Nadav Skjøndal-Bar [a,1] David R.Morris [b]

[a]*Norwegian University of Science and Technology, Department for engineering cybernetics, Trondheim, Norway*

[b]*University of Washington, Biochemistry Department, Seattle, USA*

Dynamic Model of the Process of Protein Synthesis in Eukaryotic Cells

Abstract

Protein synthesis is the final step of gene expression in all cells. In order to understand the regulation of this process, it is important to have an accurate model that incorporates the regulatory steps. The model presented in this paper is composed of set of differential equations which describe the dynamics of the initiation process and its control, as well as peptide elongation, starting with the amino acids available for peptide creation. A novel approach for modeling the elongation process permits useful prediction of protein production and consumption of energy and amino acids, as well as ribosome loading rate and ribosome spacing on the mRNA.

Key words: Protein synthesis, Initiation factors, Initiation control, elongation, eIF2, eIF4, charged-tRNA, amino acid consumption, differential equations of protein synthesis, dynamic model

1 Introduction

Translation of gene transcripts into protein is the final step of gene expression and control of this process is a key aspect of the regulation of gene expression (Hinnebusch, 2000). High throughput analysis of ribosome loading onto the individual mRNAs of the transcriptome of *Saccharomyces cerevisiae* has opened the way to investigations of translational control at the genome-wide level (Arava et al., 2003; MacKay et al., 2004). In order to maximize the information extracted from these genome-level experiments, a dynamic model of the protein synthesis process and its regulation is required. Such a model,

Email addresses: `barc@itk.ntnu.no` (Nadav Skjøndal-Bar),
`dmorris@u.washington.edu` (David R.Morris).
[1] Present address: Univ. of Washington, Biochemistry Department

which incorporates current understanding of translational control, also will allow rigorous tests of our mechanistic concepts of the important process of protein synthesis. This paper describes a dynamic model, which incorporates the individual mechanistic steps of translation: initiation, elongation and termination. This model also incorporates two key regulators of protein synthesis: the phosphorylation of eukaryotic initiation factor-2 (eIF-2) and the interaction of eukaryotic initiation factor-4E (eIF-4E) with the binding proteins that control its activity (Hinnebusch, 2004; Holcik and Sonenberg, 2005).

2 Process description

The process of protein synthesis is divided in this paper into a few sub-processes:

(1) Initiation and its controllers
(2) Elongation
(3) Reactions between tRNA and amino acids

where a fast control mechanism of initiation, performed by eIF2 and eIF2B (Hinnebusch, 2000; Trachsel, 1996), regulates the amount of 40S ribosome loaded at the 5' end of the mRNA. A second control mechanism is mediated through regulation of eIF4F. Elongation is modelled as the sum of forces affecting the motion of the ribosome as physical body. To fuel elongation, tRNAs react with amino acids to create charged-tRNAs. An overview of the initiation process and its control is shown in figure 1.
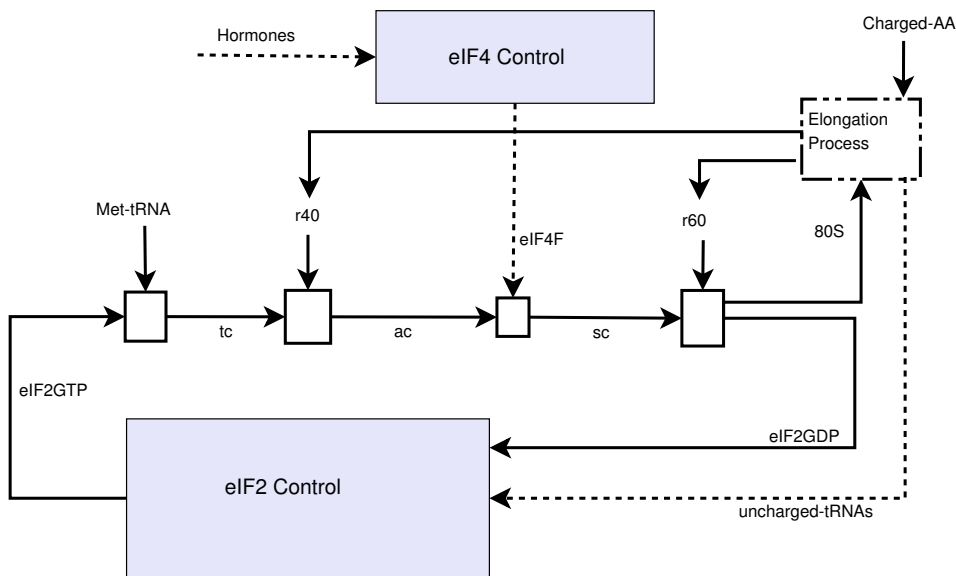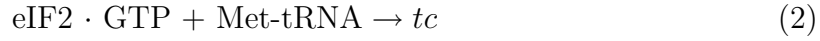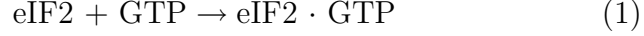

Fig. 1. Overview of the initiation process.

2.1 Initiation description

The process of translation initiation is a series of reactions that end up in identification of the initiation codon on mRNA. The initiation process is quite complex, since it involves different levels of control. It starts with the formation of the Ternary Complex as given by the following reactions:

$$\text{eIF2} + \text{GTP} \rightarrow \text{eIF2} \cdot \text{GTP} \tag{1}$$
$$\text{eIF2} \cdot \text{GTP} + \text{Met-tRNA} \rightarrow tc \tag{2}$$

where tc is the Ternary Complex (eIF2·GTP· Met-tRNA). The initiation factor eIF2 forms a binary complex with GTP but not with Met-tRNA, thus it is reasonable to assume that an eIF2·GTP binary complex is formed initially and Met-tRNA is bound subsequently (Trachsel, 1996). tc is joined then with 40S ribosome (denoted as r_{40}) to create an active site complex (ac). This pre-initiation complex is directed to the m^{7}G cap at the 5' end of the mRNA, through interaction between eIF4F (denoted E_F) and the 40S ribosome (mediated by eIF3), where it becomes a scanning complex, sc. This step has a complex control mechanism where the regulation of eIF4F dominates and will be described later. The process can be summarized as

$$tc + r_{40} \xrightarrow{k_{51}} ac \tag{3}$$
$$ac \xrightarrow{k_{52}, \ E_F} sc \tag{4}$$

where k_{51} and k_{52} are reaction rate constants of ac and sc formation, respectively. These above reactions can be described by the next set of equations:

$$\dot{tc} = - k_{51} \ tc \cdot r_{40} + k_4 \ e_T \cdot x_M \tag{5}$$
$$\dot{ac} = k_{51} \ tc \cdot r_{40} - k_{52} \ ac \cdot E_F \tag{6}$$
$$\dot{r}_{40} = - k_{51} \ tc \cdot r_{40} + \dot{r}_{40}|_{el} \tag{7}$$

where $\dot{r}_{40}|_{el}$ represents change rate of the 40S subunit during the elongation process (will be described later). During initiation, the Scanning Complex sc leaves the 5' terminal cap and moves along the 5' Untranslated Region (UTR) until the AUG codon is encountered. This UTR might be unstructured in which sc migrates along easily without encountering hinders or energy barriers. The 5' UTR might however contain barriers, for example secondary structures, which pose a energy barrier to the scanning complex sc and the scanning process is slowed down, or might even stop. The efficiency of the scanning process depends on structural features of the 5' transcript leader (Kozak, 1991), as illustrated in figure 2. The differences in time that it takes to move from the 5' terminal cap to the initiation codon can be modeled in several ways. In this paper, we assume a constant rate that is unique to each species of transcript and which can be easily implemented by introducing a

3

multiplicative constant ρ_u^{-1}, $\rho_u \geq 1$ to sc rate of change, thus delaying the initiation for mRNAs when the 5' UTR contains inhibitory structures. ρ_u can be intrepreted as the resistivity of the UTR to motion of sc on it. Table 1 presents
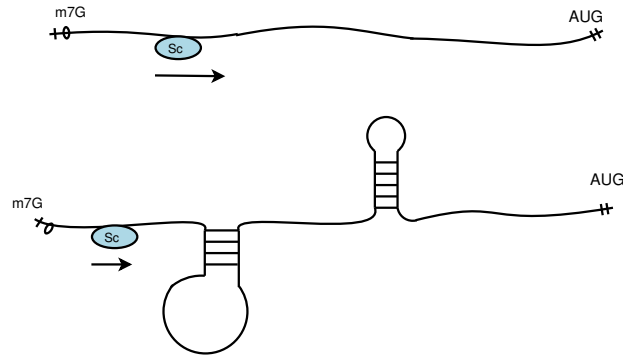

Fig. 2. Two different UTRs and sc moving on them. Top UTR is a simple strand, where $\rho_u \approx 1$. Bottom UTR has secondary complex which pose energy barriers, for which $\rho_u \gg 1$.

the variables and parameters used to described the dynamics of initiation and elongation.

Variable/par.	Size	Description
x_{aa}	$\mathbb{R}^{1x20}$	Amino acids concentration
x_u	$\mathbb{R}^{1x20}$	Uncharged tRNA
x_{ct}	$\mathbb{R}^{1x20}$	Charged tRNA
x_m	scalar	Methionyl-tRNA
tc	scalar	Ternary Complex $eIF2 \cdot GTP \cdot Met\text{-}tRNA_i$
ac	scalar	Active complex
sc	scalar	Scanning complex $40S \cdot ac$
r_{40}	scalar	Ribosome 40S
r_{60}	scalar	Ribosome 60S
r_{80}	scalar	Ribosome 80S
ρ_u	scalar	Resistivity of the UTR

Table 1
Variables in the translation dynamic model.

When the AUG codon is finally recognized, the 60S ribosome is joined to the Scanning Complex sc in the following way

$$sc + r_{60} \xrightarrow{k_6;\ \rho_u} r_{80} + e_D \tag{8}$$

where k_6 is the rate constant and e_D denotes concentration of eIF2·GDP. From the process given by (8), changes of Scanning Complex sc, r_{60} and r_{80} can be

4

described by

$$\dot{sc}(t) = k_{52}ac(t) \cdot E_F(t) - k_6 \cdot \frac{1}{\rho_u} \cdot sc(t) \cdot r_{60}(t) \tag{9}$$

$$\dot{r}_{60}(t) = -k_6 \ \rho_u^{-1} \ sc(t) \cdot r_{60}(t) + \dot{r}_{60}|_{el} \tag{10}$$

$$\dot{r}_{80}(t) = k_6 \frac{1}{\rho_u} sc(t) \cdot r_{60}(t) + \dot{r}_{80}|_{el} \tag{11}$$

where $\dot{r}_{60}|_{el}$ and $\dot{r}_{80}|_{el}$ describe changes in 60S and 80S subunits, respectively, due to elongation process. When secondary structures are present, resistivity is increased, which slows down ribosome movement and reduces the translation initiation rate. It is important to note that ρ_u is a characteristic property, thus a constant for a given mRNA.

2.2 Initiation Control

2.2.1 eIF2 control

Since eIF2·GTP is a precursor of the initiation process (see figure 1), reducing its amount will reduce translation rate. This section expands the eIF2 controller in figure 1 and describes the main idea behind the controller mechanism. The variables and parameters are described in table 2 and the controller mechanism is illustrated in figure 3. Recycling of eIF2 from the complex eIF2·GDP

Variable/parameter	Description
e_2	eIF2
e_B	eIF2B
e_p	eIF2$_p$ (Phosphorylated eIF2)
D_2	inactive complex eIF2$_p$·GDP·eIF2B
e_D	eIF2·GDP complex
e_T	eIF2·GTP complex
e_{pD}	eIF2$_p$·GDP complex
e_{com}	eIF2·GDP·eIF2B complex
G_2	GCN2
tr	tRNA
G_t	GCN2·tRNA
$k_{71}, \ k_{72}$	Rate constants for inactive complex
$k_{81} \ k_{82}$	Rate constants for GCN$_2$ reactions
$k_{11} \ k_{12}$	Rate constants for eIF2 phosphorylation and the reverse reaction

Table 2
Variables and parameters in eIF2 controller.

is modelled as formation of an intermediate eIF2·GDP·eIF2B which is broken

down quickly to its individual components in the following manner

$$e_D + e_B \xrightarrow{k_{71}} e_{com} \tag{12}$$

$$e_{com} + \text{GTP} \xrightarrow{k_{72}} e_T + e_B + \text{GDP} \tag{13}$$
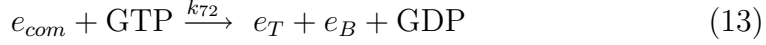
thus the initiation factor eIF2B is required in order to recycle eIF2·GDP, which is a prerequisite to the scanning process.

Controlling the amounts of eIF2B in the cell will regulate the recycling rate of eIF2·GDP to eIF2·GTP, thus promoting or inhibiting the intitation and formation of 80S ribosome. Concentrations of eIF2B can be regulated by trapping it with eIF2_p·GDP, since eIF2B has at least 150-fold greater affinity to eIF2_p·GDP than to eIF2·GDP (Rowlands et al., 1988). This creates an inactive complex in the following manner

$$e_{pD} + e_B \quad \underset{k_{22}}{\overset{k_{21}}{\rightleftharpoons}} \quad D_2 \tag{14}$$

where D_2 is a dummy (inactive) intermediate of the form eIF2_p·GDP·eIF2B. The parameters k_{21} and k_{22} should be chosen such that the forward reaction will be favorable and the reverse reaction is at constant rate, $k_{21} \geq k_{22} \; \forall t$. This way, any excess of e_p will immediately react with e_B to inhibit initiation. The total changes in e_B and D_2 are then given by

$$\frac{d(e_B)}{dt} = -k_{21} \cdot e_B \cdot e_{pD} + k_{22} \cdot D_2 - k_{71} \cdot e_D \cdot e_B + k_{72} \cdot e_{com} \; x_T$$
$$= k_{22} \cdot D_2 + k_{72} \cdot e_{com} \; x_T - e_B \cdot (k_{21} \; e_{pD} + k_{71} \cdot e_D) \tag{15}$$

$$\frac{dD_2}{dt} = -k_{21} \; e_B \; e_{pD} + k_{22} \cdot D_2 \tag{16}$$

Formation of e_p is catalyzed by GCN2·tRNA (denoted as G_t), which is created by the next process:

$$x_{ut} + G_2 \quad \underset{k_{82}}{\overset{k_{81}}{\rightleftharpoons}} \quad G_t \tag{17}$$

where x_{ut} is the concentration of uncharged-tRNA, G_2 is GCN2, and G_t is GCN2·tRNA. When amino acid levels drop in the cell, uncharged-tRNA increases, which binds to GCN2 to form the active enzyme GCN2·tRNA. The expression $k_{81} \geq k_{82}$ should hold in order to account for fast changes of the controller in the case of rapid depletion of charged-tRNAs. G_t catalyses phosphorylation of both eIF2 and eIF2·GDP (denoted as e_D), where phosphorylation of eIF2 and the reverse reaction are described as follows:

$$e_2 + \text{ATP} \xrightarrow[G_t]{k_{11}} e_p + \text{ADP} \tag{18}$$

$$e_p + \text{H}_2\text{O} \xrightarrow{k_{12}} e_2 + \text{P}_i \tag{19}$$

$$x_D + e_p \xrightarrow{k_{10}} e_{pD} \tag{20}$$

6

Assuming ATP concentration is abundant in the cell, the change of e_2 due to phosphorylation in the processes described by (18)-(19) is given by

$$\left.\frac{e_2}{dt}\right|_{phos} = k_{12} \ e_p - k_{11} \ e_2 \ G_t \tag{21}$$
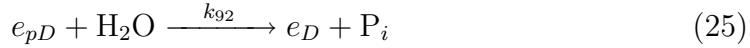
where the reaction rate k_{12} is constant. k_{11} can be chosen either as constant or a function of the enzyme GCN2tRNA. The damping term $k_{11}e_2G_t$ is increasing proportionaly to the levels of G_t, hence increasing phosphorylation rate. The total change in e_2 due to phosphorylation and reaction with GTP, is thus given by the next equation

$$\begin{aligned}
\dot{e}_2 &= \left.\frac{d(e_2)}{dt}\right|_{phos} - k_3 \cdot x_T \cdot e_2 \\
&= k_{12}e_p - e_2(k_3 x_T + k_{11}G_t)
\end{aligned} \tag{22}$$

while the change in e_p is described by

$$\dot{e}_p = -\dot{e}_2|_{phos} - k_{10}x_D e_p = k_{11}e_2 G_t - e_p(k_{12} + k_{10}x_D) \tag{23}$$

The phosphorylation of e_D and the reverse reaction are as follows:

$$e_D + \text{ATP} \xrightarrow[G_t]{k_{91}} e_{pD} + \text{ADP} \tag{24}$$

$$e_{pD} + \text{H}_2\text{O} \xrightarrow{k_{92}} e_D + \text{P}_i \tag{25}$$

where the rate of the reverse reaction depends on G_t. Since we assume that the supply of H_2O is abundant, we ignore this term so e_p and e_{pD} are converted back to e_2 and e_D, respectively, at a constant rate. Change in e_D is thus found as

$$\begin{aligned}
\dot{e}_D &= (- \text{ recycling}) - (\text{phosph.}) + (\text{de-phosph.}) \\
&= -k_{71} \ e_B \ e_D - k_{91} \ G_t \ e_D + k_{92} \ e_{pD} \\
&= k_{92} \ e_{pD} - e_D \cdot (k_{71}e_B + k_{91}G_t)
\end{aligned} \tag{26}$$

where the phosphorylation (see process in (24)) depends on the G_t concentrations while de-phosphorylation is done at a constant rate k_{92}. Therefore, increase in G_t will increase the damping term $k_{71}e_B + k_{91}G_t$, reducing rate of e_D. Changes in e_{pD} can be described by the next equation

$$\begin{aligned}
\dot{e}_{pD} &= -k_{21} \ e_B \ e_{pD} + k_{22} \ D_2 + k_{91}G_t \ e_D + k_{10} \ g_d \ e_p - k_{92}e_{pD} \\
&= k_{22}D_2 + k_{91}G_t \ e_D + k_{10} \ g_d \ e_p - e_{pD}(k_{21}e_B + k_{92})
\end{aligned} \tag{27}$$

This process can be summarized in figure 3. The internal grey box highlights the control part which regulates concentrations of e_B.
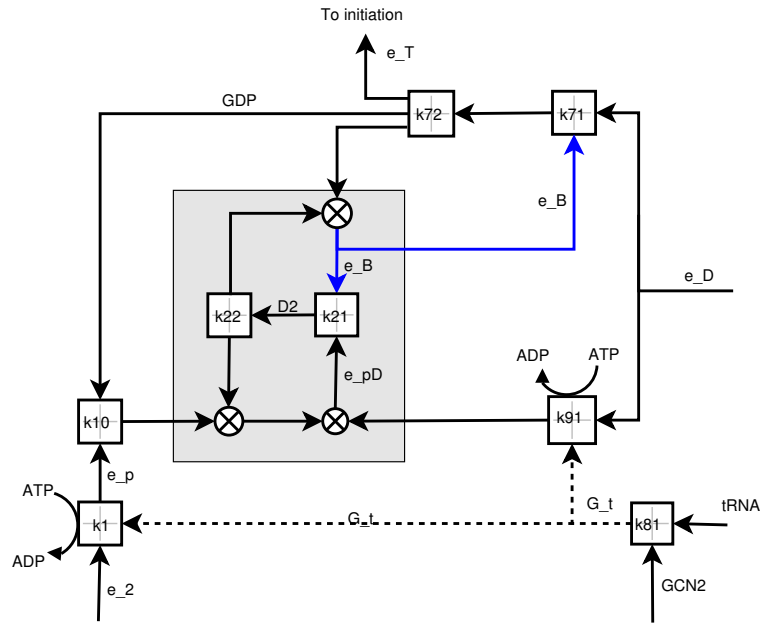
Fig. 3. eIF2 initiation control model. e_D results from the initiation process and is being recycled to e_T in a controlled manner. Reactions involve k_{12} and k_{92} are not shown here.

2.2.2 eIF4 control

Since eIF4F is a prerequisite for the preinitiation complex to load onto the mRNA, inhibiting its activity will prevent further loading of 40S ribosomes onto the m^7G cap, thus decreasing the ribosome loading rate. Figure 4 shows the pathways and dependencies of the controller while table 3 presents the variables participating in eIF4 control. Control of eIF4E concentration is me-

Variable/parameter	Description
E_4	eIF4E
E_{BP}	eIF4E-BP
E_p	eIF4E-BP$_p$ (Phosphorylated eIF4E-BP)
D_4	inactive complex
E_G	eIF4G
E_F	eIF4F
H	Stimulating hormone signal
k_{411}	Rate constants for dephosphorylation of E_p
k_{421} k_{422}	Rate constants for E_F formation and breakdown
k_{43}	Rate constants for inactive complex formation

Table 3
Variables and parameters in eIF4 controller.

diated through formation of a complex with eIF4E-BP to form an inhibited complex eIF4E·eIF4E-BP, described by the next reaction

$$E_{BP} + E_4 \xrightarrow{k_{43}} D_4 \tag{28}$$

8

External stimuli enhance phosphorylation of the eIF4E-BPs (Holcik and Sonenberg, 2005), resulting in eIF4E-BP$_p$ and breaking the dummy complex D$_4$ to phosphorylated eIF4E-BP and free eIF4E. These signals thus allow increase in free eIF4E which then directs the eIF4G to the m^{7}G cap. The eIF4E-BPs inhibit translation by binding to eIF4E (step 1 in figure 4) to prevent the association between eIF4E and eIF4G, thus blocking the assembly of a functional eIF4F complex (Raught et al., 2000). The main reaction can be summarized by the following reactions:

$$E_{BP} + \text{ATP} \xrightarrow{H} E_p + \text{ADP} \tag{29}$$

$$E_p \xrightarrow{k_{411}} E_{BP} + P_i \tag{30}$$

$$E_4 + E_G \underset{k_{422}}{\overset{k_{421}}{\rightleftarrows}} E_F \tag{31}$$

Breakdown of D_4 mediated by phosphorylation of the eIF4E-BP in the D_4 complex, thus resulting in two products, the phosphorylated eIF4E-BP$_p$ and free eIF4E, in the following manner:

$$D_4 + \text{ATP} \xrightarrow{H} E_4 + E_p + \text{ADP} \tag{32}$$

This reaction is considered here as an irreversible process. Step 3 in the process is the association of free eIF4E with eIF4G to form eIF4F (denoted as E_F) in a reversible reaction and enabling translation. Formation of E_F occurs rapidly, while the reverse reaction depends on an external signal H, such that increase in the signal decrease the rate of the reverse reaction.

Controlling the Active Complex (ac) on the mRNA m^{7}G cap can be regarded as a relative long term translation control; removing ac prevents further access of ribosomes to the initiation AUG codon. The extracellular stimuli that affect the level of phosphorylation of eIF4E-BP and the regulation of the initiation process are discussed in Gingras et al. (1999) and Raught et al. (2000).

Assume an external signal (hormones, nutrients, etc.) that regulates intracellular protein synthesis. Denote stimulating hormone levels at some time instance as h(t), and divide it by the hormone level capacity, H_{cap}, the normalized hormonal level H becomes

$$H = \frac{h(t)}{H_{cap}}, \quad \text{where} \quad H_{cap} \geq h(t) \ \forall \ t \tag{33}$$

with $H \in [0, 1]$ where low or high hormone levels result in H close to zero or one, respectively. Step 1 in figure 4 is the formation of an inactive complex which depletes the amount of free eIF4E, thus arresting the translation process. Step 2 breaks the inactive complex to phosphorylated eIF4E-BP complex and free eIF4E. Denote the variables eIF4E, eIF4G, eIF4F and the complexes
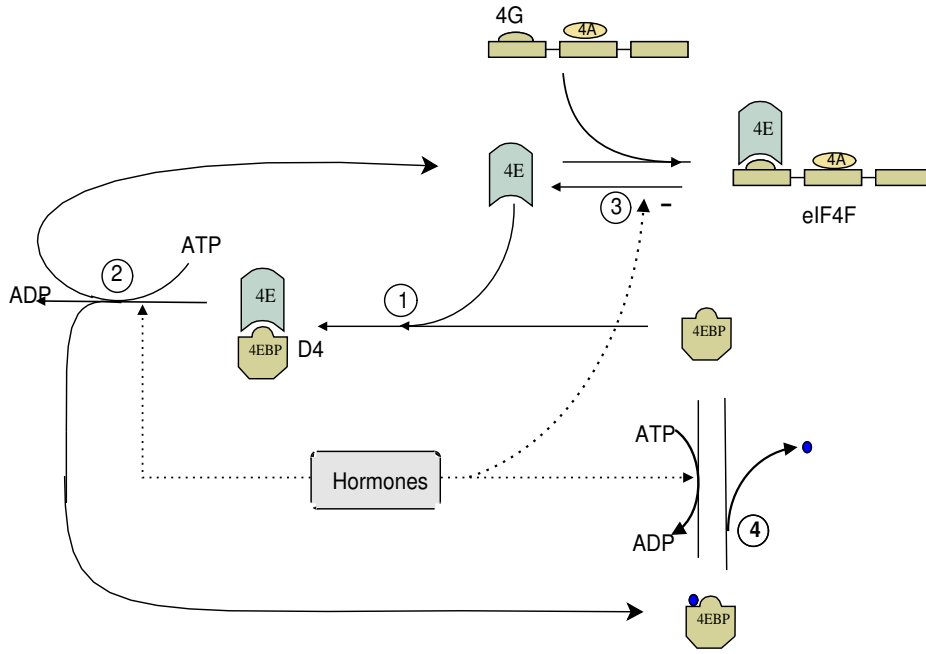
Fig. 4. Initiation control eIF4 configuration. Step 1 is inhibition of free eIF4E to D_4 complex. Hormonal signals in step 2 stimulate dissociation of D_4 complex using ATP thus promoting translation. Step 3 is the association of the eIF4E with eIF4G and the reverse reaction (which is negatively effected by the H signal). Step 4 is phosphorylation of eIF4E-BPs, stimulated by hormonal signals and its reverse reaction.

eIF4E-BP and eIF4E-BP·eIF4E as E_4, E_G, E_F, E_{BP} and D_4, respectively, then changes in E_4 concentrations can be described by

$$\dot{E}_4 = k_{422}E_F + H \cdot D_4 - E_4 \left[k_{43}E_{BP} + k_{421}E_G\right] \tag{34}$$

Since translation control is required, it is essential to include the reverse reaction, i.e. $k_{422} \neq 0$ is the rate of E_F breakdown. The hormonal signals have a negative effect on the breakdown rate of E_F such that when the signals are oriented toward synthesis, the reverse reaction is slowed down at rate

$$k_{422} = c_1 \cdot (1 - H) \tag{35}$$

where c_1 is a constant and $0 \leq H \leq 1$, which result in stable E_F structure on the mRNA cap. Changes of phosphorylated E_4 can be described by the following equation

$$\dot{E}_p = H \cdot [E_{BP} + D_4] - k_{411}E_p \tag{36}$$

therefore increasing the value of H will increase the change of E_p, preventing decrease of free E_4 due to inhibition of step 2. The change rates of the

remaining metabolics are given by

$$\dot{D}_4 = k_{43} \cdot E_{BP}E_4 - H \cdot D_4 \tag{37}$$

$$\dot{E}_{BP} = k_{411}E_p - E_{BP} \cdot [H + k_{43}E_4] \tag{38}$$

$$\dot{E}_G = k_{422}E_F - k_{421}E_4E_G \tag{39}$$

$$\dot{E}_F = -\dot{E}_G \tag{40}$$

where the rate k_{422} is given by (35).

2.3 Control of ribosome spacing

Loading rate of the ribosomes on the mRNA has a physical limitation which need to be accounted for. Denote the physical width of each ribosome on an mRNA as $L_{r_{80}}$ (in units of *codons*), then the maximum number of ribosomes that can be loaded on a given mRNA due to space limitation is $L_{mRNA}/L_{r_{80}}$. Let $C_p = L_{r_{80}}^{-1}$ represents the constant capacity. Then the space limitation is modelled continousely by using the first order filter

$$S_p(t) = \frac{C_p - d(t)}{C_p + d(t)} \tag{41}$$

where $d(t)$ is the spacing, or linear density of the ribosomes on the mRNA,

$$d(t) = \frac{r_{80}(t)}{L_{mRNA}} \tag{42}$$

in units of $[ribosomes/codons]$. The variable $S_p(t)$ reduces the loading rate as the density increases. The larger $d(t)$ is, the more difficult it is to load a ribosome, and $S_p(t)$ lies inside the interval $[0, 1]$ where

$$\lim_{d \to C_p} S_p = 0 \tag{43}$$

S_p is then cooperated in the model by affecting the rate of loading 40S on the terminal cap as $k_{51}^* = k_{51}k_p \cdot S_p$ where k_p is some constant.

2.4 Elongation Process

After recognizing the initiation codon AUG, the 60S ribosome is joined to *sc* and 80S ribosome begins to polymerize amino acids and progress towards 3' end of the mRNA. As the 80S ribosome moves codon-by-codon along the mRNA strand, charged tRNA provides the amino acid in the ribosomal A site

11

using the ternary complex eEF1A·GTP·AAtRNA, which binds in a codon-dependent manner (Merrick and Nyborg, 2000). In the process, an amino acid is added to the peptide chain, the uncharged tRNA is released and the 80S ribosome advances one codon. This is done repeatedly until the termination codon is reached, as long as the supply of AA-tRNA and energy are not depleted. The motion of the 80S ribosome depends on the supply of energy and AA-tRNAs. When the 80S ribosome reaches the termination codon, it releases the peptide, the 40S and 60S subunits. The process of elongation can be described by the next set of reactions:

$$f(AAtRNA) + r_{80} + g(GTP) \xrightarrow{elongation} f(tRNA) + r_{40} + r_{60} + g(GDP) \quad (44)$$
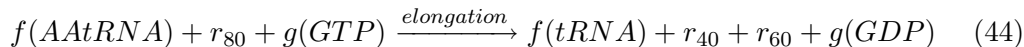
where the functions $f(\cdot)$ and $g(\cdot)$ depend on the mRNA sequence, the motion of the ribosome and the loading rate. In order to model the dynamics of this process, we will consider the 80S ribosome as a charged particle, with a continous motion between each two codons along a mRNA strand, subjected to continuous forces as presented in figure 5. The motion of the ribosome is
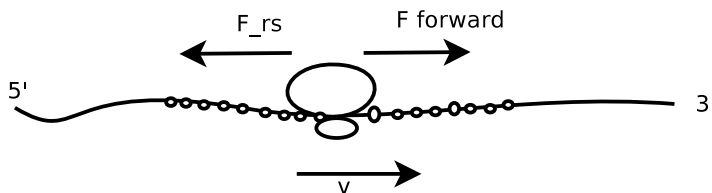

Fig. 5. Forces applied on the ribosome 80S during elongation

given from Newton's second law of motion

$$m\bar{a} = \sum F = F_{forward} - F_{rs} \quad (45)$$

where $m = \gamma$ is the mass of the ribosome, $\bar{a}$ is the acceleration and $F_{forward}$ is the force pulling the ribosome toward the 3' terminus while F_{rs} is the force resists the motion toward the 5' end. Since attachment of amino acid to the peptide is done at each codon and requires energy in the form of hydrolysis of GTP to GDP, each amino acid requires about $30Kj/mol$ (Voet, 2004). When considering number of GTP and GDP molecules (not concentrations), the amount of energy needed is $30 \cdot 10^3/Avg$ $[Joul/codon]$ where Avg is Avogadro number. Following this, the energy required to translocate a ribosome between two codons is

$$U = \frac{30 \cdot 10^3}{Avg} \quad [Joul] \quad (46)$$

The mechanisms that trigger conformational changes in the ribosome structure to create and control motion are unknown today. Recognizing this, we developed a strategy in the context of the model to implement ribosome movement. Assume a constant electrical field E exists between each two codons. Assume further that the ribosome has a net charge different than zero. Then

the force pulling the ribosome forward can be described as

$$F_{forward}(t) = q(t) \cdot E \tag{47}$$

where $q(t)$ is the charge of the ribosome at time instance t. The forces act on the ribosome against the movement direction 5' $\to$ 3' are combined from several factors. There is a resistive type of damping force that is proportional to the velocity v since the faster the ribosome moves, the more difficult it gets to mobilize the appropriate charged tRNA to the A site. The force which resist the movement is given by

$$F_{rs} = \beta \cdot v(t) \tag{48}$$

where $v(t)$ is the velocity (in units of *codons/s*) of the ribosome along the strand and

$$\beta = \frac{\gamma^2}{\alpha E}$$

defines the proportional damping constant where $1/\alpha$ is the damping coefficient and γ is the mass of the ribosome 80S.

Using equation (45), the acceleration can be described as

$$\gamma \frac{d(v(t))}{dt} = q(t)E - \beta \cdot v(t) \tag{49}$$

Solving this differential equation is done by separation of variables in (49) and integration

$$\int_0^v \frac{dv}{\frac{qE}{\gamma} - \frac{\beta v}{\gamma}} = \int_0^t dt \tag{50}$$

and the solution is given as

$$v = v_m \left[1 - \exp(-\frac{\beta}{\gamma} t) \right] \tag{51}$$

where

$$v_m = \frac{qE}{\beta} \tag{52}$$

is the velocity when $t \gg \gamma/\beta$. Assume that there exists time $t_a \approx \gamma/\beta$ such that

$$\bar{v} = 0.62 \frac{qE}{\beta_a} = \frac{\alpha_a E^2}{\gamma^2} q \tag{53}$$

is an average velocity when the ribosome is between two codons. It is also reasonable to assume that the mass of the 80S ribosome does not change during elongation since it is very large comparing to the peptide and the tRNA. Then

$$\sigma = \frac{\alpha_a E^2}{\gamma^2} \tag{54}$$

13

is a constant, depend on the mRNA itself, and not on the charge. We can define resistivity then as

$$\rho = \frac{1}{\sigma} \tag{55}$$

Resistivity is a characteristic of the specific mRNA, and not of the ribosome itself since γ is independent of the gene. It depends on the codon sequence of the mRNA. Rare codons will present difficulty in acquiring the correct charged-tRNA, whilst common codons use more abundant charged-tRNAs, thus contributing to faster elongation and motion. The 3D structure of the strand might also be an important factor, yet it is difficult to demonstrate how it affects the motion or the forces, since there is no experimental evidence relating to this assumption. If the mRNA strand is involved in a complex structure, it might present resistance to the motion toward the 3' cap, increasing α.

The position of the ribosome is then

$$pos(t) = \int_{t_0}^{t} v(\tau)d\tau \tag{56}$$

where $pos(t)$ is the position (in codons) on the mRNA starting at the AUG codon. The 80S ribosome charge $q(t)$ is effected by the amount of charged tRNA available for the next codon. If no charged-tRNA is presented to the A-site, then the ribosome will not move to the next codon sequence and the elongation stops. In this model, reducing the value of charge $q(t)$ to zero will drive the force in (49) to zero as well, rendering the acceleration to a negative value, thus reducing velocity towards zero. An example of changing the $q(t)$ dynamic is by using the following expression

$$\dot{q}(t) = \begin{cases} C_q & \text{if } \frac{x_{ct}^i}{x_{req}^i} \geq 1, \ \forall i \in \{1, 2, ..20\} \\ -C_q & \text{else} \end{cases} \tag{57}$$

where C_q is a constant, hence $q(t)$ will increase linearly as long as there is enough charged tRNAs, i.e. the amount of the i-charged-tRNA is larger than the i amino acid requirement and will decrease when charged-tRNAs are depleted. $q(t)$ is saturated at max value q_{max} and has a minimum of zero, i.e. $0 \leq q(t) \leq q_{max}$. This mechanism assures that all the ribosomes on the mRNA strand will stop in case of complete amino acid starvation.

While elongating, charged-tRNAs are mobilized to the A-site, where amino acids are added to the growing polypeptide chain and the uncharged-tRNAs are released. This process is performed at each codon, and changes in i charged-tRNA, x_{ct}^i, due to elongation are described by the next equation

$$\left. \frac{dx_{ct}^i}{dt} \right|_{el} = -r_{80}(t) \cdot v(t) \cdot x_{req}^i \cdot L_{mRNA}^{-1}, \quad i = 1, 2, .., 20 \tag{58}$$

14

where x_{req}^i is the requirement for amino acid i on the specific mRNA. Since each charged-tRNA^i releases the amino acid and becomes an uncharged-tRNA^i, the rate of change of the uncharged-tRNA^i during the elongation cycle can be described as

$$\left.\frac{dx_{ut}^i}{dt}\right|_{el} = -\left.\frac{dx_{ct}^i}{dt}\right|_{el} \tag{59}$$

Note that this model assumes that the sequence of the amino acid on the mRNA is evenly distributed, and at each time instance the reaction (58) occurs *for all* i, i.e. the entire vector $\mathbf{x}_{ct}$ is reduced by a level which correspond to the amino acid requirement for this specific gene. This does not represent a real case where at each instant only one specific amino acid is being attached to a single peptide, leaving single uncharged-tRNA. However, this approach to the problem should not pose a problem on the results, only in cases where the amino acids are arranged on the mRNA in large groups of identical amino acids. This is rarely the case and will not affect the result of peptide creation in any case.

The 80S ribosome leaves the initiation codon and moves at velocity $v(t)$. At some time, say t_e, the ribosome gets to the termination codon, where the peptide is released, and 80S is broken to 40S and 60S subunits which are released to be recycled. Denote the time instances t_0 and t_e as the time if initiation of 80S ribosome and time of termination of the same ribosome, respectively, then changes of 80S ribosome on the mRNA can be described as

$$
\begin{aligned}
\dot{r}_{80}(t) &= r_{80}(loading) - r_{80}(breaking) \\
&= \begin{cases} k_6\rho_u^{-1}\left[sc(t) \cdot r_{60}(t) - \phi \cdot v(t)sc(t-\tau)r_{60}(t-\tau)\right] & \text{if } t \geq \tau \\ k_6\rho_u^{-1}sc(t) \cdot r_{60}(t) & \text{else} \end{cases}
\end{aligned} \tag{60}
$$

where ϕ is a constant, set to be the inverse of the velocity at steady state, $v(t_s)$, and τ is the time delay, or time it takes the ribosome to travel from the initiation codon to the termination one, i.e. $\tau = t_e - t_0$. This time delay is computed using equation (56) by differentiating the time where $pos(t) = 0$ (denoted as t_0 and is usually 0) with the time where $pos(t) = L_{mRNA}$ (denoted as t_e). In other words, substracting the time the first 80S subunit meets the AUG condon with the time the same subunit reaches the termination one, at position L_{mRNA}. Denote the term

$$el = \begin{cases} \frac{k_6}{\rho_u} \cdot \phi \cdot v(t) \cdot sc(t-\tau) \cdot r_{60}(t-\tau) & \text{if } t \geq \tau \\ 0 & \text{else} \end{cases} \tag{61}$$

then we can rewrite changes in r_{40} and r_{60} as

$$\dot{r}_{40} = -k_{52} \; tc \cdot r_{40} + el \qquad (62)$$

$$\dot{r}_{60} = -\frac{1}{\rho_u} k_6 \; sc \cdot r_{60} + el \qquad (63)$$

While elongating, uncharged-tRNA is released from the E site and, if free amino acids are not present, uncharged-tRNA concentration will increase. G_2 will react then with the free tRNA, resulting in the production of the enzyme G_t. This enzyme catalyzes the phosphorylation of eIF2, thus activating the eIF2 controller to change the loading rate. Since G_2 is a scalar and x_{ut} is a vector $x_{ut} \in \mathbb{R}^{20}$, only the highest value of the uncharged-tRNA is used since it will correspond the the limiting amino acid, depleted from the AA pool. Thus, changes in G_2 is described by the next equation:

$$\dot{G}_2 = -k_{81}G_2 \cdot ||x_{ut}||_\infty + k_{82} \; G_t \qquad (64)$$

where the expression $||x_{ut}||_\infty$ is max value of the vector x_{ut} and the dynamic of G_t is the opposite of G_2, i.e.

$$\dot{G}_t = -\dot{G}_2 \qquad (65)$$

We are not considering in this model any of the elongation factors eEF1, eEF2 and eEF3, since there is no evidence today of a major control mechanism at this level (Merrick and Hershey, 1996). However, if any system of regulation using elongation factors is discovered, it can be easily corporated into the model.

2.5 Amino acid reactions and energy

Amino acids are joined with tRNAs, as described by the following reaction:

$$\text{tRNA+AA} \xrightarrow{k_k} \text{AA-tRNA} \qquad (66)$$

where k_k is the reaction rate, AA is the vector concentration of 20 amino acids, while tRNA and AA-tRNA are the concentrations of the corresponding 20 uncharged- and charged-tRNAs, respectively. The reverse reaction is not considered here, since we assume it is much slower and insignificance to this model. While the elongation process is taking place, charged-tRNAs are continuously contributing amino acids to the growing polypeptide in the A-site and departing the 80S ribosome complex as uncharged-tRNA.

16

Changes in the charged-tRNA concentrations in the cell during the elongation process can be described by the next equation

$$\dot{x}_{ct}^{\,i} = k_k \cdot x_{ut}^i \cdot x_{aa}^i + \dot{x}_{ct}^{\,i}\Big|_{el} \tag{67}$$

where the index i represents the i th element of the vectors corresponding to the i th amino acid. k_k can be chosen to be constant, or alternatively, assuming that reaction between the amino acids and the uncharged tRNA have the same rate for all the amino acids, k_k can be chosen to be a function of the amino acid concentration according to

$$k_k = k_9 \left(1 - e^{-1/c_1\|x_{aa}\|}\right) \tag{68}$$

and k_9 is maximum rate at high AA concentrations while $1/c_1$ is the concentrations of amino acids at about 0.62 saturation.

The change in concentrations of uncharged tRNA can be described by the next equation

$$\dot{x}_{ut}^{\,i} = -k_{81} \cdot \|x_{ut}\|_\infty \cdot G_2 + k_{82} \cdot G_t - k_k\, x_{ut}^i \cdot x_{aa}^i + \frac{d(x_{ut}^{\,i})}{dt}\Big|_{el} \tag{69}$$

where $\frac{d(x_{ut}^i)}{dt}\Big|_{el}$ is found using equation (59). Experiments have shown that under normal conditions, tRNAs are 90-100% charged (Surdin-Kerjan et al., 1973; Lewis and Ames, 1972) implying that k_k is high with respect to the rate of peptide elongation. In situations where x_{aa}^i is depleted (due to amino acid starvation for example), the term $k_k x_{ut}^i x_{aa}^i = 0$ in equation (67) and the consumption rate of charged-tRNAs becomes:

$$\dot{x}_{ct}^{\,i}(t \geq t_s) = 0 - r_{80}(t)\frac{v_m \cdot x_{req}^i}{L_{mRNA}} \tag{70}$$

where t_s is the time point of amino acid starvation, and v_m is still constant velocity given by (52) as long as $x_{ct}^i(t)$ is not close to zero. Thus rate of charged-tRNA reduction is proportional to the number of 80S ribosomes in the process of elongating.

Energy consumption results from transformation of e_T to e_D in the initiation process, and hydrolysis of one GTP molecule to GDP for each codon the 80S ribosome passes through. The change in GTP consumption due to elongation is proportional to the number of active ribosomes 80S moving on the strand and the velocity of the ribosomes (in units of $Codons/sec$). Thus, the rate of change of total GTP and GDP is

$$\dot{x}_T = -r_{80} \cdot v - x_T[k_3 e_2 + k_{72}e_{com}] \tag{71}$$
$$\dot{x}_D = r_{80} \cdot v + k_{72}e_{com}x_T - k_{10}x_D e_p \tag{72}$$

17

and the total amount of energy in *Joul* that has been used in protein synthesis is simply the amount of x_D produced from initial time t_0 multiplied by $30 kJ/mol$ and in this case, dividing by Avogadro number to adjust to number of molecules, as

$$\mathcal{E}(t) = \frac{30 \cdot 10^3}{Avg} \int_{t_0}^{t} [r_{80}(\eta)v(\eta) + k_{72}e_{com}(\eta)x_T(\eta) - k_{10}x_D(\eta)e_p(\eta)]\, d\eta \quad (73)$$

where t_0 is the initial process time. Rate of protein production is a result of the peptide being released when the 80S ribosome recognizes the termination codon, and is given by the following equation

$$\dot{\Pi} = \begin{cases} \frac{1}{\rho_u}k_6 \cdot \phi \cdot v(t) \cdot sc(t - \tau) \cdot r_{60}(t - \tau) & \text{if } t \geq \tau \\ 0 & \text{else} \end{cases} \quad (74)$$

and the total protein production is

$$\Pi(t) = \frac{k_6}{\rho_u}\phi \int_{t_0}^{t} v(t) \cdot sc(t - \tau) \cdot r_{60}(t - \tau)dt , \qquad t_0 \geq \tau \quad (75)$$

2.6 Summary

We summarize here the differential equations discussed in sections 2.1 through 2.5. First, the initiation and eIF4 control diff. eq. are the next set:

$$\dot{sc} = k_{52} \cdot ac \cdot E_F - k_6 \cdot \rho_u^{-1} \cdot sc \cdot r_{60} \quad (76)$$
$$\dot{tc} = k_4 \ e_T \cdot x_M - k_{51} \ tc \cdot r_{40} \quad (77)$$
$$\dot{ac} = k_{51} \ tc \cdot r_{40} - k_{52} \cdot ac \cdot E_F \quad (78)$$
$$\dot{E}_4 = k_{422}E_F + H \cdot D_4 - E_4 \left[k_{43}E_{BP} + k_{421}E_G\right] \quad (79)$$
$$\dot{E}_p = H \cdot [E_{BP} + D_4] - k_{411}E_p \quad (80)$$
$$\dot{D}_4 = k_{43} \cdot E_{BP}E_4 - H \cdot D_4 \quad (81)$$
$$\dot{E}_{BP} = k_{411}E_p - E_{BP} \cdot [H + k_{43}E_4] \quad (82)$$
$$\dot{E}_G = k_{422}E_F - k_{421}E_4E_G \quad (83)$$
$$\dot{E}_F = -\dot{E}_G \quad (84)$$

The equations for eIF2 control are the next set:

$$\dot{e}_2 = k_{12}e_p - e_2(k_3x_T + k_{11}G_t) \quad (85)$$
$$\dot{e}_p = k_{11}e_2G_t - e_p(k_{12} + k_{10}x_D) \quad (86)$$
$$\dot{D}_2 = k_{21} \ e_B \ e_{pD} - k_{22} \cdot D_2 \quad (87)$$
$$\dot{e}_B = k_{22} \cdot D_2 + k_{72} \cdot e_{com} \ x_T - e_B \cdot (k_{21} \ e_{pD} + k_{71} \cdot e_D) \quad (88)$$
$$\dot{e}_D = k_{92} \ e_{pD} - e_D \cdot (k_{71}e_B + k_{91}G_t) + k_6\rho_u^{-1}sc \cdot r_{60} \quad (89)$$
$$\dot{e}_{pD} = k_{22}D_2 + k_{91}G_t \ e_D + k_{10} \ x_D \ e_p - e_{pD}(k_{21}e_B + k_{92}) \quad (90)$$
$$\dot{e}_{com} = k_{71}e_Be_D - k_{72}e_{com}x_T \quad (91)$$
$$\dot{e}_T = k_3e_2x_T + k_{72}e_{com}x_T - k_4e_Tx_M \quad (92)$$

18

and elongation equations are the set of equations

$$\dot{pos} = v(t) \tag{93}$$

$$\gamma \cdot \dot{v} = q(t)E - \beta v(t) \tag{94}$$

$$\dot{x}_{ut}^{i} = r_{80} \cdot v(t) \cdot x_{req}^{i} \cdot L_{mRNA}^{-1} - k_{81}G_2 \cdot ||x_{ut}||_{\infty} +$$
$$+ k_{82} \ G_t - k_k \cdot x_{ut}^{i} \cdot x_{aa}^{i} \tag{95}$$

$$\dot{x}_{ct}^{i} = k_k \cdot x_{ut}^{i} \cdot x_{aa}^{i} - r_{80} \cdot v(t) \cdot x_{req}^{i} \cdot L_{mRNA}^{-1} \tag{96}$$

$$\dot{x}_{aa}^{i} = U_{aa}^{i} - k_k x_{ut}^{i} x_{aa}^{i} \tag{97}$$

$$\dot{G}_2 = k_{82} \ G_t - k_{81}G_2 \cdot ||x_{ut}||_{\infty} \tag{98}$$

$$\dot{G}_t = -\dot{G}_2 \tag{99}$$

$$\dot{x}_T = -r_{80} \cdot v - x_T[k_3 e_2 + k_{72}e_{com}] \tag{100}$$

$$\dot{x}_D = r_{80} \cdot v + k_{72}e_{com}x_T - k_{10}x_D e_p \tag{101}$$

Finally the change of ribosomes and the protein production can be described as

$$\dot{r}_{40}(t) = -k_{51} \ tc(t) \cdot r_{40}(t) + \rho_u^{-1}k_6 \cdot \phi \cdot v(t) \cdot sc(t-\tau) \cdot r_{60}(t-\tau) \tag{102}$$

$$\dot{r}_{60}(t) = \rho_u^{-1}k_6 \cdot [-sc(t) \cdot r_{60}(t) + \phi \cdot v(t) \cdot sc(t-\tau) \cdot r_{60}(t-\tau)] \tag{103}$$

$$\dot{r}_{80}(t) = \rho_u^{-1}k_6 \cdot [sc(t) \cdot r_{60}(t) - \phi \cdot v(t) \cdot sc(t-\tau) \cdot r_{60}(t-\tau)] \tag{104}$$

$$\dot{\Pi}(t) = \rho_u^{-1}k_6 \cdot \phi \cdot v(t) \cdot sc(t-\tau) \cdot r_{60}(t-\tau) \tag{105}$$

where the time delay $\tau = t_e - t_0$ is computed using eq. (56).

3   Analysis

3.1   Stability analysis

Analysing the system of equations (76)-(105) is a very complex task, mainly due to the non linearity of the system and the time delay characteristic of the model. In order to gain some insight into the stability of the controllers, we shall decouple the E_4 controller equations 79- 84 from the rest of the system and prove it corresponds to the mass balance law, i.e. the mass of E_4 control intermediates does not increase or decay. We bring first the next proposition without a proof:

Proposition 1 (non-negative system) *The system given by equations (76)-(105) is non-negative, i.e. all its variables are non-negative* $\forall t \geq 0$

One can argue that proposition 1 holds since all the variables except $pos(t)$ and $v(t)$ are in units of concentration, which does not have negative value. Note that $pos(t)$ and velocity $v(t)$ never become negative since a ribosome can not reverse its motion. By decoupling E_4 control equations and defining

it as a sub-system, one can show that this sub-system is conserving its mass. Proving this will varify that E_4 controller equations are valid and matter is not generated 'from nothing' or total mass does not dissipitate from the system. First, it can be easily shown that no isolated equilibrium points exist for the E_4 controller sub-system (or the e_2 sub-system). To see that, define a vector of variables $x(t) = (E_4, \ E_p, \ D_4, \ E_{BP}, \ E_G, \ E_F)^T$ and note that (79) to (84) can be written as $\dot{x} = f(x,t)$. Steady state occurs when $\dot{x} = 0$ (Khalil (2002)) and since $\dot{E}_F = -\dot{E}_G$ (equation 84) we find that

$$E_G(t) = \int_0^t [k_{422} E_F(\eta) - k_{421} E_4(\eta) E_G(\eta)] d\eta + E_G(0)$$

$$E_F(t) = \int_0^t [-k_{422} E_F(\eta) + k_{421} E_4(\eta) E_G(\eta)] d\eta + E_F(0)$$

so when $t_{ss} \to \infty$ at steady state, $E_G(t_{ss}) + E_F(t_{ss}) = E_G(0) + E_F(0)$ and thus the equilibrium point $x_e(t_{ss})$ depends on the initial conditions and is not isolated.

Next, we use *Lyapunov* stability theory for nonlinear systems (Khalil, 2002) to show that the E_4 sub-system is bounded. The sum of concentrations of the sub-system is the Lyapunov function

$$\begin{aligned} V_1(x,t) &= E_4 + E_p + D_4 + E_{BP} + E_G + E_F \qquad\qquad (106) \\ &= \|x\|_1 \end{aligned}$$

where the last equality holds since the vector x is positive from proposition 1. It is straightforward to see that $V_1(x,t) > 0, \ \forall x \neq 0$ since all the variables are non-negative. A quick look at figure 4 reveals that both D_4 and E_F are joined molecules of E_4, E_{BP} and E_4, E_G, respectively. We shall therfore define a second auxilary (Lyapunov) function V_2

$$V_2(x,t) = E_4 + E_p + 2D_4 + E_{BP} + E_G + 2E_F \qquad\qquad (107)$$

where $V_2(x,t) > 0, \ \forall x \neq 0$. Finding the derivatives of V_2 with respect to time gives

$$\begin{aligned} \frac{dV_2(x,t)}{dt} &= \dot{E}_4 + \dot{E}_p + 2\dot{D}_4 + \dot{E}_{BP} + \dot{E}_G + 2\dot{E}_F \\ &\equiv 0 \end{aligned}$$

which implies that $V_2(x,t)$ does not change and thus $V_2(x,0) = V_2(x,t), \ \forall t$. But since

$$\begin{aligned} V_2(x,t) &= E_4 + E_p + 2D_4 + E_{BP} + E_G + 2E_F \\ &\geq E_4 + E_p + D_4 + E_{BP} + E_G + E_F \\ &= \|x\|_1 \end{aligned}$$

then $V_2(x,0) = V_2(x,t) \geq V_1(x,t)$, $\forall t \geq 0$ and we can conclude that the total concentration $\|x\|_1$ will always be smaller than the sum of initial concentrations, i.e. the controller is bounded from above by $V_2(x,0)$. Using the same approach, define a third function $V_3(x,t) = 0.5E_4 + 0.5E_p + D_4 + 0.5E_{BP} + 0.5E_G + E_F > 0$, $\forall x \neq 0$. Since $\dot{V}_3 \equiv 0$, then $V_3(x,0) = V_3(x,t)$ and

$$V_3(x,t) \leq \|x\|_1 = V_1(x,t)$$

and we conclude that V_3 is a lower bound on the total concentration. Since D_4 and E_F are composed of two intermediates each, the total mass of the system does not change (conservation of mass), while the total concentration $\|x\|_1$ is decreased or increased when D_4 and/or E_F is synthesized or catabolized, respectively, never falling below $V_3(x,0)$ or increasing above $V_2(x,0)$. Since the sub-system does not have isolated equilibium points (or attractors), we conclude that the trajectories $x(t)$ converge to a finite value, bounded by $V_1(x,0)$ and $V_3(x,0)$, dependent on the initial value $x(0)$.

e_2 controller can also be decoupled from the rest of the system, and presented by the differential equations set (85) to (92), with the states $x_2 = (e_2,\ e_p,\ e_B,\ e_D,\ e_T,\ e_{pD},\ D_2,\ e_{com})^T$. By solving $\dot{x}_2 = 0$, it is easy to prove that similar to E_4, this sub-system does not have isolated equilibrium points and the trajectories converge to a steady state, which is dependent on the initial values. It is more complex to show that e_2 controller is bounded since this sub-system is affected by G_t. If $G_t = 0$, then proof of stability can be performed in a similar manner to E_4. For the more general case, a complete proof is a subject to future work.

3.2   Initial conditions and synthesis rate

Initiation rate depends mainly on the concentrations of e_2 and e_B, since these two variables affects the ribosome loading rate and eventually the protein synthesis rate. However, as was shown in the former section, the final concentrations of e_2 and e_B depend on their initial values. Figure 6 presents variations the rate of synthesis of an actual protein, yeast calmodulin (Cmd1p) as a function of initial e_2 and e_B concentrations for a case where amino acid supply is abundant. The figure shows a sharp increase in rate of protein synthesis (molecules per second) for initial e_2 and e_B values between $0.1nM$ to $3nM$, However, this increase flattens out due to the physical limitation of loading rate on the mRNA (see section 2.3). It is interesting to note that synthesis rate increases sharply for a very small concentrations of e_B, for example increase of e_B from 0.2 to 1 increases the synthesis 2.5 fold, while larger concentrations of eIF2 are needed to achieve the same increase. Mikami et al. (2005) reports increase of 1.7-2.5 fold in synthesis rate when added eIF2B and eIF2 to the
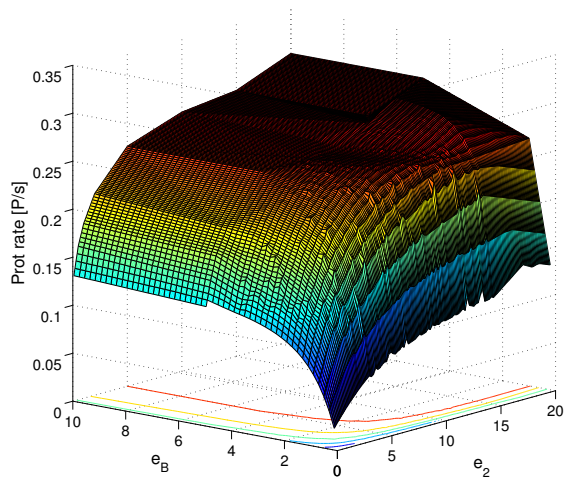
(a) front view
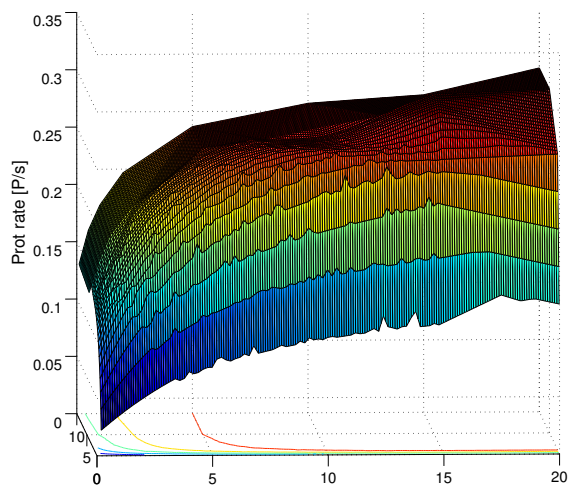
(b) Side view

Fig. 6. Rate of CMD1 protein synthesis (protein/sec) as a function of initial values of e_2 and e_B. The rate increases rapidly, but flattens out around $e_2 = 2nM$ and $e_B = 3nM$. Any additional increase in the intermediates does not affect the protein synthesis rate.

experiments. They also report stimulation of protein synthesis by addition of eIF2 alone, which is supported by figure 6.

4 Simulation results

In this section, we provide results from simulations of the model equations (76)- (105), using MATLAB. In order integrate the equations we use a simple first Euler integration method with time step 10^{-3} over 400 seconds. We chose to demonstrate the model on transcripts from the yeast *TDH3* (glyceraldehyde-3-phosphate dehydrogenase 3), *CMD1* (Calmodulin) and *ACT1* (Actin) genes. Some of the properties of the proteins are shown in table 4. Tdh3p and Cmd1p are translated quite efficiently, as judged by ribosome loading, whilst the *ACT1* mRNA is loaded less efficiently (MacKay et al., 2004), suggesting a more complicated 5'UTR structure with energy barriers (Wong et al., 2004; Mathews et al., 1999), which have to be accounted for in using higher value of ρ_u. The parameters and the rate constants used in this model are given by the table in Appendix A.1. The first simulation demonstrates feed-

Protein	Length (AA)	Limiting AA	Function
TDH3	332	Val-37 (11.1%)	Gluconeogenesis, Glycolysis
CMD1	147	Leu-18 and Ser-18 (12.2%)	Ca++ binding protein
ACT1	375	Ser- 31 (8.3%)	Actin, structural protein involved in cell polarization, endocytosis, and other cytoskeletal functions

Table 4

Some properties of the simulated proteins

ing of amino acids at a constant rate (equal feeding to all amino acids, $U_{aa}^i = 10\ aa/s$), and starvation after about 130 seconds (then $U_{aa}^i = 0\ \forall i$). Results from simulating cmd1p synthesis are presented in figures 7 and 8. The first 80S ribosome reaches the termination codon when $pos(t_e) = L_{mRNA} = 147 codons$. This occurs at $t_e = 16.1 sec$, thus $\tau = t_e - t_0 = 16.1 sec$. Then the subsequent ribosomes leave the mRNA producing sharp peaks in the density. As can be seen from figure 7, the density slightly oscillates at steady state (between 120 to 250 seconds), due to loading and unloading of ribosomes, where the steady state value shows 4-5 ribosomes on the mRNA (figure 8 top left) with density of $d = r_{80}/L_{mRNA} = 4.5/147 \approx 0.03\ rib/codon$. At steady state, number of uncharged-tRNAs is very low due to rapid reaction with amino acids, see eq. (66). This causes low concentration of GCN2·tRNA (figure 8 top right), hence low rate of e_D phosphorylation process, where concentrations of D_2 and e_{pD} are relatively low (figure 8 bottom right and center, respectively), while concentrations of e_B are high (figure 8 bottom left), enabling fast recycling of e_D to e_T and rapid initiation. After 130 seconds feeding stops, and concentrations of the limiting amino acids (*Ser*, *Glu* and *Leu*) are reduced rapidly, until depletion at time $t = 230 sec$. After this depletion, the reaction (66) does not occur for these amino acids and levels of the corresponding uncharged-tRNAs increase rapidly, which triggers fast production of the enzyme G_t (figure 8, top right). This enzyme accelerates the rate of phosphorylation of e_D to e_{pD}, which reacts with e_B to form D_2 inhibited complex, hence removing e_B from
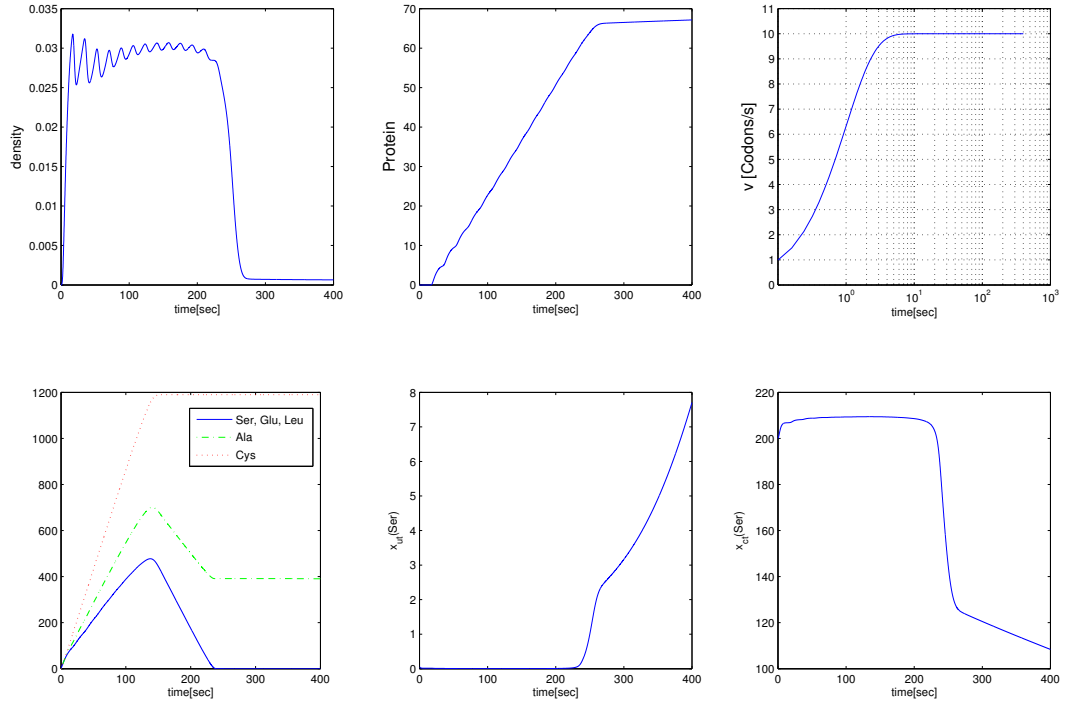
23

Fig. 7. Simulation of cmd1p synthesis: Density of r_{80} on the mRNA (top left), peptide production (top center), velocity of the 80S ribosomes (top right). Bottom figures: number of amino acid, uncharged and charged-tRNA.

the system (figure 8 bottom left). When this occurs, e_D can no longer be recycled to e_T, thus translation rate is reduced and e_2 controller stops loading ribosomes after this time. Notice the sharp drop in e_T concentrations when ribosome loading is stopped. e_T is being used in reaction (2), but not recycled back from e_D to e_T since very low concentrations of e_B are presented.

4.1 Deactivating eIF2 control

In this section we demonstrate the importance of the eIF2 control mechanism on the protein Tdh3p synthesis (see properties of this protein in table 4). Two simulations were executed, one using $k_{21} = 0.4$ rate and another using $k_{21} = 0$ (marked as $-\cdot-$ line and $-*-$, respectively, in figures 9 and 10). $k_{21} = 0.4$ simulates normal operation while $k_{21} = 0$ represents deactivation of e_2 controller, thus preventing inhibition e_B by e_{pD}. These simulations clearly demonstrate the efficiency of the eIF2 control. Note that the amount of 80S ribosomes at steady state is a bit higher with eIF2 control deactivated than when using eIF2 control. This comes from the fact that using the reaction (14) reduces the concentrations of available e_B in the system and thus reducing the
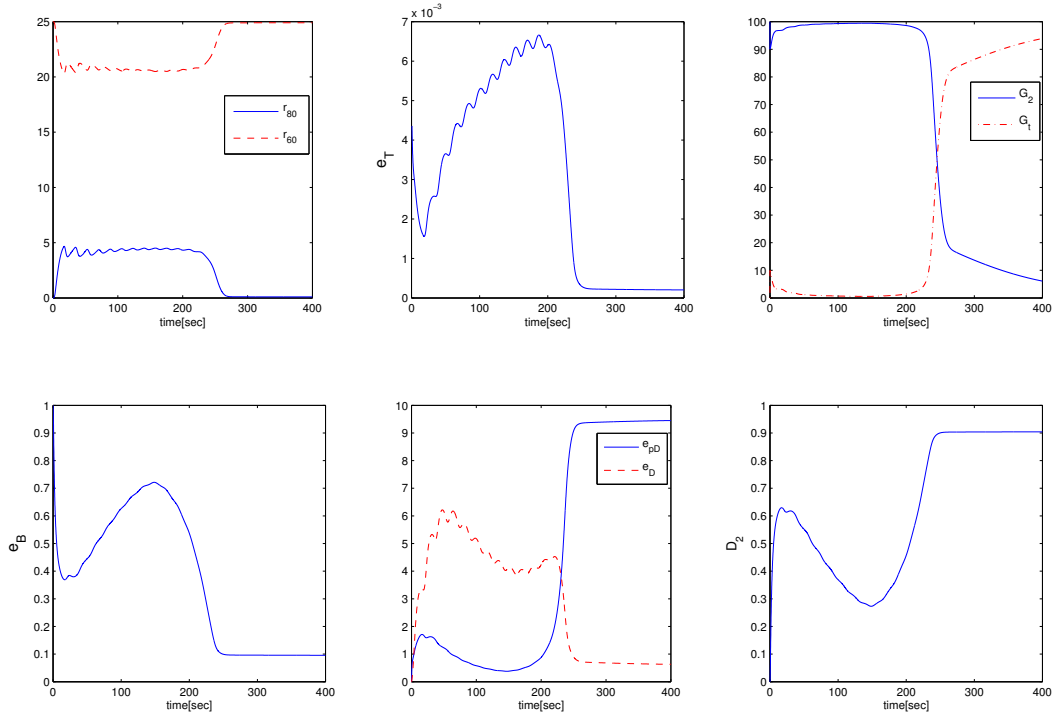
Fig. 8. eIF2 control dynamics of cmd1p synthesis

initiation rate. The advantage of the eIF2 control is obvious. In the case where $k_{21} = 0.4$, the limiting amino acid Val is depleted after about 250 seconds, which causes reduction in $x_{ct}(Val)$ and increase in $x_{ut}(Val)$, phosphorylation rate is increased (figure 10, bottom center, $-\cdot-$). In amino acid starvation, the eIF2 control reduces the loading rate of the ribosomes gradually to zero (see density plot at figure 9, top left) by joining e_B (figure 10, bottom left, $-\cdot-$) with e_{pD} and creating D_2 (bottom right, $-\cdot-$). When $k_{21} = 0$, e_B does not react with e_{PD}, and $D_2 = 0 \; \forall t$ (bottom right, $-*-$). Then e_B concentrations are always high, causing large amount of ribosomes on the mRNA to deplete the remaining charged-tRNAs rapidly (fig. 9 bottom right, $-*-$), eventually stopping elongation when no more $x_{ct}(Val)$ is available. Using eIF2 control however, causes gradual depletion of e_B and reduction of initiation rate, and by reducing loading rate of ribosomes on the mRNA, only few are stalled on the strand at starvation.

4.2   Control through eIF4F regulation

As has been described in section 2.2.2, the putative hormonal signal H affects E_F attachment to the 5' terminal cap of the mRNA. If E_G is not presented at the cap, the initiation is inhibited by preventing r_{40} attachment to the cap
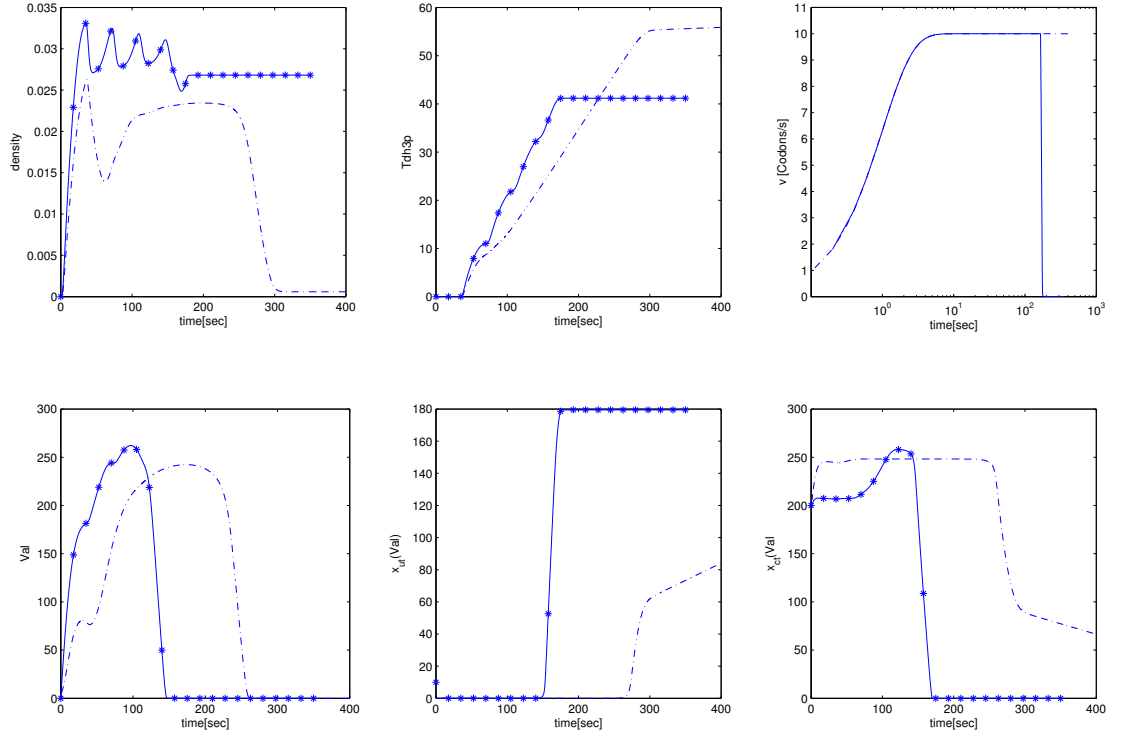
Fig. 9. Simulation of tdh3p synthesis, with eIF2 control $(-\cdot-)$ and without $(-*-)$

and scanning along the untranslated region. This behavior is demonstrated here by simulating *ACT1* synthesis (see figure 11). Since the ACT1 mRNA seems to have a complex 5'UTR structure, ρ_u is chosen to be high, $\rho_u = 50$ (see section 2.1 for discussion about the 5'UTR structure and initiation). The first simulation is operated at full synthesis, where $H = 1$ (solid lines). Rate of protein synthesis decreases after about $400sec$ due to starvation of amino acids acting through the eIF2 controller. The second simulation presents a case where hormonal signal H is reduced to an arbitrary low value of $H = 0.01$ after 120 seconds, simulating a case where the cell might inhibit protein synthesis due to stress, viral attack or other disturbances. Concentrations of D_4 are increased rapidly due to the signal through a mechanism involving reduction of E_F concentrations. The density of the ribosomes on the mRNA (figure 11, top left) is decreased then to zero after 350 seconds, and protein synthesis stops.
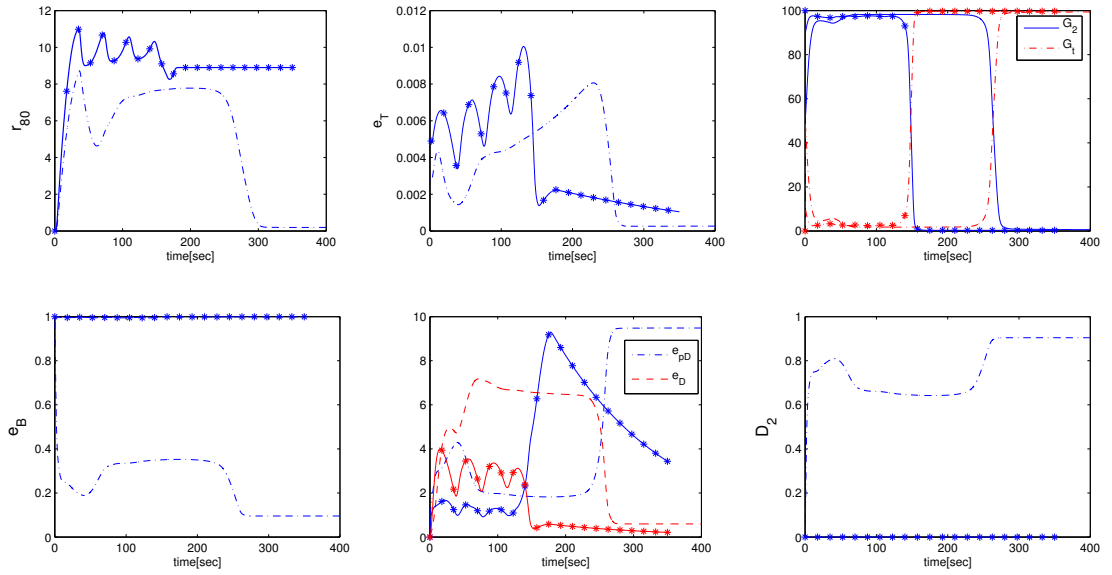
Fig. 10. Simulation of tdh3p synthesis, with eIF2 control $(-\cdot-)$ and without $(-*-)$.
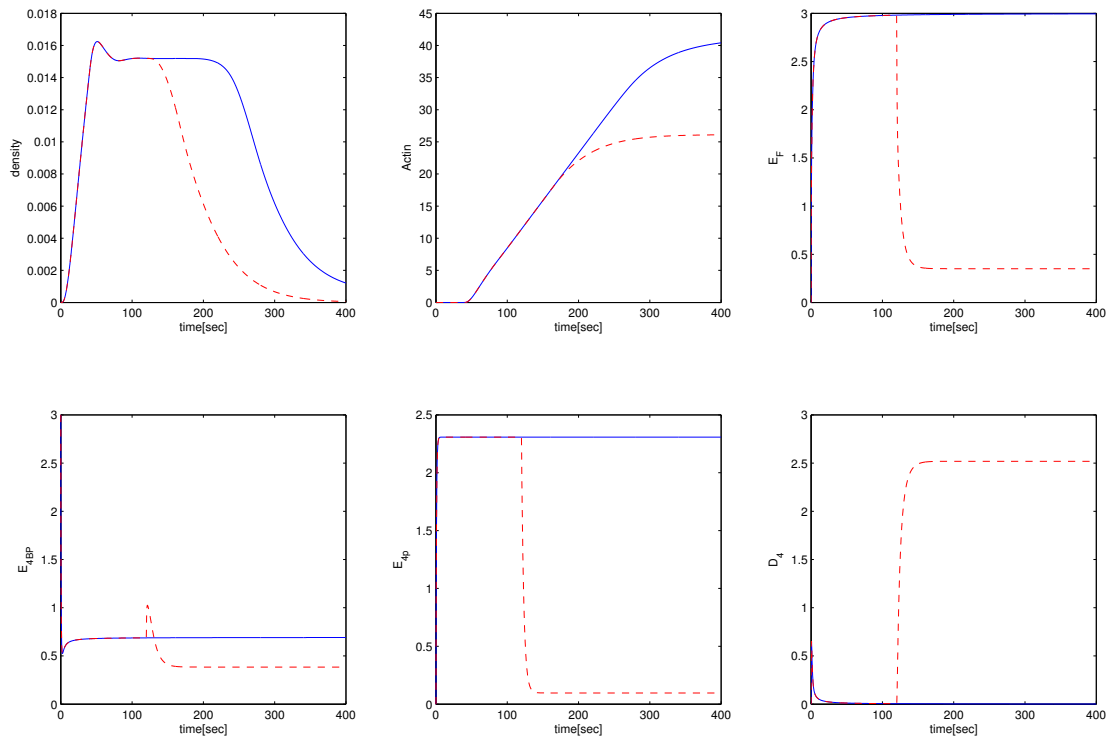

Fig. 11. Simulation results for act1p synthesis. Solid line: no changes in H are presented. Stippled line: H is reduced to 0.01 at $t = 120sec$.

5 Relationship of the model to experimental results

Genome-level experiments have provided ribosome loading data across the transcriptome of yeast (Arava et al., 2003; MacKay et al., 2004). From the average number of ribosomes loaded and the length of the open reading frame, one can calculate the ribosome density (ribosomes per codon) on individual mRNA species (MacKay et al., 2004). Examples of three transcripts, chosen for different lengths and ribosome densities, are shown in figure 12. The *CMD1* and *TDH3* transcripts differ by approximately a factor of 2 in both length and number of ribosomes loaded, resulting in very similar ribosome densities (0.027 and 0.024, respectively). These two transcripts are clearly much more efficiently loaded than the transcriptome average of 0.014 in yeast (MacKay et al., 2004). *ACT1* is the longest transcript of the three, but is loaded with fewer ribosomes than *TDH3* (figure 12) and has a ribosome density of 0.016, which is not significantly different than the average. These ribosome densities from experiment compare favorably with steady state densities derived from simulation of 0.030, 0.022 and 0.15 for CMD1, TDH3 and ACT1, respectively (figures 7, 9 and 11).
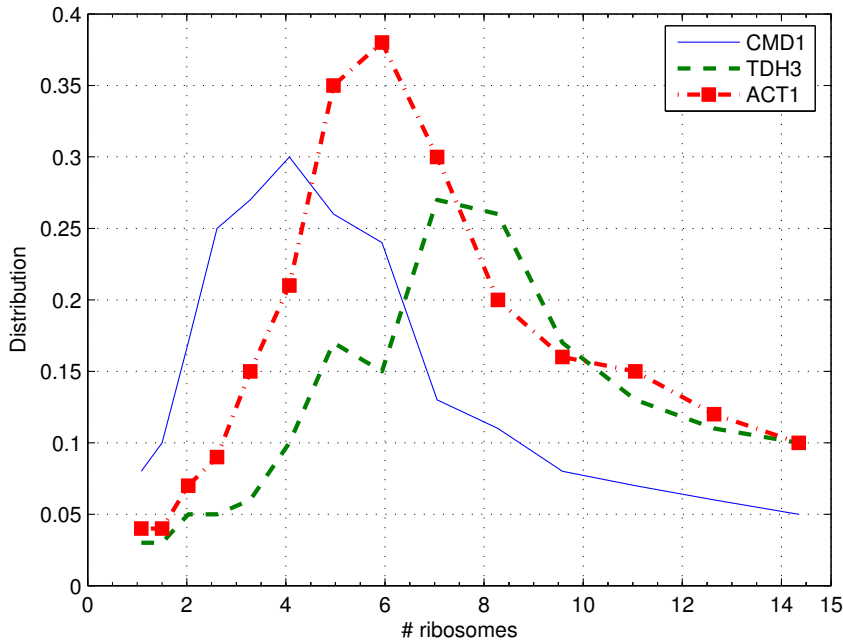

Fig. 12. Ribosome loading profiles of the transcripts from the *CMD1*, *TDH3* and *ACT1* genes. The data are derived from microarray experiments (MacKay et al., 2004) are plotted as relative transcript level (ratio Cy3/Cy5 fluorescence) as a function of ribosome number. The sizes of the open reading frames of these transcripts are 444, 999 and 1128 nucleotides, respectively.

6 Discussion

6.1 *eIF2 control*

This paper demonstrates the advantages of developing and using a dynamical model of protein synthesis. Section 3.2 predicts increase of protein synthesis as a function of e_2 and e_B. Very small concentrations of e_B were needed to increase the synthesis rate 1-3 fold, comparing with e_2 which was needed in larger concentrations. This can be explained by the fact that very small concentrations of e_B are needed to recycle eIF2·GDP to eIF2·GTP, and these amounts are depleted rapidly when reacting with phosphorylated eIF2·GDP. Thus eIF2 controller is an important mechanism for regulating initiation. If an uncontrolled loading of ribosomes on the mRNA occurs, the ribosome density will be large and consumption of amino acids will be rapid, without ability to inhibit initiation in case of depletion of a limiting amino acid. This will create a situation where possibly large numbers of ribosomes are stalled on the mRNA, as has been demonstrated by simulation of *TDH3* with and without control. In this example, removing the eIF2 control causes 9 ribosomes to be stalled on the gene, vs. none when eIF2 control is operative (see figure 9 and 10). Under these conditions, translation would be very inefficient and energy is wasted in form of stalled ribosomes. If amino acid starvation continues or a stress situation arises where mRNA needs to be broken down and synthesis aborted, then energy will be wasted in the form of production of incomplete peptides. The total number of incomplete peptides can be quite large: $r_{80} \cdot [mRNA]$. Even in an instance where turnover does manage to supply the limiting amino acid to a certain extent, large numbers of ribosomes will either require large turnover rate or will cause the elongation velocity v to oscillate, causing the ribosomes to stall frequently and the elongation to be inefficient.

6.2 *Difference in eIF2 and eIF4 control*

eIF2 and eIF4 controllers operate with two different input signals. Amino acid level (U_{aa}) is the input for the eIF2 controller in the present model, recognizing that other signals such as virus infection and E.R. stress can activate eIF2 phosphorylation. An external signal, represented as H, is the input signal for regulation through eIF4. e_D concentration affects the eIF2 controller directly, which the controller converts to e_T, but these signals are not controlled by the cell (see figure 3). U_{aa} is only partially controllable through protein turnover, but not entirely since the cell does not have direct control on the access to essential amino acids. eIF2 controller can be viewed as endogenous control mechanism, sensitive to rapid fluctuations in amino acid concentrations. This

controller responds rapidly to different amino acid conditions. Hence, eIF2 control links initiation to the availability of the charged tRNA for elongation and thereby guarantees optimal performance through efficient use of ribosomes. Amino acid levels are not the only endogenous signals affecting the eIF2 controller. Another example is discussed by Scheuner et al. (2005), where regulation of the endoplasmic reticulum (ER) and insulin production in the pancreas are coordinated in order to meet the demand imposed by a high-fat diet in mice.

The eIF4 controller is regulated by changes in exogenous hormonal signals, which are controlled at a higher level by signal transduction pathways triggered by the hormones. This second pathway is independent of eIF2. It can be advantagous in response to a long term amino acid starvation or stress due to viral attack, where energy must be saved and resources diverted from protein synthesis to other processes for uncertain time period. Furthermore, continuous elongation in amino acid limitation can increase the risk for misincorporation. When the situation is back to normal and amino acids are available, the organism can resume normal translation rate by stimulating hormones.

6.3 Response to variations in amino acid supply

Regulation of eIF2 by GCN2 is activated in low concentrations of amino acids. Under these conditions, uncharged-tRNA concentrations are high, increasing phosphorylation of eIF2 and inhibiting the initiation process. This mechanism is not responsive to fluctuations at high amino acid concentrations, where tRNA would be expected to be in a fully charged state. The high eIF2 concentrations will permit the initiation controller to load ribosomes without considering variations of high amino acid concentrations. This is supported by the simulations of the model, where the amino acid levels of *Val* are high at $t = 150$ seconds (see figure 7, bottom left), yet the controller loads ribosomes at a constant level during the entire steady state period, without regard to increase in *Val*). This is also noted by Kimball and Jefferson (2000). Ribosome loading is limited only by the concentrations of ribosomes and the time it take to mobilize them and load them on the 5' terminal cap. Consistant with this concept, ribosome loading profiles are similar in yeast growing in either minimal medium, or in medium supplemented with high concentrations of amino acids (G. Mize and D.R. Morris, unpublished results)

6.4 Elongation model

The elongation model is based on the assumption that amino acid composition is evenly distributed across a gene, an assumption which is not true in reality.

The model does not take into account situations where one 80S ribosome is stalled due to lack of amino acid, while the ribosomes elongating behind it close the distance, creating a non-uniform density along the mRNA. The model assumes a uniform ribosomes distribution where the velocity $v(t)$ is the same for all the ribosomes, where stall means $v(t) = 0$ and *all* the ribosomes stop at the same time. This fact does not change the protein production results or the loading rate. The model can account however for different types of genes, where resistivity will change the steady state velocity of the ribosomes due to secondary structures available. Furthermore, the few protein synthesis dynamic models deal with only one ribosome elongating at a time, a fact which poses a limitation on the results (as discussed by Drew (2001)). The model presented here overcomes this limitation, enabling description of properties such as density of ribosomes, steady state velocity, etc.

6.5   Motion along the 5'UTR

The motion of the *sc* on the 5'UTR was modeled in this paper as a simple integrator, where the constant ρ_u represent the resistivity of the UTR. The UTR structure will vary between different genes, and the motion of the *sc* along the UTR may not be constant. This will impose different loading rates between genes and will result in different densities. It is possible to model the motion of *sc* the same way elongation of r_{80} was modelled on the mRNA. Knowledge about the structure of the UTR of individual transcripts will permit more accurate description of loading rates and densities. This information can be incorporated into the model in place of ρ_u.

7   Conclusion

Through this dynamic model, we are able to understand not only steady state but different transitions and relations between the amino acids, tRNAs, and other variables. This dynamic model can be used to predict not only density of ribosomes on an mRNA strand, but protein production in optimal and non-optimal conditions. Special cases such as amino acid starvation, viral attacks, and stress are all known to affect the protein synthesis process, and this model can predict results from such situations. Furthermore, the model can be incorporated with other dynamic metabolism models in order to predict properties such as growth and energy consumption, under normal and pathological conditions.

8 Acknowledgment

Appendix A List of Variables / parameters

Variable/par.	Value	Size/dim.	Description
E_4	var	nM	eIF4E
E_{BP}	var	nM	eIF4E-BP
E_p	var	nM	eIF4E-BP$_p$ (Phosphorylated eIF4E-BP)
D_4	var	nM	inactive complex
E_G	var	nM	eIF4G
E_F	var	nM	eIF4F
H	$[0,1]$	unitless	Stimulating hormone signal
k_{411}	0.3	s^{-1}	Rate constants for dephosphorylation of E_p
k_{421}	0.6	s^{-1}	Rate constant for E_F formation
k_{422}	var	s^{-1}	Rate constant for E_F breakdown
c_1	0.6	s^{-1}	Constant for the rate k_{422}
k_{43}	0.5	s^{-1}	Rate constants for inactive complex formation
e_2	var	nM	eIF2
e_B	var	nM	eIF2B
e_p	var	nM	eIF2$_p$ (Phosphorylated eIF2)
D_2	var	nM	inactive complex eIF2$_p$GDPeIF2B
e_D	var	nM	eIF2GDP compound
e_T	var	nM	eIF2GTP complex
e_{pD}	var	nM	eIF2$_p$·GDP complex
e_{com}	var	nM	eIF2GDPeIF2B complex
G_2	var	nM	GCN$_2$
G_t	var	nM	GCN$_2$tRNA
k_{11}, k_{12}	1, 0.2	s^{-1}	Rate constants for eIF2 phosphorylation and the reverse reaction
k_{21}, k_{22}	1, 0.2	s^{-1}	Rate constants for D_2 formation and breakdown
k_{51}	0.5	s^{-1}	Rate constant for ac formation
k_p	0.05	s^{-1}	Constant for rate k_{52}
k_6	0.05	s^{-1}	
k_{71}, k_{72}	0.1, 1	s^{-1}	Rate constants for inactive complex
k_{81}, k_{82}	0.8, 0.4	s^{-1}	Rate constants for GCN$_2$ reactions

Table Appendix A.1: Variables and parameters in the dynamic model.'var' represents variable.

Variable/par.	Value	Size/dim.	Description
k_{91}, k_{92}	0.08, 0.5	s^{-1}	Rate constants for phosphotylation and de-phosphorylation of e_D
k_{10}	0.1	s^{-1}	Rate constant of reaction between e_p and x_D
x_{aa}	var	$\mathbb{R}^{1x20}$	Amino acids concentration
x_u	var	$\mathbb{R}^{1x20}$	Uncharged tRNA
x_{ct}	var	$\mathbb{R}^{1x20}$	Charged tRNA
x_D	var	nM	GDP concentrations
x_T	var	nM	GTP concentrations
x_M	var	nM	Methionyl-tRNA
tc	var	nM	Ternary Complex $eIF2 \cdot GTP \cdot Met\text{-}tRNA_i$
ac	var	nM	Active Complex
sc	var	nM	Scanning complex $40S \cdot ac$
r_{40}	var	nM	Ribosome 40S
r_{60}	var	nM	Ribosome 60S
r_{80}	var	nM	Ribosome 80S
d	var	$rib \cdot codons^{-1}$	Density of the 80S ribosomes on the mRNA
x_{req}	const.	$\mathbb{R}^{1x20}$	vector of the amino acid requirement
i	-	-	index for the i amino acid
k_k	4	s^{-1}	Rate constant for amino acid reaction
C_{cap}	1/30	$codons^{-1}$	
α^{-1}	100	s^{-1}	Damping coefficient (mRNA specific)
ρ_u	const.≥ 1	unitless	mRNA resistivity constant. For *TDH3*, *CMD1* and *MYO1*, $\rho_u = 1$, and for *ACT1*, $\rho_u = 50$.
ϕ	0.1	$s \cdot codon^{-1}$	Constant
v	≥ 0	$codon/s$	Velocity of the 80S subunit on the mRNA
q	≥ 0 var		Charge of 80S subunits
E	$30^3/avg$	$J/codon$	Electrical field
γ	$3^3/avg$	kg	Mass of on 80S subunit
β	$\gamma^2/\alpha E$	kg/s	Proportional damping constant (mRNA spesific)
L_{mRNA}	const.	codons	Length of the mRNA in codons (mRNA specific)
τ	var	s	Time delay variable, time it takes 80S subunit to elongate. Computed using eq. (56)

Table Appendix A.1: Variables and parameters in the dynamic model.'var' represents variable.

References

Arava, Y., Wang, Y., Storey, J., Liu, C., Brown, P., Herschlag, D., 2003. Genome-wide analysis of mRNA translation profiles in saccharomyces cerevisiae. Proc. Natl.Acad.Sci.U.S.A. 100 (7), 3889–3894.

Drew, D. A., 2001. A mathematical model for prokaryotic protein synthesis. Bulletin of Mathematical Biology (63), 329–351.

Gingras, A. C., Raught, B., Sonenberg, N., 1999. eif4 initiation factors: Effectors of mrna recruitment to ribosomes and regulators of translation. Annu. Rev. Biochem. 68, 913–963.

Hinnebusch, A. G., 2000. Mechanism and regulation of initiator methionyl-tRNA binding to ribosomes. In: N.Sonenberg, Hershey, J., M.B.Mathews (Eds.), Translational control of gene expression. Cold Spring Harbor Press, Ch. 5, pp. 185–243.

Hinnebusch, A. G., 2004. Study of translational control in eukaryotic gene expression using yeast. Ann. N.Y. Acad.Sci. 1038, 60–74.

Holcik, M., Sonenberg, N., 2005. Translational control in stress and apoptosis. Nat. Rev. Mol. Cell Biol. 6 (4), 318–327.

Khalil, H. K., 2002. Nonlinear Systems, 3rd Edition. Prentice Hall, Upper Saddle River, New Jersey.

Kimball, S. R., Jefferson, L., 2000. Regulation of translation initiation in mammalian cells by amino acids. In: Sonenberg, N., Hershey, J., Mathews, M. B. (Eds.), Translational Control of Gene Expression. Translational Control. Cold Spring Harbour Laboratory Press, Ch. 16, pp. 561–579.

Kozak, M., October 1991. Structural features in eukaryotic mrnas that modulate the initation of translation. The Journal of Biological Chemistry 266 (30), 19867–19870, the American Society for Biochemistry and Molecular Biology, Inc.

Lewis, J., Ames, B., 1972. The percentage of transfer RNA his charged in vivo and its relation to the repression of the histidine operon. Journal of Molecular Biology 66, 131–142, histidine regulation in *Salmonella typhimurium.*

MacKay, V. L., Li, X., Flory, M. R., Turcott, E., Law, G. L., Serikawa, K. A., Xu, X. L., Lee, H., Goodlett, D. R., Aebersold, R., Zhao, L. P., Morris, D. R., 2004. Gene expression analyzed by high-resolution state array analysis and quantitative proteomics: response of yeast to mating pheromone. Mol Cell Proteomics 3 (5), 478–489.

Mathews, D. H., Sabina, J., Zuker, M., Turner, D. H., 1999. Expanded sequence dependence og thermodynamic parameters improves prediction of RNA secondary structure. Journal Mol. Biol. (288), 911–940.

Merrick, W. C., Hershey, J. W., 1996. The pathway and mechanism of eukaryotic protein synthesis. In: Hershey, J. W., Mathews, M. B., Sonenberg, N. (Eds.), Translational Control. Vol. 1 of Translational Control. Cold Spring Harbor Laboratory Press, Ch. 2, pp. 31–69.

Merrick, W. C., Nyborg, J., 2000. The protein biosynthesis elongation cycle. In: N.Sonenberg, Hershey, J., M.B.Mathews (Eds.), Translational Control of

Gene Expresison. Cold Spring Harbor Laboratory Press, Ch. 3, pp. 89–125.

Mikami, S., Masutani, M., Sonenberg, N., Yokoyama, S., Imataka, H., October 2005. An effecient mammalian cell-free translation system supplemented with translation factors. Protein Expression and Purification In press, available online 25 October 2005.

Raught, B., Gingras, A. C., Sonenberg, N., 2000. Regulation of ribosomal recruitment in eukaryotes. In: N.Sonenberg, Hershey, J., M.B.Mathews (Eds.), Translational Control of Gene Expresison. Cold Spring Harbor Laboratory Press, Ch. 6, pp. 245–293.

Rowlands, A. G., Panniers, R., Henshaw, E. C., April 1988. The catalytic mechanism of guanine nucleotide exchange factor action and competitive inhibition by phosphorylated eukaryotic initiation factor 2. The Journal of Biological Chemistry 263 (12), 5526–5533.

Scheuner, D., Mierde, D., Song, B., Flamez, D., Creemers, J., Tsukamoto, K., Ribick, M., Schuit, F., Kaufman, R., Jun 2005. Control of mRNA translation preserves endoplasmic reticulum function in beta cells and maintains glucose homeostasis. Natural MedicineNature Publishing Group.

Slotine, J. J., Li, W., 1991. Applied nonlinear control. Prentice Hall, Upper Saddle River, NJ 07458.

Surdin-Kerjan, Y., Cherest, H., Robichon-Szulmajster, H., 1973. Relationship between methionyl transfer ribonucleic acid cellular content and synthesis of methionine enzyme in *Saccharomyces cerevisiae*. Journal Bacteriol. 113, 1156–1160.

Trachsel, H., 1996. Binding of Initiator Methionyl-tRNA to Ribosomes. In: Hershey, J. W., Mathews, M. B., Sonenberg, N. (Eds.), Translational Control. Translational Control of Gene Expression. Cold Spring Harbour, Ch. 4, pp. 113–138.

Voet, D., 2004. Biochemistry, 3rd Edition. John Wiley & sons, Inc., Wiley International Edition.

Wong, K. K. Y., Bouwer, H. G. A., Freitag, N. E., 2004. Evidence implicating the 5' untranslated region of *Listeria monocytogenes actA* in the regulation of bacterial actin-based motility. Cellular Microbiology 6 (2), 155–166.