*IOSR Journal Of Pharmacy And Biological Sciences (IOSR-JPBS)*
*e-ISSN:2278-3008, p-ISSN:2319-7676. Volume 14, Issue 5 Ser. II (Sep – Oct 2019), PP 10-20*
*www.Iosrjournals.Org*

# Studies on Protein Synthesis within a Theoretical Model

Tumpa Saha[1] and Biplab Chattopadhyay[2]
[1]*Department of Chemistry, University of North Bengal, Raja Rammohanpur, Siliguri – 734013, West Bengal, INDIA*
[2]*Department of Physics, Taki Government College, Taki, North 24 Parganas, Taki – 743429, West Bengal, INDIA*
*Corresponding Author; Dr. Biplab Chattopadhyay*

***Abstract:*** *Synthesis of proteins through the phenomenological pathway of transcription and translation, that constitutes the central dogma, has been explored theoretically. Emphasizing the closed-cycle character of the phenomenon, a prototype of the same is framed in mathematical terms with the biological inputs from allied literature. The mathematical prototype is actually a set of three coupled time differential equations signifying time rate of change of DNA, RNA and protein densities occurring in the biological cells of eukaryotes. The prototype has been scrutinized by well-set mathematical tools in regard of its sustainability under detailed stability tests. To judge exact behavioural pattern of the prototype solutions, rigorous numerical simulations of the time differential equations are carried out. Analyses of numerical simulation results with various changing parameters lead to predictive conclusions about different regulatory mechanisms existing in the protein synthesis phenomenology. Some future research directions are indicated too.*

***Keywords****: Proteins as key to Biological growth; Protein Synthesis under Central Dogma; Theoretical Modelling; Synthesis regulation by Model Parameters; Bio-medical applications of Synthesis regulation.*

---

Date of Submission: 16-09-2019 Date of acceptance: 01-10-2019

---

## I. Introduction

Biological proteins are regarded as similes of life. In every living organism proteins signify the livingness and are thus of paramount relevant importance [1-7]. As a matter of consequence the process through which proteins are synthesized is also of substantial interest in case of eukaryotic cells, more so, because protein synthesis leads to cell cycle and growth as well. To increase our knowledge about biological ingredients of living systems such as cells, tissues, organs or most complex system of human body, one needs to focus on identifying the chemical constituents (assembly of molecules) of these biological entities at finer scales [8]. And, the knowledge of growth would mean not only the identification of chemical constituents, but characterization of the interactions within these constituents too. Thus, depiction of biological growth would need understanding of chemical constituents intertwined with the interactions existing amongst them. The process of protein synthesis, which gives vent to various biological matters and thus implies growth, could as well be understood in terms of relevant biological constituents and their interactions.

The biochemical process of protein synthesis, in eukaryotes, is fundamental to biological growth and is essential for regulation of biological processes manifesting life. Synthesis of protein is also a prime necessity for sustenance of life. Pursuit through the avenue of molecular biology produces considerable knowledge about the various molecular assemblies involved in the process of deducing protein from the DNA, including identifying the individual building blocks in the assembly as well as their characterization. But interaction involving various molecular assemblies including DNAs, RNAs and Proteins that constitute the composite transcription-translation pathway of central dogma [9-10] is needed to be investigated in-depth to acquire thorough understanding about the same. We intend to picturize the transcription of DNA to RNA and subsequent translation of RNA to protein in terms of a theoretical prototype involving the molecular abundance of DNA, RNA and proteins and their interactions mediated by various proximal molecular assemblies. Such investigations would reveal as to how mutual interaction between relevant molecular populations adds to the regulatory mechanism of protein synthesis and would thus probably generate informations regarding the control mechanism of protein synthesis, necessary at times, to alleviate development of certain abnormalities. Our motto is to frame a dynamical model involving time differentials of three variables DNA, RNA, protein and to study the model to bring out various characteristic features inherent to the dynamics of protein deduction through transcription-translation pathway. There have been earlier attempts to cast the dynamics of protein production within the relevant theoretical framework, but such efforts were limited only to the part of translation of mRNA into protein [11-14]. Our interest lies in the composite process of transcription of DNA to RNA and translation of a part of RNA to protein and we would be implementing these concepts within a theoretical framework

comprising of differential equations to depict the phenomenology of the flow of gene expression in steps from DNA to protein.

The paper is organized in seven different sections. Section-I provides introductory description. In Section-II the detailed dynamic process of protein synthesis phenomenon under central dogma is elaborated and the cyclic characteristic of the process is emphasized. Formulation of the theoretical model representing the closed chain dynamic process of protein synthesis is discussed in Section-III. Results of analytical studies of the model equations are presented in Section-IV. Section-V includes detailed outcomes from numerical simulation of the model equations. In Section-VI, we discuss in detail the results obtained from the numerical simulation of the theoretical model. Concluding remarks along with some predictive assertions are presented in Section-VII.

## II. Dynamical and Cyclic Progression of Protein Synthesis

Let us now look into the greater details of dynamical process of protein synthesis through the transcription-translation pathway with the intention to explore in detail as to how the dynamicity is maintained through the closed chain or cyclic progression.

Gene expression or genetic codes are contained in the double helix DNA that makes part of the nucleus of a biological cell. Cells are considered to be the basic building blocks of any biological entity. Inside a eukaryotic cell a large nucleus is embedded within a semi fluid medium called cytoplasm which also includes other cell-organelles. Dimensional measure of a typical eukaryotic cell ranges around 20 micrometer (μm) and that of a nucleus is 5 μm in diameter [15]. In the nucleus a grainy looking material, which is actually thread-like, called chromatin, resides within the semi fluid medium of nucleoplasm. Chromatin consists of DNA and associated binding proteins. Nucleus in a cell is being surrounded by a double membrane envelop which has nuclear pores of typical dimension of 100 nanometre (nm) [15] and through these pores, passage of proteins and ribosome take place between the nucleus and the cytoplasm of the cell.

Within the chromatin matter DNAs (tangled with proteins) are organized into long structures called chromosomes. Each DNA consists of two helical chains which coil around the same axis with the coil radius approximately ~1nm [9]. DNA molecules are constituted by four bases as adenine (abbreviated as A), guanine (G), thymine (T) and cytosine (C) each of which are attached to sugar and phosphate to complete the formation of corresponding nucleotide. Many nucleotides are organized in specific sequence, being bound to each other through phosphate groups, thereby gives rise to DNA strands. Two DNA strands remain stabilised in helical shape through hydrogen bonding between the specific nucleotide bases in separate strands. Stretch of a DNA chain could vary in length so as to be comprised of 14 to about 100 base pairs [10]. Sequence of nucleotides in DNA strands remain in the form of many triplets and these triplets are termed as genes which bear within them the instructions to produce specific proteins [16-17].

Double helix DNAs in the nucleus give in to the synthesis of single stranded RNAs through the process of transcription. Actually genetic information in the triplets of DNA nucleotide sequence is copied to the RNA nucleotide (triplet) sequence with the substitution of the base thymine (T) (in DNA) by the base uracil (U) in RNA and this substitution becomes uniform throughout the strand. In eukaryotic nucleus, three different enzymatic proteins, such as RNA-polymerase of labels I, II, III initiate the process of transcription and thus very many RNAs result [10,18]. RNA-polymerase I & III lead to the synthesis of transfer RNA (tRNA), ribosomal RNA (rRNA) and various small RNAs. RNA-polymerase II transcribes most genes giving vent to abundant messenger RNA (mRNA) that encode proteins and some supporting RNAs. Though RNA-polymerase initiates the process of transcription, the same is carried through and completed by various transcription factors made of composite proteins. In an mRNA, each group of three nucleotides in consecutive sequence (triplets) is called a codon which prevalently specifies an amino acid, the building block of a polypeptide molecule or protein. Synthesis of mRNA from DNA through transcription is balanced by dilution and degradation processes [1]. Degradation of mature mRNA yields corresponding nucleotides which further inflicts transcription of genes, thus adding to the abundance of RNA. It should be noted here that most RNAs synthesized in the nucleus of the cell, diffuse through the nuclear pore complexes to the relevant location of cytoplasm. Small molecules (dimension less than 50000 Daltons) would easily pass through the pores whereas large molecules (those complexed with proteins) are passed by the nuclear membrane through the special energy-diffusion mechanism [10]. Various protein matters in the form of polymerase and transcription factors that take part in the transcription process in the nucleus, also moves through the pores from cytoplasm to the nucleus. Actually, movement of various macromolecules, through nuclear pores from or to the nucleus in a cell, is maintained by various nuclear transport receptors which actually escort such movement of macromolecules [9].

Production of mRNA through transcription clears the pathway to proteins. The genetic informations stored in the nucleotide sequence of mRNA are translated into amino acids. In transcription of DNA to RNA, as if a coarse-grained message is converted into a fine-tuned text without changing the language. However, in translation, the texted information in mRNA is changed into a different language in protein. This is because there are only four different nucleotides in mRNA whereas there are twenty different types of amino acids in

proteins which signify that a clear one-to-one correspondence between the two is not possible. Essentially, through the transcription-translation pathway, specific sequences of nucleotides in a gene are carried to the proteins through the mRNA, following a set of rules collectively known as genetic code [17].

The RNA is like a linear polymer comprised of four different nucleotides where a unit of polymer is constituted by three nucleotides and these triplets are called codons. When translation of mRNA to protein takes place, each codon either yields to one of the twenty possible amino acids or may specify a step to the translation process. Four different nucleotides of RNA may give rise to $4^3 = 64$ possible combinations of nucleotide triplets or codons. Of the 64 codons, only three (UAA, UAG & UGA) signal stop to translation and the rest 61 codons corresponds to the 20 different amino acids. However, one of the 61 triplet codons AUG plays the role of initiation codon signalling the start of translation and also may yield to the methionine amino acid. The set of values, as described above, under which messages from genes are transported to amino acids through triplet codons, is actually termed as genetic code [19-21]. Having thus described the genetic flow, let us consider the specifics of translation of mRNA to amino acids.

Translation of the sequenced nucleotide triplet of an mRNA takes place in the cytoplasm of a cell on the ribonucleoproteins assembly called ribosome. A ribosome is composed of two subunits, one large and one small and it is a complex comprising of 67% RNA and 33% proteins [10]. In the structure of ribosome, signatures of rRNA is prevalent than the associated proteins. An mRNA suffers translational transformation in certain well-defined steps [22-27]. First, mRNA strand binds to the large and small subunits of a ribosome, with small subunit being attached to the initiation codon (AUG). A tRNA molecule with an attached amino acid, termed as aminoacyl-tRNA, first recognizes the ribosomal attachment of an initiation codon (AUG) and thus bind to a triplet of mRNA strand selectively. Rule of selection is that an mRNA codon allows only the complementary anticodon triplet of tRNA to bind with it forming the codon-anticodon pair. Note that association of small ribosomal subunit to (AUG) codon actually signals the start of aminoacyl-tRNA binding to an mRNA. Then, another aminoacyl-tRNA binds to the adjacent codon of mRNA and thereby a peptide bond [28] is formed between the amino acids. Just after this bonding the first tRNA is released and another tRNA gets attached adjacent to the second tRNA forming another peptide bond thereby increasing the polypeptide chain one by one in the sequence. This process continues until a stop codon is reached by the bond-affine tRNA, after which completed polypeptide chain is released. It has been observed that during the translation phase of protein synthesis, ribosome operate with considerable efficiency, that is a single ribosome of an eukaryotic cell adds about two amino acids to a polypeptide chain in just one second [10]. After the release of polypeptide chain, the ribosome gets dissociated into two separate subunits with the help of certain protein matters, and in the next moment the two subunits come to an optimal state to reinitiate translation on the same mRNA molecule. The stretch length of a polypeptide chain typically varies between (being comprised of) 40-600 amino acid units [29]. Synthesis of most protein molecules requires between 20 seconds to several minutes [10]. Within the stretch of this time, it is possible that many initiations for translation may take place on each mRNA molecule. Observations point to the fact that a sizeable number of polypeptide chains may be yielded by a single mRNA molecule [30].

The whole process of protein synthesis is thus a thoroughly dynamic phenomenon in which genetic information flows in steps through three different bio-molecules DNA, RNA and Protein and they make up to a closed cycle till the requirement of a specific protein is fulfilled. Synthesis of a specific protein is started owing to signalling from central nervous system of body as per requirement of such protein by body. Such signalling inflicts DNA population to take up the task of replication so as to increase its abundance to required level. Then the DNA molecule population give in for the making of RNA molecule population by transcription and these RNAs further give in to the population of proteins through translation. Most proteins start folding during the time they are being synthesized and acquire compact structure to become operative [29]. Functional proteins and transcriptional proteins, making together the part of transacting proteins [31], remain ceaseless and take part in various activities. The transcriptional proteins regulate activities of RNA-polymerase and hence the process of transcription. Functional proteins, when accumulated in required loads, proceed to perform pre-assigned activities. The cis-acting part of produced protein acts as in-situ gene [31] in the sense that these proteins hold the nucleotides in the DNA double helix at chromatin congregation in the nucleus. Various enzymatic proteins also play their respective roles in progressing further synthesis of proteins. DNA polymerase inflicts replication of DNA, whereas RNA polymerase plays role in transcription and ribosomal proteins play its part in the process of translation for production of proteins. Process of protein synthesis is thus not an open sequence process, but practically a closed cycle process where synthesized proteins keep on cycling the transcription-translation pathway to further enhance the profusion of necessary proteins [9].

## III. Formulation of the Theoretical Model

It is apparently conclusive from the pursued discussions in previous sections that the synthesis of proteins under central dogma follows the closed chain dynamical pathway. Considering the density of the DNA,

RNA and Protein molecules in typical eukaryotic cells as variables, a dynamical model for protein synthesis is proposed. We assume densities of DNA (D), RNA (R), and Proteins (P) to be assigning the weight of these biomolecular populations at a specific time and set up three coupled differential equations to represent the dynamical model in question. Time differential of each variable gives the time rate of change of respective population and the same is defined in terms of interactions of these populations with each interaction either enriching or degrading a population [32-33]. For the present case of the dynamic process of protein synthesis, inter-population two-body type interactions are governed by the law of mass action and are mostly mediated by enzymatic proteins or other relevant biological matters.

Interactive biological processes or bio-molecular entities that alter the measure of active DNA in the cell are predominantly four fold. Chromatin matter inside the cell nucleus consists of DNA double helix in large proportions and this adds to the DNA load at an assumed constant rate (a). The active portion of DNA, being influenced by DNA-polymerase, gets replicated [9] at a constant assumed proportion ($\chi$ ). Proteins interact with active DNA at a definite considered proportion ($\alpha$) to regulate its measure. To our perception, cis acting proteins [31] that have tangling association with DNA inside the chromatin, actually hold DNAs and thus keep these DNAs active for transcription. Regulation of the measure of active DNA is achieved by incorporating an inverse exponential function with dimensionless argument $P/D$ with DNA-protein interacting term. Depending on the proportion of proteins, DNAs proportion is such regulated that, for small proportion of proteins less than that of DNA ($P < D$), the DNA abundance gets enhanced as proteins. However, for large loads of proteins $P > D$, the relevant enhancement of DNA becomes inversely proportional to P, and for $P \gg D$ the contribution becomes insignificant. DNAs also interact with various transcriptional proteins, under the influence of RNA-polymerase, to add to the abundance of RNAs, and thereby, degrade the measure of DNA [5,29] at a constant rate ($\beta$). Incorporating all the above processes, time-based change of DNA density can be cast as a time-differential equation

$$\frac{dD}{dt} = a + \chi D + \alpha DPe^{-(P/D)} - \beta DP \qquad (1)$$

where the number density of DNA molecules (D) as well as densities of RNA (R) and protein (P) molecules are defined in $\mu m^{-3}$ (per cubic micrometer).

Abundance of RNA in the system is modified through four different processes. Interactions of DNA with proteins lead to the transcription of DNA and this enhances the RNA measure [5,31] at an assumed fixed rate ($\gamma$). Mature mRNA suffers degradation [1,34,35] and then interacting with DNA they transcribe DNA to RNA. We thus consider mutual interaction of DNA and RNA to enhance RNA measure [1] at a rate ($\nu$). Ribosomal proteins interact with RNAs and through the process of translation, lead to protein production [9], which makes us to consider degradation of RNA population at a definite rate ($\delta$). There is a natural loss of RNAs, and we consider this to happen at a rate ($\mu$). All the processes thus inflicting change in the measure of RNA or its density, if taken together, lead to the time-differential equation

$$\frac{dR}{dt} = \gamma DP + \nu DR - \delta RP - \mu R \qquad (2)$$

Proteins, which are produced in the cytoplasm of a cell, may suffer modification in three different ways. Interaction of ribosomal proteins with RNA gives rise to the enrichment of proteins [9] (through translation) at a constant presumed rate ($\lambda$). Cis-acting proteins that bind to the chromatinic DNA through proportional interaction are assumed to reduce the protein abundance [9,31] by a constant factor ($\eta$). A fraction of protein is lost from its contemporary loads through natural processes at a rate $\mu'$. Change in protein density, thus, can be represented as a time-differential equation

$$\frac{dP}{dt} = \lambda RP - \eta DP - \mu' P \qquad (3)$$

Clubbing together the equations (1), (2) and (3), representing time rate of change of the measures of bio-molecular populations of DNA, RNA and proteins respectively, we have the theoretical model for protein-synthesis dynamics as a set of three coupled differential equations

$$\begin{aligned} \frac{dD}{dt} &= a + \chi D + \alpha DPe^{-\frac{P}{D}} - \beta DP \\ \frac{dR}{dt} &= \gamma DP + \nu DR - \delta RP - \mu R \\ \frac{dP}{dt} &= \lambda RP - \eta DP - \mu' P \end{aligned} \qquad (4)$$

To better understand the dynamical trait of the whole process of protein synthesis and to identify the significance of various intermolecular interactions involved in the process, one needs to analyze the set of

equations (4) representing the mathematical archetype of protein synthesis [12]. A fundamental purpose of such analysis is to find out the regulatory mechanisms embedded within the process of protein synthesis and to apply such mechanisms to alleviate bodily abnormalities arising out of or associated with protein synthesis.

## IV. Analytical Studies of the Model Equations

As emphasized, application of knowledge-yields from protein synthesis requires thorough analysis of model equations (4). Such analysis is carried out in two different fronts, one being the theoretical analysis to gain confidence about exactness and sustainability of the model and the other being computational simulation of model equations to gather precise understanding about its solutions. The part of theoretical analysis revolves around judging the stability of model solutions by exploring various equilibria and through satisfying the standard prescription of Routh-Hurwitz criteria [36-38] for stability.

Model equations (4) are observed to be smooth functions of variables D, R and P aa well as the parameters involved. Further, enforcing the condition that the variables and parameters of the model would assume values only from the positive part of real line, we arrive at the conclusion that model solutions hold the existence, uniqueness and continuity properties in the positive octant of the coordinate space.

Stability of asymptotic model solutions bears ramifying significance for the sustainability of the corresponding biological system or process. To judge sustainability, one needs to explore various equilibria points in the variable space involving asymptotic stable solutions (fixed point solutions $D^*$, $R^*$ and $P^*$) such as E (0,0,0) , $E_1(D_1^*,0,0)$ , $E_2(D_2^*,R_2^*,O)$ and $E^*(D^*,R^*,P^*)$. Out of these equilibria, E (0, 0, 0) is the trivial equilibrium that exists, but with minimal significance to model stability. Fixed point equilibria are obtained from asymptotic model equations

$$\begin{aligned} a + \chi D^* + \alpha D^* P^* e^{-\frac{P^*}{D^*}} - \beta D^* P^* &= 0 \\ \gamma D^* P^* + \nu D^* R^* - \delta R^* P^* - \mu R^* &= 0 \\ \lambda R^* P^* - \eta D^* P^* - \mu' P^* &= 0 \end{aligned} \tag{5}$$

Equilibria $E_1$ and $E_2$ exist with the solutions being

$$D_1^* = \left|\frac{a}{c}\right|$$

where the modulus sign is for making the solution biologically meaningful.

and $$D_2^* = \frac{\mu}{\nu}, \quad R_2^* = 0$$

Solutions for the general non-trivial equilibrium are,

$$D^* = \frac{\beta P^* - \chi}{\alpha} \tag{6}$$

$$R^* = \frac{\eta\beta P^* - \eta\chi + \alpha\mu'}{\alpha\lambda} \tag{7}$$

$$P^* = \frac{B \pm \sqrt{B^2 - 4AC}}{2A} \tag{8}$$

where, $A = (\alpha\beta\gamma\lambda + \beta^2\nu\eta - \alpha\beta\delta\eta)$, $B = (\alpha\chi\gamma\lambda + 2\beta\chi\nu\lambda - \alpha\beta\nu\mu' - \alpha\chi\delta\eta + \alpha^2\delta\mu' + \alpha\beta\mu\eta)$ and $C = (\alpha\chi\nu\mu' - \chi^2\nu\eta - \alpha\chi\mu\eta + \alpha^2\mu\mu')$

In the above solutions for general non-trivial equilibrium, we expand the exponential function and keep only the leading term in $e^{P^*/D^*}$, which suffices to incorporate the full effect of negative exponential function. Thus, the exponential function yields a multiplicative factor $D^*/P^*$ wherever it occurs. Solution for $P^*$ in the general equilibrium, often termed as interior equilibrium, is entirely in terms of the model parameters. For $P^*$ to assume real values confined in the positive part of the real line one needs to satisfy the parametric condition

$$4(\alpha\beta\gamma\lambda + \beta^2\nu\eta - \alpha\beta\delta\eta)(\alpha\chi\nu\mu' - \chi^2\nu\eta - \alpha\chi\mu\eta + \alpha^2\mu\mu') \geq 0 \tag{9}$$

This imposing condition provides a cap on the values the model parameters can assume and is thus significant for assigning default values for the model parameters.

Stability of the general equilibrium solution could further be judged by evaluating Routh-Hurwitz condition for which we need to linearise the model equations around the fixed point solutions in the equilibrium and obtain the Jacobian matrix $(J)$ [38]. Linearization is achieved by introducing new variables as $X(t) = D(t) - D^*$, $Y(t) = R(t) - R^*$ and $Z(t) = P(t) - P^*$ and this leads to the matrix equation

$$\begin{pmatrix} \dot{X} \\ \dot{Y} \\ \dot{Z} \end{pmatrix} = \begin{pmatrix} \chi - \beta P^* & 0 & -\beta D^* \\ \gamma P^* + \nu R^* & \nu D^* - \delta P^* - \mu & \gamma D^* - \delta R^* \\ -\eta P^* & \lambda P^* & \lambda R^* - \eta D^* - \mu' \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = J \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \tag{10}$$

Where the matrix $J$ represents the Jacobian of transformation. While getting into the Jacobian matrix, we are being guided by the observation that the ratio $P^*/D^*$ is quite large and a negative exponential of this ratio, that is $e^{-\frac{P^*}{D^*}}$ have marginal or no effects in terms of the asymptotic solutions. Thus, in the Jacobian, we

neglect terms involving $e^{-\frac{P^*}{D^*}}$ for all practical purposes. Evaluation of trace of $J$, that is $Tr\ J$ and determinant of $J$, that is $det\ J$, yield

$$Tr\ J = (\chi - \mu - \mu') + (\nu - \eta)D^* + \lambda R^* - (\beta + \delta)P^* \quad (11)$$

$$Det\ J = (\chi - \beta P^*)[(\nu D^* - \delta P^* - \mu)(\lambda R^* - \eta D^* - \mu') - \lambda P^*(\gamma D^* - \delta R^*)] - \beta D^*[\lambda P^*(\gamma P^* + \nu R^*) + \eta P^*(\nu D^* - \delta P^* - \mu)] \quad (12)$$

The above Jacobian matrix of transformation is observed to follow the Routh-Hurwitz criteria, that is $Tr\ J < 0, det\ J > 0$, for a chosen set of values of model parameters, termed as default values. It should be noted here that default values of parameters are obtained by complying to imposing parametric conditions (9) yielded by stability analysis as well as monitoring numerical solutions of model variables for changing parameter values. Default values of parameters are also judged by looking at the magnanimity of various considered interaction from allied literature, as far as practicable.

## V. Outcomes of the Numerical Simulation of Model Equations

Gaining confidence about the sustainability of the proposed model for protein synthesis, through analytical parleys, we proceed for simulation solutions of the model, in order to understand exactly the characteristic features of the dynamical process of protein synthesis starting from the DNA. Estimated default set of parameters, used for numerical calculations, are as listed in Table-1. Model equations are solved numerically implementing a fourth order Runge-Kutta method [39] as functions of time $(t)$ with initial $t = 0$ seeds for variables as $D_0 = 100,\ R_0 = 50,\ P_0 = 20$. However, as we checked, change of seeds does not affect the time varying characteristics of solutions or their asymptotic measures. Note that, default values to parameters (as in Table-1) are normally assigned, but, in practice, their variations are also considered whenever necessary. This is to explore extensively the various characteristics of protein synthesis dynamics as well as regulatory properties, if any, of various model parameters.

In Figure-1, we have presented time varying solutions of model variables $D(t),\ R(t),\ P(t)$ where plot in the left pane refers to default model parameters and that in the right pane for $\beta = 0.004$ with all other parameters as in Table-1. We observe that, for default parameters, model variables solution first show oscillatory nature of diminishing amplitude along the increasing time-axis and at large enough time, i.e. asymptotically, become stable and single-valued. In the right pane, at small time, solutions show large amplitude oscillations which decrease with time, but asymptotically the solutions are still oscillatory with fixed amplitude. Here, asymptotic solutions are stable but multivalued limit-cycles having an upper and a lower bound. These oscillating features of model solutions are considered to be characteristic to theoretical model of biological systems. We have checked time series solutions of model variables for each parameter being assigned values other than the default one and found that in each case, the asymptotic solutions are always stable and either single valued or multivalued. Single and multi valuedness of solutions are indicative of specific role played by the relevant parameter in the process of protein synthesis.

**Table-1:** Set of parameters used with their default values

| Parameter | Definition | Default Value (UNIT) |
|---|---|---|
| $a$ | Initial concentration of DNA in cell nucleus | 10.0 $\mu m^{-3} sec^{-1}$ |
| $\alpha$ | Strength of DNA-protein interaction contributing to DNA population | 0.01 $\mu m^3 sec^{-1}$ |
| $\beta$ | Rate of loss of DNA contributing to RNA population through interaction with transcriptional protein | 0.0001 $\mu m^3 sec^{-1}$ |
| $\chi$ | DNA lysis constant | 12.0 $sec^{-1}$ |
| $\gamma$ | Strength of DNA-protein interaction contributing to RNA population | 0.01 $\mu m^3 sec^{-1}$ |
| $\nu$ | Strength of DNA-RNA interaction contributing to RNA population | 0.2 $\mu m^3 sec^{-1}$ |
| $\delta$ | Strength of RNA-protein interaction contributing to protein population | 0.005 $\mu m^3 sec^{-1}$ |
| $\mu$ | Rate of natural loss of RNA including through degradation & dilution | 0.01 $sec^{-1}$ |
| $\lambda$ | Strength of RNA-protein interaction contributing to protein population | 0.0099 $\mu m^3 sec^{-1}$ |
| $\eta$ | Strength of protein-DNA interaction contributing to DNA population | 0.02 $\mu m^3 sec^{-1}$ |
| $\mu'$ | Rate of natural loss of protein including through degradation & dilution | 0.2 $sec^{-1}$ |

**Table-1:** Set of parameters used in solving Model Equations with their default values

Next, we embark on varying each of the parameters and scrutinize any corresponding changes in the model solutions. This is pursued to better judge the regulatory effects of individual parameters. A consequent planar plot of asymptotic fixed point solutions as functions of a model parameter is termed as phase diagram. Analyses of such phase-diagram data produce precise characteristic bearing of the parameter on the solutions and indicate the regulatory effectiveness of the corresponding parameter.

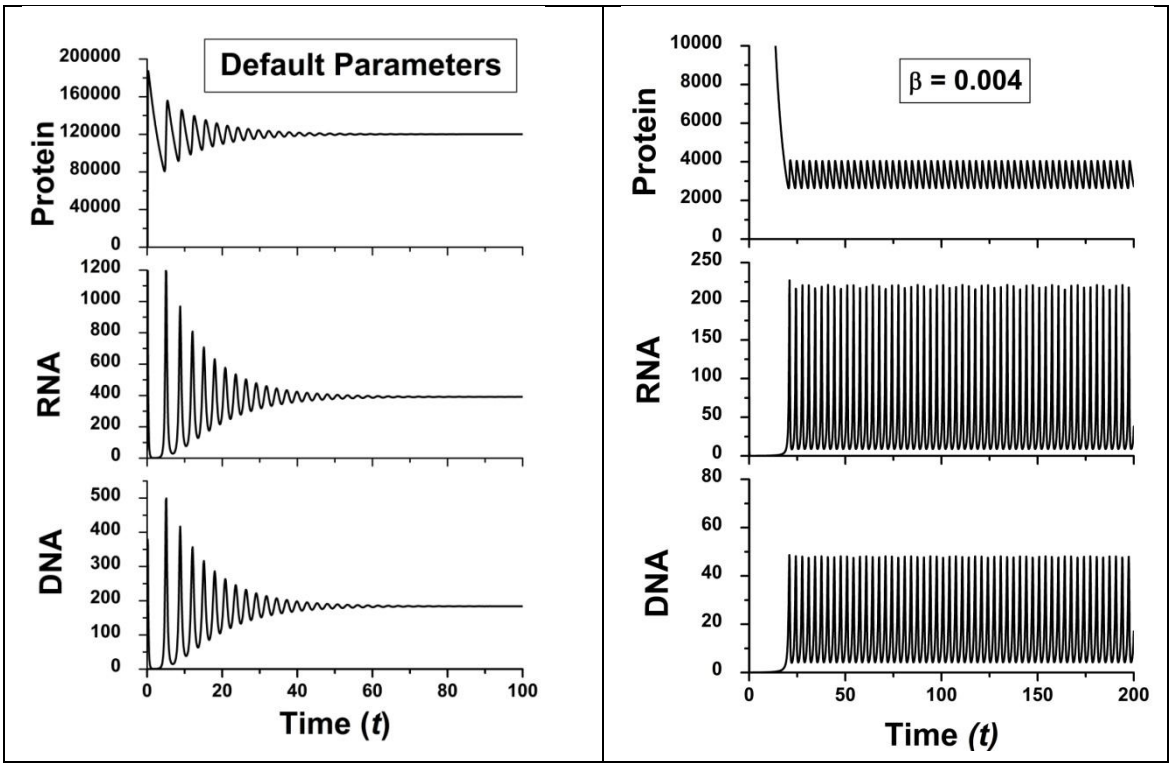


**Figure-1:** In the left pane, time varying solutions of model variables are shown with the time taken in units of seconds. Here, model parameters are assigned to their default values as in Table-1. In the right pane**, s**olutions for archetype variables are plotted as functions of time (in seconds) with parameter $\beta$ set to the value as shown in the figure and all other parameters being at their default values.

In Figure-2, we plot phase diagrams corresponding to parameters $\chi$ (left pane) and $\beta$ (right pane). In reference to the plot with varying $\chi$, we see that asymptotic solutions for DNA, RNA and Proteins, $D^*, R^* \; and \; P^*$ respectively, are single valued fixed-point solutions. For $\chi$ very small (close to zero), asymptotic solutions acquire non-zero fixed values and with increasing $\chi$, all stable solutions increase monotonically but in marginally non-linear fashions. An observation on selective numerical values of stable solutions reveals that for small $\chi$, $P^*$ is about 200 times the $D^*$, but for $\chi \sim 5$, $P^*$ turns out to be 400 times the $D^*$. This means that, with increasing $\chi$ protein density grows at a much faster rate than the corresponding DNA density.

The phase diagram in terms of parameter $\beta$ shows that all the asymptotic stable solutions are single valued till about $\beta \sim 0.0015$, after which the solutions become multivalued limit cycles asymptotically. Beyond $\beta \sim 0.0015$, solutions are bound within a definite range which are depicted as bifurcated curves for each of $D^*, R^* \; and \; P^*$ plots. It is also observed here that $D^*$ and $P^*$ nearly keep a definite relative proportion $1:200$ before these solutions attain the asymptotic multivalued state.

Plots of $D^*, R^* \; and \; P^*$ as functions of $\gamma$ (left pane) and $\lambda$ (right pane) are presented in Figure-3. In both the plots here, asymptotic solutions are single valued. For $\gamma$ extremely small, all the solutions acquire fixed values keeping a definite proportion between them $D^*: R^*: P^* \sim 1: 2: 40$. With increasing $\gamma$, $D^*$ and $R^*$ fall sharply but $P^*$ increases very slowly. For $\gamma \sim 0.015$, both $D^*$ and $R^*$ population densities become insignificant, whereas $P^*$ remains at its high normal value. Corresponding to the parameter $\lambda$, representing RNA-protein interaction that contributes to protein population density and signify the rate of translation, phase diagram plot shows characteristics similar to that of $\gamma$, the rate of transcription. As can be observed, for $\lambda$ very small close to zero, $D^*$ assumes fixed value but $R^*$ is quite high, whereas $P^*$ acquires a fixed value maintaining $D^*: P^* \sim 1: 40$. Increase of $\lambda$ makes $R^*$ falls more sharply than $D^*$, and $P^*$ remains nearly same or increases monotonically.

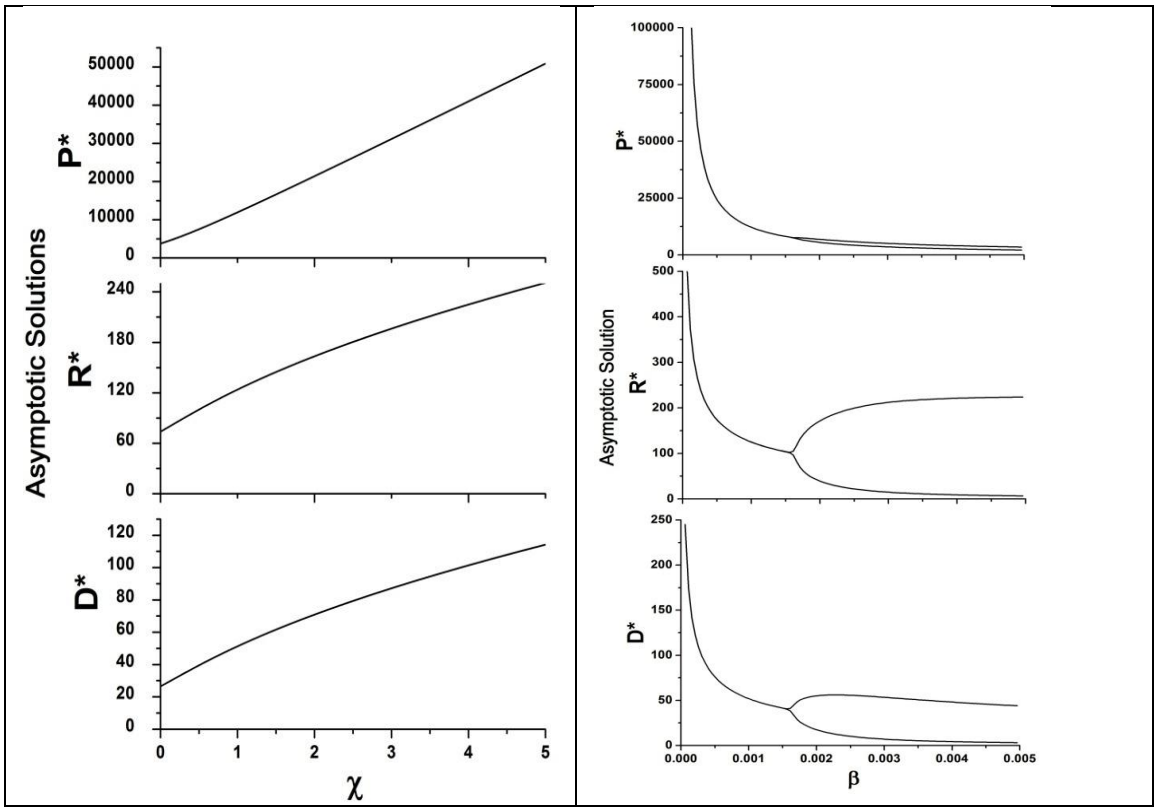


**Figure-2:** Phase diagram depicting asymptotic stable values of archetype variables as functions of the parameter $\chi$, is incorporated in the left pane of the figure, with all other parameters (excepting $\chi$) being kept at their default values as in Table-1. In the right pane**,** asymptotic stable values of model variables are plotted as functions of $\beta$ with other parameters kept at their default levels as in Table-1.

It is to be noted here that, phase diagram data, for all other remaining parameters are plotted and the planar viewgraphs are analysed. Important changing characteristic features are found in case of phase diagrams corresponding to parameters $\nu, \delta, \eta \; and \; \mu'$. Our observations for all such phase diagrams with significantly altering solutions generate a perception that all these allied parameters could be utilized to control protein synthesis in various pertinent manners as per specific requirements.

## VI. Numerical Simulation Results and Discussion

In all the viewgraphs, it has been commonly observed that the asymptotic value of protein population density is quite a many times that of the DNA. As a specific example, refer to the model solutions for default parameters (Figure-1, left pane) where asymptotic stable values of variables are $D^* \sim 200$ and $P^* \sim 120000$. Further, refer to the phase diagram corresponding to the parameter $\gamma$ (Figure-3, left pane), where, for very small $\gamma$, $D^* \sim 3000$ and $P^* \sim 120000$. These observations are imperative on us to justify the normally occurring high value of protein population density $P^* \sim 10^5$ $\mu m^{-3}$ (per cubic micrometer). An order of magnitude estimation of protein molecule units per $\mu m^{-3}$ here would be relevant to judge the very high protein density in our model.

It is reported in the allied literature of protein synthesis that the mass of a protein molecule ranges normally between$10 - 100$ Kilo Dalton [40]. This estimation is such that $50$ KiloDalton could be considered as an average of the protein molecule mass. The mass unit Dalton actually represents the atomic mass unit (amu) such that $1$ Dalton is equivalent to $1$ amu [1]. In this sense, the average level of molecular mass of a protein unit could be $50000$ gms. Further, literature shows that, in case of eukaryotes, concentration of cytoplasmic protein is about $180$ mg/ml [9,41]. Molar mass of proteins, that is $50000$ gms (1 Mole), contains Avogadro number ($N = 6.023 \times 10^{23}$) of protein molecules. Thus $180$ mg could have protein molecules as high as $2 \times 10^{18}$ /cc. Our model variables are assigned the unit number density per micrometer cube ($\mu m^3$), which is logical in comparison of the dimension of a common eukaryotic cell that is few hundred $\mu m^3$ [15]. Now, the protein molecule number density, when converted in terms of $\mu m^3$, yields the highest average measure as $2 \times 10^6$ $\mu m^{-3}$. Within our model, the order of magnitude of asymptotic protein population density $P^* \sim 10^5$ $\mu m^{-3}$ is of comparable order with the above estimation (based on experimental findings in the literature).

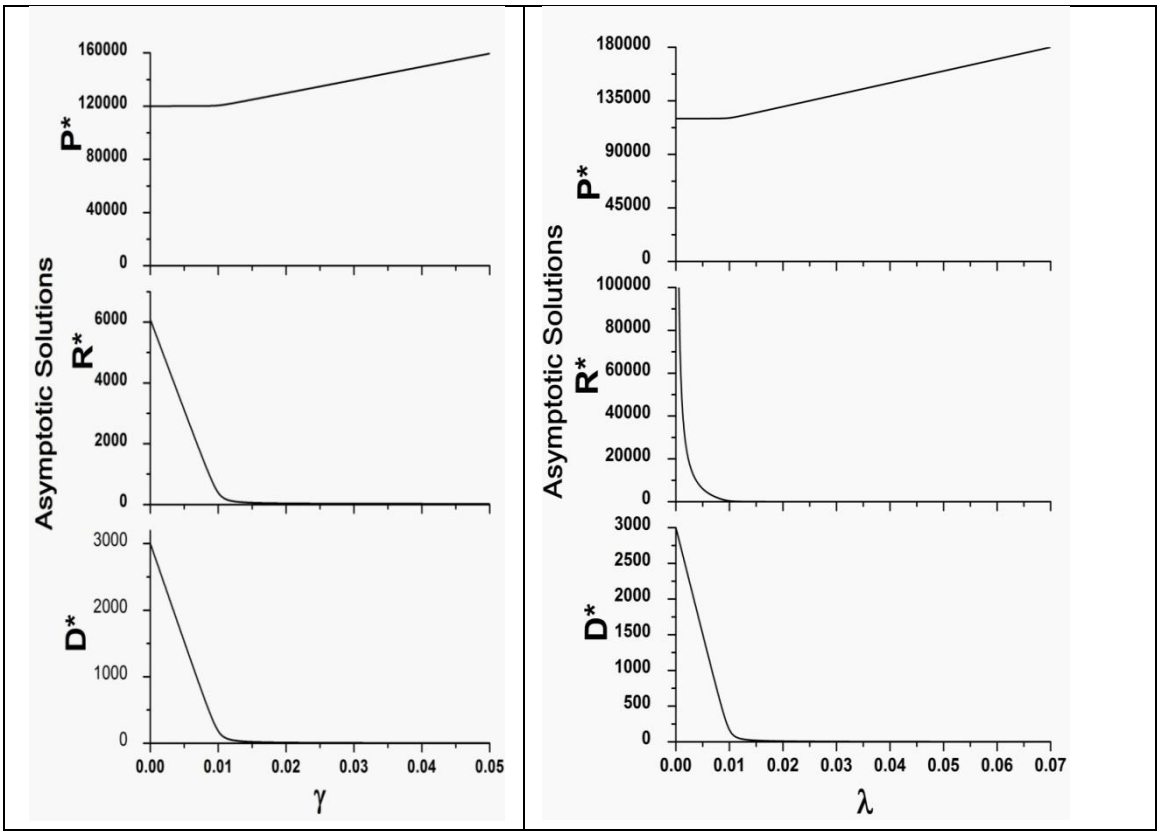


**Figure-3:** Left pane: Stable values of DNA, RNA and Protein populations are illustrated in planar viewgraph with increasing values of parameter γ. All other parameters are kept at their respective standard values as in Table-1. Right pane: Phase diagram for model variables are portrayed as functions of the parameter λ with other parameters assuming fixed default values as in Table-1.

Critical observations on various phase-diagram data reveal that those having significant varying characteristics, could be grossly clubbed in different categories. To be specific, phase diagrams corresponding to χ and δ have certain apparent similarities, phase diagrams corresponding to γ and λ have nearly similar features and so are those corresponding to η and $\mu'$. In terms of parameters β and ν, asymptotic solutions of model variables are multi-valued and stable in specific ranges of parameter values. Actually, the act of categorization of phase-diagram data is to emphasize that the parameters in the same category could be utilized to regulate (or control) the protein synthesis in similar fashion. However, exact quantitative details of regulation would definitely be parameter specific.

In case of the parameter χ signifying the replication rate of DNA, any increment of χ causes protein production to grow at an ever increasing rate whereas DNA and RNA measures grow at decreasing pace. Thus, if we need a very large proportional loads of proteins as compared to the DNA, we have to set the numerical value of replication rate very high which would probably necessitate enhancement of DNA-polymeraze in the cell involved in synthesizing the required protein. The parameter χ, representing DNA replication rate, could thus be exploited to regulate protein synthesis. In terms of parameter δ, representing the rate at which RNA measure is degraded by taking part in the process of translation, we find that for a range of values of $\delta\ (\sim 0.01 - 0.03)$ protein population density could be kept constant even though DNA and RNA population densities vary considerably. The parameter δ thus could provide associated characteristic regulation.

Phase diagrams corresponding to parameters γ, the transcription rate, and λ, the rate of translation, have almost similar characteristic features even in terms of numerical values. Increase in either of the rates of transcription (γ) or translation (λ), lead to drastic fall of both DNA and RNA abundances. Till the values of γ or λ $(\sim 0.01)$ when DNA and RNA face degradation, protein population density remains constant. However, as DNA and RNA abundances become marginal, protein density rises linearly. These parameters (γ and λ) could thus be tuned as per necessity to get high protein density keeping DNA and RNA at marginal levels and hence the allied regulation. Phase diagrams for η (rate of protein loss owing to their chromatinic DNA attachment) and $\mu'$ (natural protein degradation rate) show that by tuning these parameters, one could have a nearly steady protein density while DNA and RNA could vary in wide ranges. Here, again, associated parametric regulation could be enforced as per necessity.

The parameter β stands for the rate of DNA degradation by way of transcription. Phase diagram for β shows drastic reduction of all three populations, DNA, RNA and Proteins, with the increasing rate of DNA degradation. Thus, where the purpose is to downgrade the protein abundance, we can suitably tune the parameter β, keeping in mind that beyond some value of β ($\sim 0.0015$) the asymptotic solutions are stable but multivalued. Corresponding to the parameter ν, representing the rate of transcription through matured mRNA channel, we observe that suitable tuning of the parameter (in a wide range) could retain normal level of protein population density along with moderate densities of DNA and RNA.

## VII. Concluding Remarks and Predictive Assertions

In the concluding overture, we stress that, we have formulated a theoretical dynamic model of the synthesis of protein from a DNA population density (within a eukaryotic cell) through the transcription-translation pathway. Also, by virtue of DNAs being activated by proteins through chromatinic binding, the whole process of protein synthesis eventually becomes a cyclic process. Through our analytical calculation on the theoretical model, we gain confidence about its sustainability and by numerical simulation of model equations, we obtain precise characteristic features of the model (as well as its solutions) in detail. We find that many parameters of the model could be significantly used to regulate protein synthesis in varied ways as per necessity. Through detailed observations on numerical data, certain qualitative predictions are made, which could be put to experimental verification.

Further work relating to the present area of protein synthesis, would include, finding out precise biological analogue of various parameters used in the model and obtaining direct numerical correspondence of specific bio-molecular activities with each of the model parameters. We would also endeavour further to explore, in detail, the influence of various bio-molecular and bio-chemical organelles in regulating the process of protein synthesis through transcription-translation pathway. Efforts would also be made to incorporate the effects of DNA super coiling and protein folding within our model.

## Acknowledgements:

One of the authors, Biplab Chattopadhyay, would like to thank UGC, INDIA for motivating assistance in the form of a recommended Research Project (No. PSW-270/15-16 (ERO), Date:19-Jan-2017) for which , however, the financial assistance remains to be released yet.

## References:

[1]. Michael M. Cox & David L. Nelson, Lehninger's Principles of Biochemistry, 5th edition (2008), W.H. Freeman & Company
[2]. Lodish H., Berk A., Zipursky SL., et. al., Molecular Cell Biology, 4th edition, (2000), W. H. Freeman & Company
[3]. Haurowitz, The Chemistry & Function of Proteins, 2nd edition, (1963), Academic Press
[4]. Gil Alterovitz, Ehsan Afkhami & Marco Ramoni, Robotics, Automation & Statistical Learning for Proteomics, Nova Science Publishers, 2005
[5]. D. Del Vecchio & E. D. Sontag, Dynamics & Control of Synthetic Bio-molecular Networks, American Control Conference, 2007, Acc 07, July 9-13, 2007, pp.1577-1588
[6]. C. Hoog & M. M., Proteomics, Annual Review of Genomics & human Genetics, Vol.5, 2004, pp.267-293
[7]. Kim Sneppen & Giovanni Zocchi, Physics in Molecular Biology, 2005, Cambridge University Press
[8]. Morrison & Boyd, Organic Chemistry, 6th edition, Prentice Hall of India Private Limited
[9]. B. Lewin, Genes VIII, International edition, (2004), Pearson Prentice Hall
[10]. Bruce Alberts, Alexander Johnson, Julian Lewis, Martin Raff, Keith Roberts & Peter Walter, Molecular Biology Of The Cell, 4th edition (2002), U.S. National Library of Medicine
[11]. Richard J Dimelow & Stephen J Wilkinson, Control of translation initiation: a model-based analysis from limited experimental data, J. R. Soc. Interface, Vol.6, No.30, 2009, pp. 51-61
[12]. N. S. Bar & D. R. Morris, Dynamic model of the process of protein synthesis in eukaryotic cells, Bulletin of Mathematical Biology, Vol. 69, 2007, pp. 369-393
[13]. Kozak M., Regulation of Translation in Eukaryotic systems, Annual Review Cell Biology, Vol.8, 1992, pp. 197-225
[14]. Nadav S Bar & Rahmi Lale, Modeling & control of the protein synthesis process in eukaryotic cells, 47 th IEEE Conference on decision & control, Cancum, Mexico, Dec. 9-11,2008
[15]. G. M. Cooper & R. E. Hausman, The Cell: A Molecular Approach, 6th edition, Sinauer Associates, Inc. (2013)
[16]. Robert J. Robbins, Molecular Biology Fundamentals, John's Hopkins University, 1994
[17]. P. Schimmel & K. Ewalt, Genetic Code, (2008), Mc. Graw-Hill Companies
[18]. Devarajan Thirumalai & Jie Chen, Role of RNA-polymerase in Gene Transcription Demonstrated, Science Daily, July 15,2010
[19]. A. J. Griffiths, An Introduction to Genetic Analysis, 7th edition, (2000), W. H. Freeman, New York
[20]. V. I. Lim & J. F. Curran, Analysis of codon:anticodon interactions within the ribosome provides new insights into codon reading & the genetic code structure, RNA ( a publication of RNA society), Vol.7, 2001, pp.942-957
[21]. F. H. C. Crick, Codon-Anticodon Pairing: The Wobble Hypothesis, Journal of Molecular Biology, Vol.19, 1966, pp.548-555
[22]. Dreyfuss G, Kim VN, Kataoka N., Messenger-RNA-binding proteins & the messages they carry, Nature Reviews, Molecular Cell Biology, Vol.3,Mar.3,2002, pp.195-205
[23]. Valery I. Lim & Ceslovas Venclovas, Codon-anticodon pairing: A model for interacting codon-anticodon duplexes located at the ribosomal A- and P- sites, Federation of European Biochemical Societies, Vol.313, No.2,1992, pp.133-137
[24]. Brener S, Jacob F, Meselson M., An unstable intermediate carrying information from genes to ribosomes for protein synthesis, Nature, Vol.190, May 13, 1961, pp.576-581
[25]. Noller, H.F., RNA structure: Reading the Ribosome, Science, Vol.309, 2005, pp.1508-1514

[26]. V. M. Pain, Initiation of protein synthesis in eukaryotic cells, European J. Biochem. Vol. 236, 1996, pp 747-771
[27]. L. McKendrick, V. M. Pain, S.J.Morley, Translation initiation factor 4E, The International Journal of Biochemistry & Cell Biology, Vol.31, Jan.1991, pp. 31-35
[28]. I. L. Finar, Organic Chemistry, Vol.-2, 5th edition, (2004), Pearson Education Pte. Ltd.
[29]. Christian B. Anfinsen, Principles that govern the folding of protein chains, Science, new series, Vol.181, No.4096, July 20, 1973, pp.223-230
[30]. F. Gebauer and M. W. Hentze, Molecular mechanisms of translational control, Nature Rev. Mol. Cell Biol. Vol.5**,** 2004, pp 827-835
[31]. F. Jacob & J. Monod, Genetic regulatory mechanisms in the synthesis of proteins, Journal of Molecular Biology, Vol.3, 1961, 318-369
[32]. E. Renshaw, Modeling Biological populations in Space & Time, Cambridge University Press, 1991
[33]. Darren J Wilkinson, Stochastic modeling for quantitative description of heterogeneous biological systems, Nature Reviews Genetics, Vol.10, Feb. 2009, pp.122-133
[34]. M. J. Moore, From birth to death: The complex lives of eukaryotic mRNAs, Science, Vol.309**,** 2005, pp 1514-1518
[35]. Weijun Dai, Gen Zhang & Eugene V Makeyev, RNA-binding protein HuR autoregulates its expression by promoting alternative polyadenylation site usage, Nucleic Acids Research, Vol.40, No.2, 2012, pp. 787-800
[36]. H. Kitano, Computational systems biology, Nature, Vol.420, 2002, pp.206-210
[37]. E. D. Sontag, Molecular systems biology & control, Eur. J. Control, Vol.11(4-5), 2005, pp.396-435
[38]. Mark Kot, Elements of Mathematical Ecology, Department of Applied Mathematics, University of Washington, Cambridge University Press, 2001
[39]. B. Chattopadhyay & N. Hui, Immunopathogenesis in Psoriasis through a Density-type Mathematical Model, WSEAS TRANSACTIONS on MATHEMATICS, Vol.11, 2012, pp.440-450
[40]. Andrew B. Huhges, Amino acids, Peptides & Proteins in Organic Chemistry. Vol.5: Analysis & Function of Amino acids & Peptides, 1st edition, (2012), Wiley-VCH Verlag Gmbh & Co. KGaA
[41]. Y. C. Fung, Biomechanics, (1981), Springer New York.

"Tumpa Saha & Biplab Chattopadhyay" Studies on Protein Synthesis within a Theoretical Model" IOSR Journal of Pharmacy and Biological Sciences (IOSR-JPBS) 14.5 (2019): 10-20.