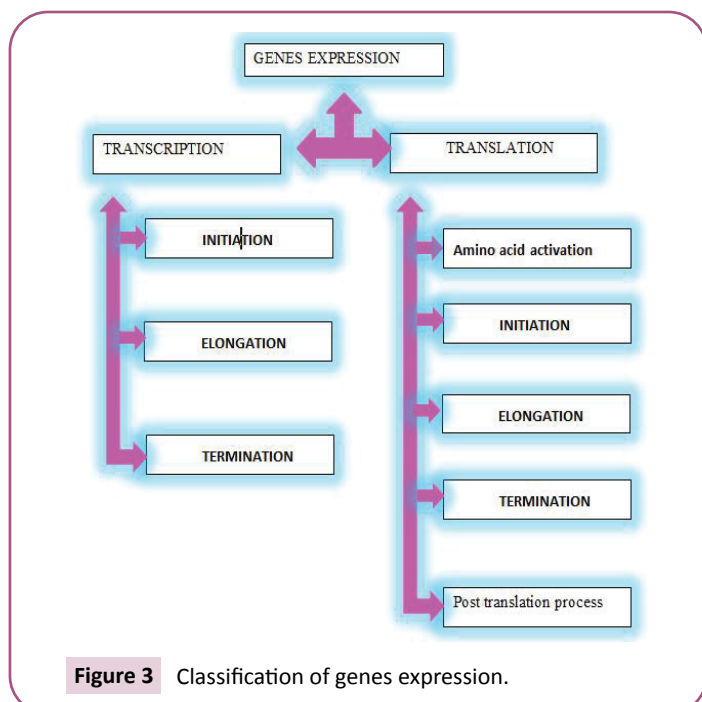
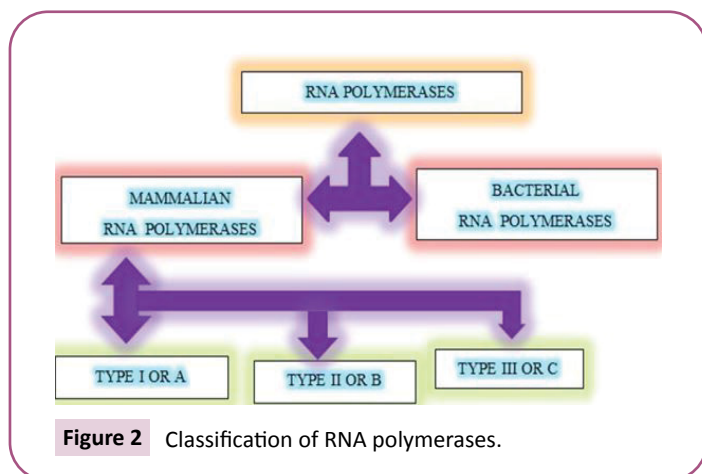
gene, the messenger RNA transcript is complete and the RNA polymerase, the DNA strand and as the messenger RNA transcript dissociate from each other (The specific signals are recognized by termination protein called rho factor). The strand of messenger RNA that is made during transcription includes regions called Exons that code for a protein and non-coding sections called introns In order for the messenger RNA to be used in translation, the noncoding introns need to be removed and modifications such as a five prime cap and a three prime poly a tail are added. This process is called intron splicing and is performed by a complex made up of proteins and RNA called a spliceosome. This complex removes the electron segments and joins the adjacent Exons to produce a mature messenger RNA strand that can leave the nucleus through a nuclear pore and enter the cytoplasm to begin translation (Figure 5).

tRNA or sRNA

tRNA molecules are soluble so there are also called soluble RNA or sRNA. it present in cytoplasm of 73-93 nucleotides, shorter than mRNA undergo post transcriptional modification. Structure is clover leaf with 4 arms (Figure 6).

Translation

The process of translation occurs within every single cell. Each cell has a nucleus. After transcription,



Activation of amino acid

mRNase move out of the nucleus and enter the cytoplasm. mRNA strand acts as a template for protein synthesis present in the cytoplasm is an enzyme, a aminoacyl tRNA synthetase. The enzyme macro molecule has two binding sites. Once I recognize this, the amino acid methionine this is followed by the binding of the ATP molecule and release of pyrophosphate resulting in activation of amino acid. Finally, the tRNA and the activated amino acid bind together. This aminoacyl ated tRNA is known as met tRNA and is released from the enzyme. This marks the commencement of first stage of protein synthesis (Figure 7).

Initiation stage

During the initiation stage, a small sub unit of a ribosome binds to the mRNA strand. The mRNA strand is made up of codons, which are sequences of three bases. Then the ribosome sub unit moves along the mRNA in five prime to three prime directions until it recognizes the Aug codon or the initiation codon. At this point met tRNA possessing, the anti-codon UAC bears up with the Aug codon of the mRNA, sorry, a large sub unit of ribosome combines with a small ribosomal subunit. The lab subunit shows three sides, the acceptor site or the A site, the peptidyl side, or the P side, the exit side, or the E side, this whole unit forms, the initiation complex. This is followed by the elongation stage (Figure 8).

Elongation stage

During this stage another tRNA carrying molecule of an amino acid approaches mRNA ribosome complex and fits in the A site. Then a bond is formed between Methionine and the amino acid molecule on the tRNA as a result met tRNA becomes d isolated the ribosome then advances a distance of one codon and D isolated tRNA shifts to the E side from where it dissociate. Meanwhile, another tRNA carrying an amino acid molecule attaches to the A site. This is followed by the binding of the amino acid molecules (Figure 9).

Repetition of this process leads to the formation of an amino acid chain. This event is called elongation. Finally, when the UAG codon or the stop codon reaches the A site, elongation is terminated.

Termination

It is the last stage of protein synthesis. The chain of amino acid molecules is released from the ribosome. This released amino acid chain is the protein and this part of protein synthesis is known as translation. Then the tRNA detaches from the mRNA. Ribosome detaches and dissociates into its small and large sub units (Figure 10).

Genetic Code

Genetic code is a dictionary that corresponds with sequence of nucleotides and sequence of Amino Acids. Words in dictionary are in the form of codons, each codon is a triplet of nucleotides, 64 codons in total and three out of these are Non Sense codons, 61 codons for 20 amino acids (Figure 11) [4-6].

Predicting of Secondary Structure of tRNA

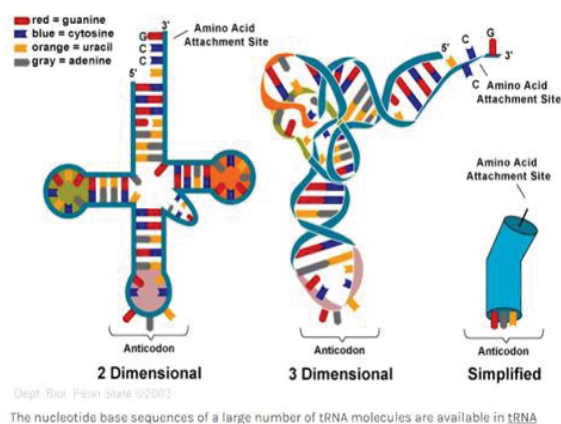


Figure 6 Secondary structure of tRNA.

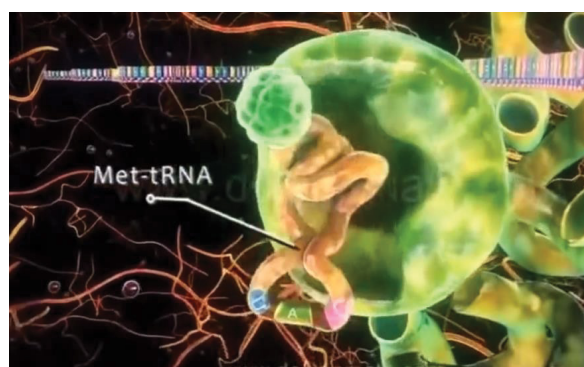


Figure 7 Activation of amino acid.

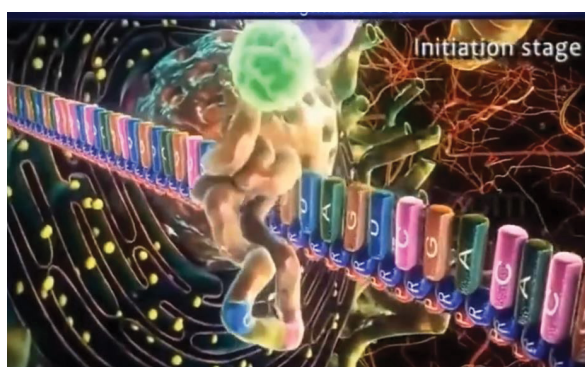


Figure 8 Initiation stage.

Genetic code table

The letters A, G, T and C correspond to the nucleotides found in DNA. They are organized into codons. The collection of codons is called Genetic code (Figure 12).

For 20 amino acids there should be 20 codons. Each codon should have 3 nucleotides to impart specificity to each of the amino acid

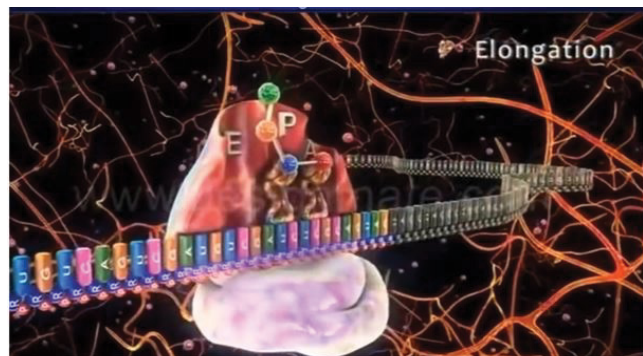


Figure 9 Elongation stage.

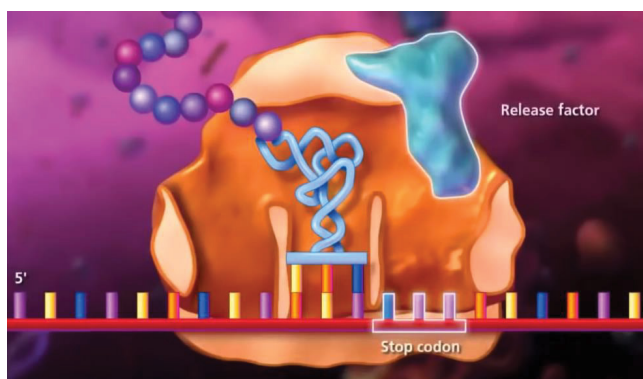


Figure 10 Termination stage.

for a specific codon:

- Nucleotide- 4 combinations
- Nucleotides 16 combinations
- Nucleotides- 64 combinations (Most suited for 20 amino acids)

Genetic code characteristics

Specificity- Genetic code is specific (Unambiguous), a specific codon always codes for the same amino acid. e.g. UUU codes for Phenyl Alanine, it cannot code for any other amino acid.

In all living organism Genetic code is the same, the exception to universality is found in mitochondrial codons where AUA codes for methionine and UGA for tryptophan, instead of isoleucine and termination codon respectively of cytoplasmic protein synthesizing machinery. AGA and AGG code for Arginine in cytoplasm but in mitochondria they are termination codons.

Redundant - Genetic code is redundant, also called degenerate. Although each codon corresponds to a single amino acid but a single amino acid can have multiple codons. Except Tryptophan and Methionine each amino acid has multiple codons.

All codons are independent sets of 3 bases. There is no overlapping; Codon is read from a fixed starting point as a continuous sequence of bases, taken three at a time. The starting point is extremely important and this is called Reading frame.

All codons are independent sets of 3 bases. There is no overlapping; Codon is read from a fixed starting point as a continuous sequence of bases, taken three at a time. The starting point is extremely important and this is called Reading frame.

Non-sense codon

There are 3 codons out of 64 in genetic code which do not encode for any Amino Acid. These are called termination codons or stop codons or non-sense codons. The stop codons are UAA, UAG, and UGA. They encode no amino acid. The ribosome pauses and falls off the mRNA (Figure 13).

Initiator codon

AUG is the initiator codon in majority of proteins - In a few cases GUG may be the initiator codon, Methionine is the only amino

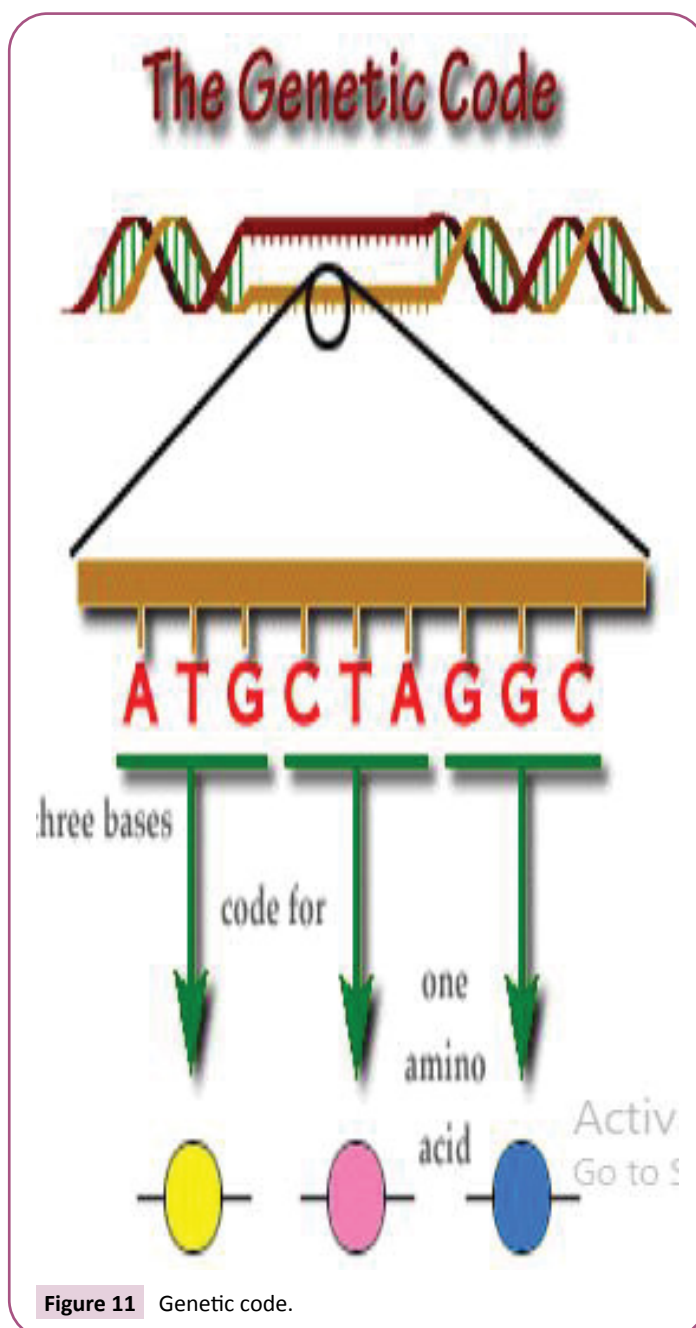


Figure 11 Genetic code.

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

Figure 12 Code table.

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

Figure 13 Stop codon.

Non-sense codon

There are 3 codons out of 64 in genetic code which do not encode for any Amino Acid. These are called termination codons or stop codons or non-sense codons. The stop codons are UAA, UAG, and UGA. They encode no amino acid. The ribosome pauses and falls off the mRNA (Figure 13).

Initiator codon

AUG is the initiator codon in majority of proteins - In a few cases GUG may be the initiator codon, Methionine is the only amino acid specified by just one codon, AUG.

Wobbling phenomenon

The rules of base pairing are relaxed at the third position, so that a base can pair with more than one complementary base. Some

References

- 1 Text Book of Biochemistry for Medical Student by D.M. Vasudevan Sreekumari S. Kannan Vaidyanathan (9th edn), 2019.
- 2 Arnold B, Harvey L, James DE (2000) Molecular cell biology (4th edn). New York: W.H. Freeman, NY, USA.
- 3 Protein biosynthesis: From Wikipedia, The free encyclopedia.

tRNA anticodons have Inosine at the third position. Inosine can pair with UC or A. This means that we don't need 61 different tRNA molecules, only half as many are required. Biochemistry For Medics 13 Wobbling phenomenon First two bases in Codon in mRNA (5'-3') base pair traditionally with the 2nd and 3rd base of the Anticodon in tRNA (5'-3') Non-traditional base pairing is observed between the third base of the codon and 1st base of anticodon. The reduced specificity between the third base of the codon and the complementary nucleotide in anticodon is responsible for wobbling.

Clinical Significance

Mutations can be well explained using the genetic code. a) Point Mutations – Silent, Missense and Nonsense, b) Frame shift mutations.

- 4 Hartman PE (1965) Gene Action. Sigmund R. Suskind: Publisher: Prentice-Hall, USA.
- 5 Martin R (1998) Protein synthesis: Methods and protocols. Totowa, NJ: Humana Press, USA. p. 77.
- 6 Dr. Chhabra N (2001) Genetic code by MD HOD Biochemistry, SSR Medical College, Mauritius.