



Research

Cite this article: Raveh A, Margaliot M, Sontag ED, Tuller T. 2016 A model for competition for ribosomes in the cell. *J. R. Soc. Interface* **13**: 20151062.
<http://dx.doi.org/10.1098/rsif.2015.1062>

Received: 12 December 2015

Accepted: 15 February 2016

Subject Category:

Life Sciences – Engineering interface

Subject Areas:

systems biology, synthetic biology, biophysics

Keywords:

competition for resources, systems biology, monotone dynamical systems, first integral, entrainment, context-dependence in mRNA translation

Authors for correspondence:

Eduardo D. Sontag

e-mail: eduardo.sontag@gmail.com

Tamir Tuller

e-mail: tamirtul@post.tau.ac.il

Electronic supplementary material is available at <http://dx.doi.org/10.1098/rsif.2015.1062> or via <http://rsif.royalsocietypublishing.org>.

A model for competition for ribosomes in the cell

Alon Raveh¹, Michael Margaliot², Eduardo D. Sontag⁴ and Tamir Tuller³

¹School of Electrical Engineering, ²School of Electrical Engineering and the Sagol School of Neuroscience, and ³Department of Biomedical Engineering and the Sagol School of Neuroscience, Tel-Aviv University, Tel-Aviv 69978, Israel

⁴Department of Mathematics and the Center for Quantitative Biology, Rutgers University, Piscataway, NJ 08854, USA

A single mammalian cell includes an order of 10^4 – 10^5 mRNA molecules and as many as 10^5 – 10^6 ribosomes. Large-scale simultaneous mRNA translation induces correlations between the mRNA molecules, as they all compete for the finite pool of available ribosomes. This has important implications for the cell's functioning and evolution. Developing a better understanding of the intricate correlations between these simultaneous processes, rather than focusing on the translation of a single isolated transcript, should help in gaining a better understanding of mRNA translation regulation and the way elongation rates affect organismal fitness. A model of simultaneous translation is specifically important when dealing with highly expressed genes, as these consume more resources. In addition, such a model can lead to more accurate predictions that are needed in the interconnection of translational modules in synthetic biology. We develop and analyse a general dynamical model for large-scale simultaneous mRNA translation and competition for ribosomes. This is based on combining several ribosome flow models (RFMs) interconnected via a pool of free ribosomes. We use this model to explore the interactions between the various mRNA molecules and ribosomes at steady state. We show that the compound system always converges to a steady state and that it always entrains or phase locks to periodically time-varying transition rates in any of the mRNA molecules. We then study the effect of changing the transition rates in one mRNA molecule on the steady-state translation rates of the other mRNAs that results from the competition for ribosomes. We show that increasing any of the codon translation rates in a specific mRNA molecule yields a local effect, an increase in the translation rate of this mRNA, and also a global effect, the translation rates in the other mRNA molecules all increase or all decrease. These results suggest that the effect of codon decoding rates of endogenous and heterologous mRNAs on protein production is more complicated than previously thought. In addition, we show that increasing the length of an mRNA molecule decreases the production rate of all the mRNAs.

1. Introduction

Various processes in the cell use the same finite pool of available resources. This means that the processes actually compete for these resources, leading to an indirect coupling between the processes. This is particularly relevant when many identical intracellular processes, all using the same resources, take place in parallel. The indirect coupling induced by competition may affect the cell's structure and functioning in ways that cannot be understood when studying each process as a single, isolated process.

Biological evidence suggests that the competition for RNA polymerase (RNAP) and ribosomes, and various transcription and translation factors, is a key factor in the cellular economy of gene expression. The limited availability of these resources is one of the reasons why the levels of genes, mRNA and proteins produced in the cell do not necessarily correlate [1–7].

It was estimated that in a yeast cell there are approximately 60 000 mRNA molecules. These can be translated in parallel [8,9], with possibly many

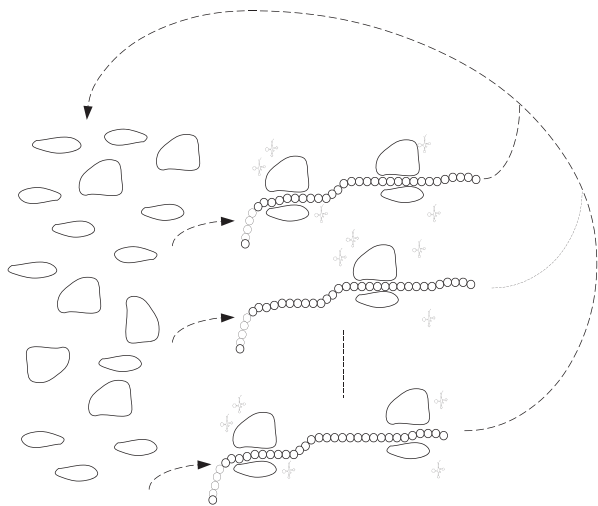


Figure 1. Simultaneous translation of mRNA chains (right) interconnected via a pool of free ribosomes (left). If more ribosomes bind to a certain mRNA molecule then the pool of free ribosomes in the cell is depleted, and this may lead to lower initiation rates in the other mRNAs.

ribosomes scanning the same transcript concurrently. The number of ribosomes is limited (in a yeast cell it is approx. 240 000) and this leads to a competition for ribosomes. For example, if more ribosomes bind to a certain mRNA molecule then the pool of free ribosomes in the cell is depleted, and this may lead to lower initiation rates in the other mRNAs (figure 1).

There is a growing interest in computational or mathematical models that take into account the competition for available resources in translation and/or transcription (see e.g. [10–15]). One such model that explicitly considers the movement of the ribosomes (RNAP) along the mRNA (DNA) is based on a set of *asymmetric simple exclusion processes* (ASEPs) interconnected to a pool of free ribosomes. ASEP is an important model from non-equilibrium statistical physics describing particles that hop randomly from one site to the next along an ordered lattice of sites, but only if the next site is empty. This form of ‘rough exclusion’ models the fact that the particles cannot overtake one another. ASEP has been used to model and analyse numerous multiagent systems with local interactions including the flow of ribosomes along the mRNA molecule [16,17]. In this context, the lattice represents the mRNA molecule, and the particles are the ribosomes. For more on mathematical and computational models of translation, see the survey paper [18].

Ha & den Nijs [19] considered a closed system composed of a single ASEP connected to a pool (or reservoir) of ‘free’ particles. The total number of particles is conserved. This is sometimes referred to as the *parking garage problem*, with the lattice/particles/pool modelling a road/cars/parking garage, respectively. Cook & Zia [20] studied a similar system using domain wall theory. Cook & Zia [21] (see also [22]) considered a network composed of two ASEPs connected to a finite pool of particles. The analysis in these papers focuses on the phase diagram of the compound system with respect to certain parameters, and on how the phase of one ASEP affects the phase of the other ASEPs. These studies rely on the phase diagram of a single ASEP that is well understood only in the case where all the transition rates inside the chain (the elongation rates) are equal. Thus, the network is typically composed of *homogeneous*

ASEPs. Another model [15] combines non-homogeneous ASEPs in order to study competition between multiple species of mRNA molecules for a pool of tRNA molecules. This study was based on the *Saccharomyces cerevisiae* genome. However, in this case (and similar models, such as [14]) analysis seems intractable and one must resort to simulations only.

Our approach is based on the *ribosome flow model* (RFM) [23]. This is a deterministic, continuous-time, synchronous model for translation that can be derived via the mean-field approximation of ASEP [24]. The RFM includes n state variables describing the ribosomal density in n consecutive sites along the mRNA molecule, and $n + 1$ positive parameters: the translation initiation rate λ_0 , and the translation elongation rate λ_i from site i to site $i + 1$, for $i = 1, \dots, n$. The RFM has a unique equilibrium point $e = e(\lambda_0, \dots, \lambda_n)$, and any trajectory emanating from a feasible initial condition converges to e [25] (see also [26]). This means that the system always converges to a steady-state ribosomal density that depends on the rates, but not on the initial ribosomal density profile. In particular, the production rate converges to a steady-state value denoted by R . The mapping from the rates to R is a concave function, so maximizing R subject to a suitable constraint on the rates is a convex optimization problem [27,28]. Sensitivity analysis of the RFM with respect to the rates has been studied in Poker *et al.* [29]. These results are important in the context of optimizing the protein production rate in synthetic biology. Margaliot *et al.* [26] have shown that when the rates λ_i are time-periodic functions, with a common minimal period T , then every state variable converges to a periodic solution with period T . In other words, the ribosomal densities *entrain* to periodic excitations in the rates (due e.g. to periodically varying abundances of tRNA molecules).

In ASEP with *periodic boundary conditions*, a particle that hops from the last site returns to the first one. The mean-field approximation of this model is called the *ribosome flow model on a ring* (RFMR). The periodic boundary conditions mean that the total number of ribosomes is conserved. Raveh *et al.* [30] analysed the RFMR using the theory of monotone dynamical systems that admit a first integral.

Both the RFM and the RFMR model mRNA translation on a single mRNA molecule. In this paper, we introduce a new model, called the *RFM network with a pool* (RFMNP), that includes a network of RFMs, interconnected through a dynamical pool of free ribosomes, to model and analyse simultaneous translation and competition for ribosomes in the cell. To the best of our knowledge, this is the first study of a network of RFMs. The total number of ribosomes in the RFMNP is conserved, leading to a first integral of the dynamics. Applying the theory of monotone dynamical systems that admit a first integral, we prove several mathematical properties of the RFMNP: it admits a continuum of equilibrium points, every trajectory converges to an equilibrium point, and any two solutions emanating from initial conditions corresponding to an equal total number of ribosomes in the system converge to the same equilibrium point. These results hold for any set of rates and in particular when the RFMs in the network are not necessarily homogeneous. These stability results are important because they provide a rigorous framework for studying questions such as how does a change in one RFM affects the behaviour of all the other RFMs in the network? Indeed,

since a steady-state exists, this can be reduced to asking: how does a change in one RFM in the network affects the *steady-state* behaviour of the network?

To analyse competition for ribosomes, we consider the effect of increasing one of the rates in one RFM, say RFM #1. This means that the ribosomes traverse RFM #1 more quickly. We show that this always leads to an increase in the production rate of RFM #1. All the other RFMs are always affected in the same manner, that is, either all the other production rates increase or they all decrease. Our analysis shows that this can be explained as follows. Increasing the rate λ_i in RFM #1 tends to increase the steady-state density in sites $i + 1, i + 2, \dots$, and decrease the density in site i of this RFM. The *total* density (i.e. the sum of all the densities on the different sites along RFM #1) can either decrease or increase. In the first case, more ribosomes are freed to the pool, and this increases the initiation rates in all the other RFMs leading to higher production rates. The second case leads to the opposite result. The exact outcome of increasing one of the rates thus depends on the many parameter values defining the pool and the set of RFMs in the network.

Our model takes into account the dynamics of the translation elongation stage, yet is still amenable to analysis. This allows to develop a rigorous understanding of the effect of competition for ribosomes. Previous studies on this topic were either based on simulations (see, for example, [14,15]) or did not include a dynamical model of translation elongation (e.g. [11,13]). For example, in an interesting paper, combining mathematical modelling and biological experiments, Gyorgy *et al.* [11] study the expression levels of two adjacent reporter genes on a plasmid in *Escherichia coli* based on measurements of fluorescence levels, that is, protein levels. These are of course the result of all the gene expression steps (transcription, translation, mRNA degradation, protein degradation) making it difficult to separately study the effect of competition for ribosomes or to study specifically the translation elongation step. Their analysis yields that the attainable output p_1, p_2 of the two proteins satisfies the formula

$$\alpha p_1 + \beta p_2 = Y, \quad (1.1)$$

where Y is related to the total number of ribosomes (but also other translation factors and possibly additional gene expression factors), and α, β are constants that depend on parameters such as the plasmid copy number, dissociation constants of the ribosomes binding to the ribosomal binding site (RBS), etc. This equation implies that increasing the production of one protein always leads to a decrease in the production of the other protein (although more subtle correlations may take place via the effects on the constants α and β). A similar conclusion also has been derived for other models ([31], ch. 7).

In our model, improving the translation rate of a codon in one mRNA may either increase or decrease the translation rates of *all* other mRNAs in the cell. Indeed, the effect on the other genes depends on the change in the *total density* of ribosomes on the modified mRNA molecule, highlighting the importance of modelling the dynamics of the translation elongation step. We show, however, that when increasing the initiation rate in an RFM in the network, the total density in this RFM always increases and, consequently, the production rate in all the other RFMs decreases. This special case agrees with the results in [11].

Another recent study [32] showed that the hidden layer of interactions among genes arising from competition for shared resources can dramatically change network behaviour. For example, a cascade of activators can behave like an effective repressor, and a repression cascade can become bistable. This agrees with several previous studies in the field (e.g. [7,33]).

The remainder of this paper is organized as follows. Section 2 summarizes the biological implications of our theoretical results. This includes no equations, and was written in order to make this paper more accessible to a wider audience. Section 3 describes the new model, and demonstrates using several examples how it can be used to study translation at the *cell level*. Section 4 describes our main theoretical results, and details their biological implications. To streamline the presentation, the proofs of the results are in the electronic supplementary material. The final section concludes and describes several directions for further research, including suggestions for biological experiments inspired from the theoretical results.

2. Biological rationale and summary of the main results

The goal of our study is to understand the intricate coupling between the translation rates of different mRNA molecules induced by the competition for the finite pool of available resources and, in particular, ribosomes. We develop a new mathematical model that includes a finite pool of ‘free’ ribosomes, and an arbitrary number of mRNA molecules. This implies that the different mRNA molecules ‘compete’ for the finite pool of ribosomes (figure 1).

Our model encapsulates the fundamental dynamical aspects of the translation process: the codon decoding rates and initiation rates may vary among different genes, and within a coding region; the movement of ribosomes is unidirectional from the 5′ end to the 3′ of the coding region; when there are multiple ribosomes on the mRNA, a ribosome cannot overtake another ribosome, etc.

The initiation and codon translation rates in the model are related to various biophysical mechanisms and transcript features including adaptation of codons to the tRNA pool (codons that are recognized by a tRNA with a higher intracellular abundance tend to be translated more quickly [34]); local folding of the mRNA (stronger folding tends to decrease elongation rate [35]); interaction/hybridization between the ribosomal RNA and the mRNA [36] (there are nucleotide sub-sequences that tend to interact with the ribosomal RNA); the interaction between the nascent peptide and the exit tunnel of the ribosome [37,38]; and more. The initiation rates may also be related to various mechanisms and transcript features including ribosome binding sites in prokaryote [39]; mRNA folding near the start codon [2,40] and nucleotide composition surrounding the start codon in eukaryotes [41,42]. However, the theoretical results proved here hold for any set of feasible parameter values covering any possible physiological condition. For example, they hold when the initiation rate is very low (and thus rate limiting), but also when the initiation rate is high, causing ribosomal ‘traffic jams’ along the mRNA. Similarly, the results hold if the codon decoding rates are constant along different codons and when they significantly vary along the codons. Thus, the reported results are relevant for

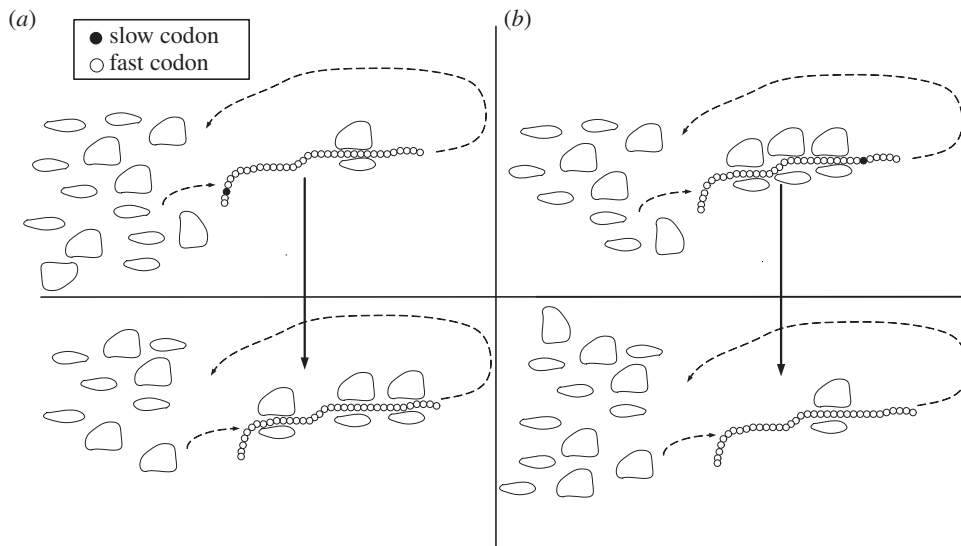


Figure 2. Local and global effects of improving the rate of a site in a specific mRNA. (a) Improving the *initiation rate* of an mRNA molecule increases its translation rate and also the number of ribosomes along this mRNA; owing to the increased ribosome density on the modified mRNA the pool is depleted and consequently the initiation and translation rates in all the other mRNA molecules *decrease*. Thus, the local and global effects are opposite. (b) A very slow codon in an mRNA molecule generates a ribosome traffic jam. Improving this codon improves the translation rate of this mRNA. In addition, the traffic jam disappears, so there are more free ribosomes in the pool, and thus the initiation and the translation rates of *all* the other mRNA molecules *increase*.

various organisms (e.g. bacteria, fungi, mammals) in various regimes and conditions (e.g. normal conditions, starvation, viral infection and diseases such as cancer, etc.). In addition, they are relevant both for endogenous and heterologous/engineered genes.

We show that the compound system always converges to a steady state. In other words, the ribosome density profile on every mRNA molecule and the number of ribosomes in the pool converge, as time goes to infinity, to some fixed value. We then analyse how this steady state changes as we change a transition rate in one of the mRNA molecules. Note that the change in one mRNA induces a change throughout the system because of the mutual coupling via the competition for the available ribosomes.

The first fundamental conclusion of our study describes the local and global effects of modifying the *initiation rate* in one mRNA molecule on the total protein production in the cell (figure 2a). We show that increasing [decreasing] the initiation rate in a certain mRNA molecule always increases [decreases] the ribosome density profile and the translation rate of that mRNA, and also decreases [increases] the ribosome density and translation rate in *all* the other mRNA molecules. In other words, the local and global effects are opposed. This conclusion is important in the context of biotechnology and synthetic biology, as it implies that increasing the initiation rate of an engineered (e.g. heterologous) mRNA molecule will lead to a decrease in the translation rates of all the other mRNA molecules. This is expected to yield a decrease in the growth rate of the host, and may eventually decrease the production rate of the heterologous gene. This conclusion is also relevant in the context of molecular evolution, as it shows that a mutation near the 5' end of the coding region in one mRNA affects the organismal fitness not only locally but also globally.

The second fundamental conclusion of our study considers increasing [decreasing] a codon elongation rate (but not the initiation rate) in a specific mRNA molecule. The local effect is *always* an increase [decrease] in the translation

rate of the modified mRNA. The global effect is homogeneous: either all the translation rates of the other mRNAs increase or they all decrease (figures 2b and 3b). Which of these two outcomes will indeed take place depends on all the rates in the system, and can be predicted using our model. We can provide an intuitive explanation of the two possible outcomes as follows. Suppose that the modified rate is a bottleneck rate, i.e. it leads to ribosomal traffic jams. Then increasing this rate reduces the traffic jams, thus freeing more ribosomes to the pool. This increases the initiation rates in *all* the other mRNAs leading to a global increase in translation rates (figure 2b).

On the other hand, suppose that the modified codon is located close to the beginning of the coding region and that there is a bottleneck codon, generating a traffic jam, further away along the coding region. Then an increase in the transition rate of the modified codon actually worsens the traffic jam. This depletes the pool, and thus yields a global decrease in the translation rates of all the other mRNAs (figure 3b).

This second conclusion is important in the context of biotechnology and synthetic biology, as it demonstrates that evaluating the effect of (heterologous or endogenous) gene engineering on its production rate should take into account the change in the ribosomal pool in order to estimate the global, and not only local, effect of the modification. Our model allows to predict this effect.

Specifically, consider the effect of introducing a 'ramp', i.e. a region with relatively slow codons at the beginning of the coding region [1,4]. This is expected to reduce the traffic jams in the modified mRNA because fewer ribosomes can enter and pile up in the modified mRNA. This then should increase the number of free ribosomes and thus the global effect is expected to be an increase in the initiation and translation rates in all the other mRNAs, leading to an increase in organismal growth rate and fitness (figure 3a).

This conclusion is also relevant in the context of molecular evolution, as it allows a better understanding of the *trans* effect of mutations in the coding region on the

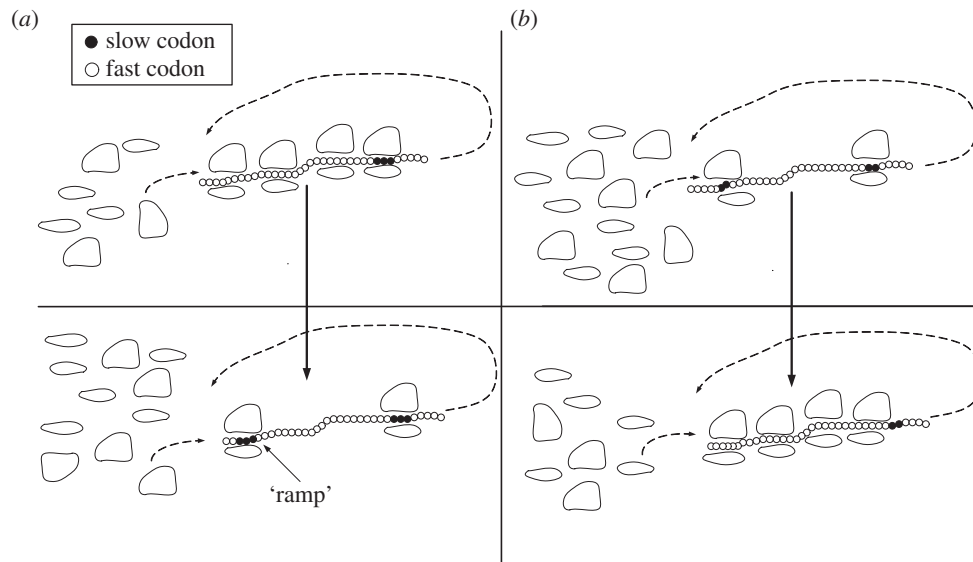


Figure 3. Local and global effects of modifying an mRNA molecule. (a) Introducing ramps, i.e. codons at the beginning of the coding sequence that are slower as (but not much slower than) the slowest codon in the molecule is expected to decrease traffic jams in the modified mRNA, thus increasing the number of free ribosomes, leading to a global increase in the initiation and translation rates in all other mRNA molecules. (b) Introducing a set of mutations in the coding region that improves the translation rates of the first codons, when there are slower codons afterward, is expected to increase the ribosome density on the modified mRNA, thus decreasing the number of free ribosomes, and leading to a global decrease in the initiation and translation rates of all other mRNA molecules.

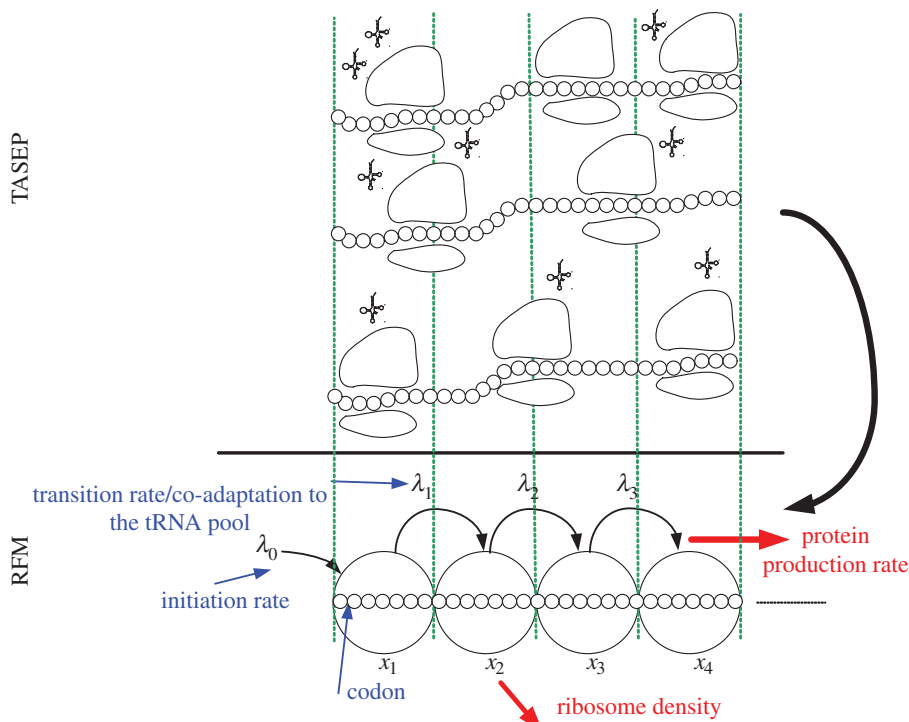


Figure 4. Topology of the RFM and comparison with the TASEP model. The RFM (lower part of the figure) is a mean-field approximation of the TASEP model (upper part of the figure). State variable $x_i(t) \in [0, 1]$ describes the normalized ribosome occupancy level in site i at time t . The initiation rate is λ_0 , and λ_i is the elongation rate between sites i and $i+1$. Production rate at time t is $R(t) := \lambda_n x_n(t)$. (Online version in colour.)

organismal fitness via the global, and not only local, effect of the modification.

Our model and its analysis also suggest several novel biological experiments. Some of these are described in §5.

3. The model and some examples

Since our model is based on a network of interconnected RFMs, we begin with a brief review of the RFM.

3.1. Ribosome flow model

The RFM models the traffic flow of ribosomes along the mRNA (figure 4). The mRNA chain is divided into a set of n compartments or sites, where each site may correspond to a codon or a group of codons. The state variable $x_i(t)$, $i = 1, \dots, n$, describes the ribosome occupancy at site i at time t , normalized such that $x_i(t) = 0$ [$x_i(t) = 1$] implies that site i is completely empty [full] at time t . Since $x_i(t)$ takes values in $[0, 1]$, one may also view $x_i(t)$ as the

probability that site i is occupied at time t . The dynamical equations of the RFM are

$$\left. \begin{aligned} \dot{x}_1 &= \lambda_0(1 - x_1) - \lambda_1 x_1(1 - x_2), \\ \dot{x}_2 &= \lambda_1 x_1(1 - x_2) - \lambda_2 x_2(1 - x_3), \\ \dot{x}_3 &= \lambda_2 x_2(1 - x_3) - \lambda_3 x_3(1 - x_4), \\ &\vdots \\ \dot{x}_{n-1} &= \lambda_{n-2} x_{n-2}(1 - x_{n-1}) - \lambda_{n-1} x_{n-1}(1 - x_n), \\ \dot{x}_n &= \lambda_{n-1} x_{n-1}(1 - x_n) - \lambda_n x_n. \end{aligned} \right\} \quad (3.1)$$

These equations describe the movement of ribosomes along the mRNA chain. The *transition rates* $\lambda_0, \dots, \lambda_n$ are all positive numbers (units = 1/time). To explain this model, consider the equation $\dot{x}_2 = \lambda_1 x_1(1 - x_2) - \lambda_2 x_2(1 - x_3)$. The term $\lambda_1 x_1(1 - x_2)$ represents the flow of particles from site 1 to site 2. This is proportional to the occupancy x_1 at site 1 and also to $1 - x_2$, i.e. the flow decreases as site 2 becomes fuller. In particular, if $x_2 = 1$, i.e. site 2 is completely full, the flow from site 1 to site 2 is zero. This is a ‘soft’ version of the rough exclusion principle in ASEP. Note that the maximal possible flow rate from site 1 to site 2 is the transition rate λ_1 . The term $\lambda_2 x_2(1 - x_3)$ represents the flow of particles from site 2 to site 3.

The dynamical equations for the other state variables are similar. Note that λ_0 controls the initiation rate into the chain, and that

$$R(t) := \lambda_n x_n(t)$$

is the rate of flow of ribosomes out of the chain, that is the translation (or protein production) rate at time t . The RFM topology is depicted in figure 4.

The RFM encapsulates simple exclusion and uni-directional movement along the lattice just as in ASEP. This is not surprising, as the RFM can be derived via a mean-field approximation of ASEP (e.g. [24], p. R345, and [43], p. 1919). However, the analysis of these two models is quite different, as the RFM is a deterministic, continuous-time, synchronous model, whereas ASEP is a stochastic, discrete-type, asynchronous one.

In order to study a network of interconnected RFMs, it is useful to first extend the RFM into a single-input single-output control system:

$$\left. \begin{aligned} \dot{x}_1 &= \lambda_0(1 - x_1)u - \lambda_1 x_1(1 - x_2), \\ \dot{x}_2 &= \lambda_1 x_1(1 - x_2) - \lambda_2 x_2(1 - x_3), \\ \dot{x}_3 &= \lambda_2 x_2(1 - x_3) - \lambda_3 x_3(1 - x_4), \\ &\vdots \\ \dot{x}_{n-1} &= \lambda_{n-2} x_{n-2}(1 - x_{n-1}) - \lambda_{n-1} x_{n-1}(1 - x_n), \\ \dot{x}_n &= \lambda_{n-1} x_{n-1}(1 - x_n) - \lambda_n x_n, \end{aligned} \right\} \quad (3.2)$$

and $y = \lambda_n x_n$.

Here the translation rate becomes the output y of the system, and the flow into site 1 is multiplied by a time-varying control $u: \mathbb{R}_+ \rightarrow \mathbb{R}_+$, representing the flow of ribosomes into the mRNA from the ‘outside world’. The time-varying function $u(t)$ is thus related to the rate ribosomes diffuse to the 5’ end (in eukaryotes) or the RBS (in prokaryotes) of the mRNA at time t . Of course, mathematically one can absorb λ_0 into u , but we do not do this because we think of λ_0 as representing some intrinsic local mRNA-specific features

(e.g. the strength of the Kozak sequences in eukaryotes or the RBS in prokaryote).

The set of admissible controls \mathcal{U} is the set of bounded and measurable functions taking values in \mathbb{R}_+ for all time $t \geq 0$. Equation (3.2), referred to as the RFM with input and output (RFMIO) [44], facilitates the study of RFMs with feedback connections. We note in passing that (3.2) is a *monotone control system* as defined in [45]. From now on we write (3.2) as

$$\text{and} \quad \left. \begin{aligned} \dot{x} &= f(x, u) \\ y &= \lambda_n x_n. \end{aligned} \right\} \quad (3.3)$$

Let

$$C^n := \{z \in \mathbb{R}^n: z_i \in [0, 1], i = 1, \dots, n\}$$

denote the closed unit cube in \mathbb{R}^n . Since the state variables in the RFM represent normalized occupancy levels, we always consider initial conditions $x(0) \in C^n$. It is straightforward to verify that C^n is an invariant set of (3.3), i.e. for any $u \in \mathcal{U}$ and any $x(0) \in C^n$, the trajectory satisfies $x(t, u) \in C^n$ for all $t \geq 0$. This means that if the initial density is well defined then it remains so for all time.

3.2. Ribosome flow model network with a pool

To model competition for ribosomes in the cell, we consider a set of $m \geq 1$ RFMIOs, representing m different mRNA molecules in the cell (figure 5). The i th RFMIO has length n_i , input function $u^i \in \mathcal{U}$, output function y^i and rates $\lambda_0^i, \dots, \lambda_{n_i}^i$. The dynamics of these RFMIOs is thus given by

$$\left. \begin{aligned} \dot{x}^1 &= f(x^1, u^1), \quad y^1 = \lambda_{n_1}^1 x_{n_1}, \\ &\vdots \\ \dot{x}^m &= f(x^m, u^m), \quad y^m = \lambda_{n_m}^m x_{n_m}. \end{aligned} \right\} \quad (3.4)$$

The m RFMIOs are interconnected through a pool of free ribosomes (i.e. ribosomes that are not attached to any mRNA molecule). We use $z(t)$ to denote the pool occupancy at time t (this may also be interpreted as the average number of free ribosomes in the pool at time t). The pool feeds the initiation location in all the mRNAs (figure 5). To model this we can simply take $u^i(t) = z(t)$ for all i (recall that u^i describes the rate at which ribosomes from the ‘outside world’ feed RFMIO # i), but we allow a somewhat more general setting modelled by

$$u^j(t) = G_j(z(t)), \quad j = 1, \dots, m. \quad (3.5)$$

We assume that each $G_j(\cdot): \mathbb{R}_+ \rightarrow \mathbb{R}_+$ satisfies the following properties: (i) $G_j(0) = 0$; (ii) G_j is continuously differentiable and $G_j'(z) > 0$ for all $z \geq 0$ (so G_j is strictly increasing on \mathbb{R}_+) and (iii) there exists $s > 0$ such that $G_j(z) \leq sz$ for all $z > 0$ sufficiently small. The first property implies that if the pool is empty then no ribosomes can bind to the mRNA molecules, and the second means that as the pool becomes fuller the initiation rates to the RFMIOs increase. The third property is a technical assumption that is needed for one of our proofs below.

Typical examples for functions satisfying these properties include the linear function, say, $G_j(z) = z$, and also the function $G_j(z) = a_j \tanh(b_j z)$, with $a_j, b_j > 0$. In the first case, the flow of ribosomes into the first site of RFM # i is given by $\lambda_0^i z(1 - x_1^i)$, and the product here can be justified via mass-action kinetics. The use of \tanh may be suitable for modelling a saturating function. This is in fact a standard function in ASEP models with a pool [21,46], because it is zero when z

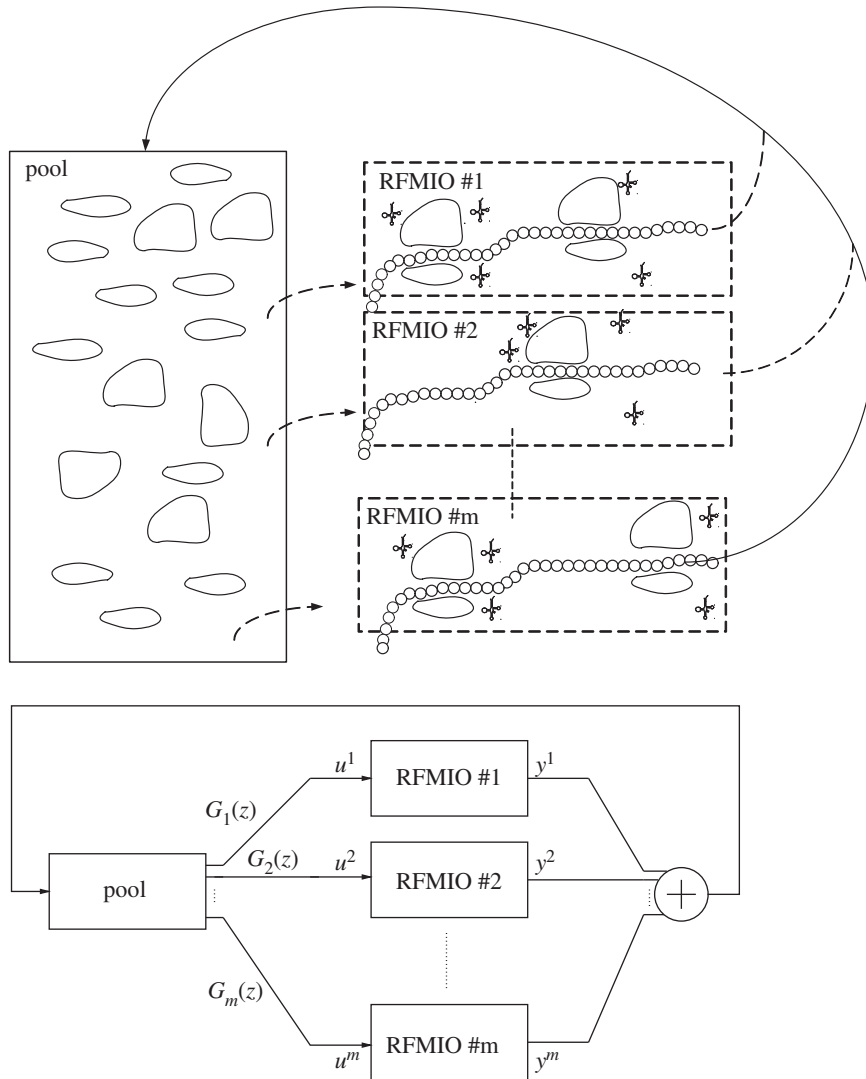


Figure 5. Topology of the RFMNP. Upper part: the RFMNP describes a set of mRNA molecules that are translated by a finite pool of ribosomes. Lower part: each mRNA is described by an RFMIO; the output of each RFMIO is fed to the pool and the pool feeds the initiation rate of each RFMIO. See details in the main text.

is zero, uniformly bounded and strictly increasing for $z \geq 0$. Also, for $z \geq 0$ the function $\tanh(z)$ takes values in $[0, 1)$ so it can also be interpreted as a probability function [47].

In the context of a shared pool, it is natural to consider the special case where $G_j(z) = G(z)$ for all $j = 1, \dots, m$. This assumption is reasonable since there is no *a priori* reason to expect a ‘preference’ between different mRNA species. However, the differences between the initiation sites in the mRNA molecules may still be modelled by taking different λ_0^j s.

The output of each RFMIO is fed into the pool, and the pool feeds the initiation locations in the mRNAs (figure 5). Thus, the pool dynamics is described by

$$\dot{z} = \sum_{j=1}^m y^j - \sum_{j=1}^m \lambda_0^j (1 - x_1^j) G_j(z). \quad (3.6)$$

The above equation means that the flow into the pool is the sum of all output rates $\sum_{j=1}^m y^j$ of the RFMIOs minus the total flow of ribosomes that bind to an mRNA molecule $\sum_{j=1}^m \lambda_0^j (1 - x_1^j) G_j(z)$. Recall that the term $(1 - x_1^j)$ represents the exclusion, i.e. as the first site in RFMIO # j becomes fuller, less ribosomes can bind to it. Summarizing, the RFMNP is given by equations (3.4), (3.5) and (3.6). This is a dynamical system with $d := 1 + \sum_{i=1}^m n_i$ state variables.

Note that combining the properties of the functions $G_j(\cdot)$ with (3.6) implies that if $z(0) \geq 0$ then $z(t) \geq 0$ for all $t \geq 0$. Thus, the pool occupancy is always non-negative.

Example 3.1. Consider a network with $m = 2$ RFMIOs, the first [second] with dimension $n_1 = 2$ [$n_2 = 3$]. Then the RFMNP is given by

$$\left. \begin{aligned} \dot{x}_1^1 &= \lambda_0^1 (1 - x_1^1) G_1(z) - \lambda_1^1 x_1^1 (1 - x_2^1), \\ \dot{x}_2^1 &= \lambda_1^1 x_1^1 (1 - x_2^1) - \lambda_2^1 x_2^1, \\ \dot{x}_1^2 &= \lambda_0^2 (1 - x_1^2) G_2(z) - \lambda_1^2 x_1^2 (1 - x_2^2), \\ \dot{x}_2^2 &= \lambda_1^2 x_1^2 (1 - x_2^2) - \lambda_2^2 x_2^2 (1 - x_3^2), \\ \dot{x}_3^2 &= \lambda_2^2 x_2^2 (1 - x_3^2) - \lambda_3^2 x_3^2 \end{aligned} \right\} \text{and } \dot{z} = \lambda_2^1 x_2^1 + \lambda_3^2 x_3^2 - \lambda_0^1 (1 - x_1^1) G_1(z) - \lambda_0^2 (1 - x_1^2) G_2(z). \quad (3.7)$$

Note that this system has $d = 6$ state variables. ■

An important property of the RFMNP is, that being a closed system, the total occupancy

$$H(t) := z(t) + \sum_{j=1}^m \sum_{i=1}^{n_j} x_i^j(t) \quad (3.8)$$

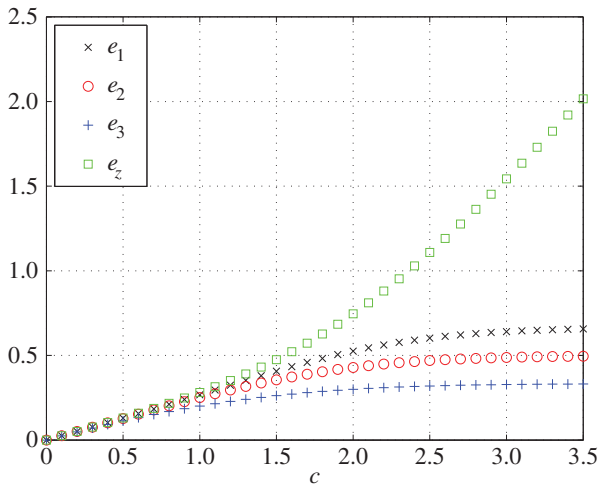


Figure 6. Steady-state values in the RFMNP in example 3.2 as a function of the total occupancy $H(0) = c$. (Online version in colour.)

is conserved (i.e. the total number of ribosomes is conserved), that is,

$$H(t) \equiv H(0), \quad \text{for all } t \geq 0. \quad (3.9)$$

In other words, H is a first integral of the dynamics. In particular, this means that $z(t) \leq H(t) = H(0)$ for all $t \geq 0$, i.e. the pool occupancy is uniformly bounded by the constant $H(0)$.

The RFMNP models mRNAs that compete for ribosomes because the total number of ribosomes is conserved. As more ribosomes bind to the RFMIOs, the pool depletes, $G_j(z)$ decreases and the effective initiation rate to all the RFMIOs decreases (figure 5). This allows to systematically address important biological questions on large-scale simultaneous translation under competition for ribosomes. The following examples demonstrate this. We prove in §4 that all the state variables in the RFMNP converge to a steady-state. Let $e_j^i \in [0, 1]$ denote the steady-state occupancy in site j in RFMIO $\#i$, and let $e_z \in [0, \infty)$ denote the steady-state occupancy in the pool. In the examples below we always consider these steady-state values (obtained numerically by simulating the differential equations).

Example 3.2. Although we are mainly interested in modelling large-scale simultaneous translation, it is natural to first consider a model with a single mRNA molecule connected to a pool of ribosomes. From a biological perspective, this models the case where there is one gene that is highly expressed with respect to all other genes (e.g. an extremely highly expressed heterologous gene).

Consider an RFMNP that includes a single RFMIO (i.e. $m = 1$), with dimension $n_1 = 3$, rates $\lambda_i^1 = 1$, $i = 0, 1, 2, 3$, and a pool with output function $G(z) = \tanh(z)$. We simulated this system for the initial condition $x_i^1(0) = 0$ for all i , and $z(0) = c$ for various values of c . Note that $H(t) \equiv H(0) = c$. Figure 6 depicts the steady-state values e_1, e_2, e_3 of the state-space variables in the RFMIO, and the steady-state pool occupancy e_z . It may be seen that for small values of c the steady-state ribosomal densities and thus the production rates are very low. This is simply because there are not enough ribosomes in the network. The ribosomal densities increase with c . For large values of c , the output function of the pool saturates, as $\tanh(z) \rightarrow 1$, and so does the initiation rate in the RFMIO. Thus, the densities

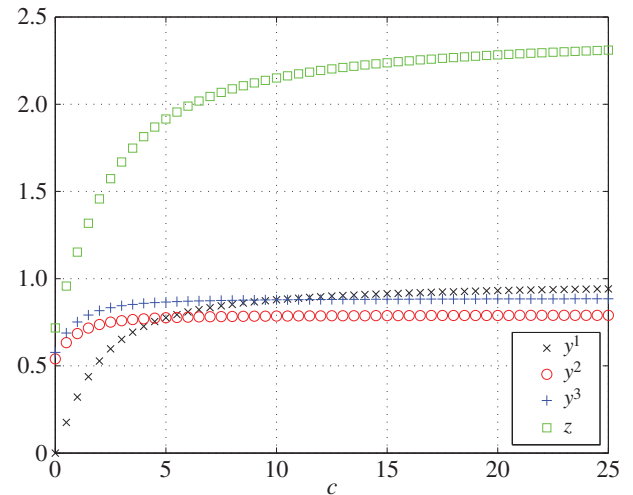


Figure 7. Steady-state outputs y^j of the three RFMIOs and pool occupancy z as a function of the homogeneous transition rate c in RFMIO $\#1$. (Online version in colour.)

in the RFMIO saturate to the values corresponding to the initiation rate $\lambda_0 = 1$, and then all the remaining ribosomes accumulate in the pool. Using a different pool output function, for example $G(z) = z$, leads to the same qualitative behaviour, but with higher saturation values for the ribosomal densities in the RFMIO. (Note that the ribosomal densities in an RFM are finite even when $\lambda_0 \rightarrow \infty$ [48].) ■

This simple example already demonstrates the coupling between the ribosomal pool, initiation rate and elongation rates. When the ribosomal pool is small the initiation rate is low. Thus, the ribosomal densities on the mRNA are low and there are no interactions between ribosomes (i.e. no ‘traffic jams’) along the mRNA. The initiation rate becomes the rate limiting step of translation. On the other hand, when there are many ribosomes in the pool the initiation rate increases, the elongation rates become rate limiting and ‘traffic jams’ along the mRNA evolve. At some point, a further increase in the number of ribosomes in the pool will have a negligible effect on the production rate.

It is known that there can be very large changes in the number of ribosomes in the cell during, e.g. exponential growth. For example, Bremer & Dennis [49] report changes in the range 6800–72 000. The example above demonstrates how these large changes in the number of ribosomes are expected to affect the translational regimes; specifically, it may cause a switch between the different regimes mentioned above.

The next example describes an RFMNP with several mRNA chains. Let $\mathbf{1}_n \in \mathbb{R}^n$ denote the vector of n ones.

Example 3.3. Consider an RFMNP with $m = 3$ RFMIOs of dimensions $n_1 = n_2 = n_3 = 3$, and rates

$$\lambda_i^1 = c, \quad \lambda_i^2 = 5 \quad \text{and} \quad \lambda_i^3 = 10, \quad i = 0, \dots, 3.$$

In other words, every RFMIO has homogeneous rates. Suppose also that $G_i(z) = \tanh(z)$, for $i = 1, 2, 3$. We simulated this RFMNP for different values of c with the initial condition $z(0) = 0$, $x^1(0) = (1/2)\mathbf{1}_3$, $x^2(0) = (1/3)\mathbf{1}_3$ and $x^3(0) = (1/4)\mathbf{1}_3$. Thus, $H(0) = 3.25$ in all the simulations. For each value of c , every state variable in the RFMNP converges to a steady state. Figure 7 depicts the steady-state value e_z and the steady-state output y^j in each RFMIO. It may be seen that increasing c , i.e. increasing all the elongation

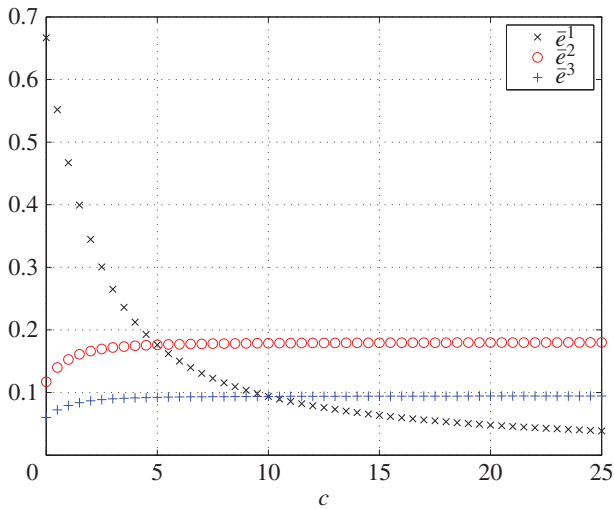


Figure 8. Average steady-state occupancy in the three RFMIOs as a function of the homogeneous transition rate c in RFMIO #1. (Online version in colour.)

rates in RFMIO #1 leads to an increase in the steady-state translation rates in *all* the RFMIOs in the network. Also, it leads to an increase in the steady-state occupancy of the pool. It may seem that this contradicts (3.9) but this is not so. Increasing c indeed increases all the steady-state translation rates, but it decreases the steady-state occupancies inside each RFMIO so that the total $H(t) = H(0) = 3.25$ is conserved.

Let $\bar{e}^j := (1/n_j) \sum_{i=1}^{n_j} e_i^j$, i.e. the averaged steady-state occupancy (ASSO) on RFMIO # j . Figure 8 depicts the ASSO in each RFMIO as a function of c . It may be seen as c increases the ASSO in RFMIO #1 decreases quickly, yet the ASSOs in the other two RFMIOs slowly increase. Indeed, since the ribosomes spend less time on RFMIO #1 (due to increased c) they are now available for translating the other RFMIOs, leading to the increased ASSO in the other mRNAs. ■

From a biological point of view this example corresponds to a situation where accelerating *one* of the mRNA chains increases the protein production rates in *all* the mRNAs and also increases the number of free ribosomes. Surprisingly, perhaps, it also suggests that a relatively larger number of free ribosomes in the cell corresponds to higher protein production rates. This agrees with evolutionary, biological and synthetic biology studies that have suggested that (specifically) highly expressed genes (that are transcribed into many mRNA molecules) undergo selection to include codons with improved elongation rates [1,2,50]. Specifically, two mechanisms by which improved codons affect translation efficiency and the organismal fitness are the following [50]: (i) *global mechanism*: selection for improved codons contributes toward improved ribosomal recycling and global allocation; the increased number of free ribosomes improves the effective translation initiation rate of all genes, and thus improves global translation efficiency, and (ii) *local mechanism*: the improved translation elongation rate of an mRNA contributes directly to its protein production rate.

The example above demonstrates both mechanisms, as improvement of the translation elongation rates of one RFM increases the translation rate of this mRNA (local translation efficiency), and also of the other RFMs (global translation efficiency). In addition, as can be seen, the decrease in ASSO in RFMIO #1 is significantly higher than the increase in ASSO in

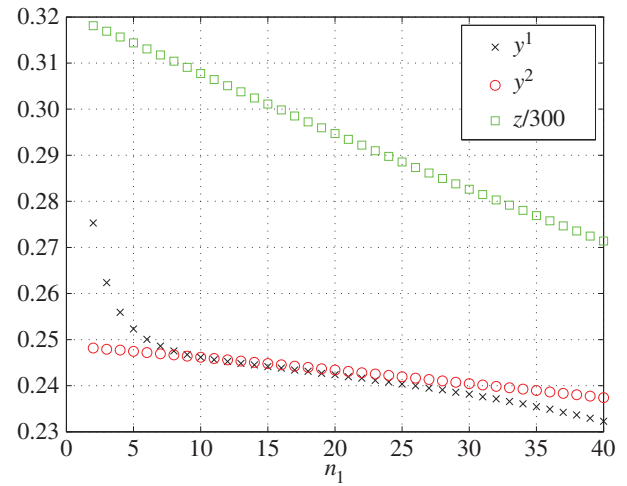


Figure 9. Steady-state outputs y^j of the two RFMIOs and steady-state pool occupancy $z/300$ as a function of the length n_1 of RFMIO #1. The normalization of z by 300 is used to obtain similar magnitudes for all the values only. (Online version in colour.)

the other RFMIO. Thus, the simulation also demonstrates that increasing the translation rate c may contribute to decreasing ribosomal collision (and possibly ribosomal abortion).

We prove in §4 that when one of the rates in one of the RFMIOs increases two outcomes are possible: either all the production rates in the other RFMs increase (as in this example) or they all decrease. As discussed below, we believe that this second case is less likely to occur in endogenous genes, but may occur in heterologous gene expression.

The next example describes the effect of changing the length of one RFMIO in the network. Let 0_n denote a vector of n zeros.

Example 3.4. Consider an RFMNP with $m = 2$ RFMIOs of dimensions n_1 and $n_2 = 10$, rates

$$\begin{aligned} \lambda_i^1 &= 1, & i &= 0, \dots, n_1, \\ \lambda_j^2 &= 1, & j &= 0, \dots, 10 \end{aligned}$$

and $G_i(z) = \tanh(z/200)$, $i = 1, 2$. In other words, both RFMIOs have the same homogeneous rates. We simulated this RFMNP for different values of n_1 with the initial condition $z(0) = 100$, $x^1(0) = 0_{n_1}$ and $x^2(0) = 0_{10}$. Thus, $H(0) = 100$ in all the simulations. For each value of n_1 , every state variable in the RFMNP converges to a steady state. Figure 9 depicts the steady-state values of z , and the steady-state output y^j in each RFMIO. It may be seen that increasing n_1 , i.e. increasing the length of RFMIO #1 leads to a decrease in the steady-state production rates and in the steady-state pool occupancy. This is reasonable, as increasing n_1 means that ribosomes that bind to the first chain remain on it for a longer period of time. This decreases the production rate y^1 and, by the competition for ribosomes, also decreases the pool occupancy and thus decreases y^2 . ■

From a biological point of view this suggests that decreasing the length of mRNA molecules contributes locally and globally to improving translation efficiency. A shorter coding sequence improves the translation rate of the mRNA and, by competition, may also improve the translation rates in all other mRNAs. Thus, we should expect to see selection

for shorter coding sequences, specifically in highly expressed genes and in organisms with large population size. Indeed, previous studies have reported that in some organisms the coding regions of highly expressed genes tend to be shorter [14]; other studies have shown that other (non-coding) parts of highly expressed genes tend to be shorter [51–54]. Decreasing the length of different parts of the gene should contribute to organismal fitness via improving the energetic cost of various gene expression steps. For example, shorter genes should improve the metabolic cost of synthesizing mRNA and proteins; it can also reduce the energy spent for splicing and processing of RNA and proteins. However, there are of course various functional and regulatory constraints that also contribute to shaping the gene length (e.g. [55]). Our results and these previous studies suggest that in some cases genes are expected to undergo selection also for short *coding* regions, as this reduces the required number of translating ribosomes.

The next section describes various fundamental properties related to the RFMNP. All the proofs are placed in the Appendix (see the electronic supplementary material).

4. Properties of the ribosome flow model network with a pool

Let

$$\Omega := [0, 1]^{n_1} \times \dots \times [0, 1]^{n_m} \times [0, \infty)$$

denote the state space of the RFMNP (recall that every x_i^j takes values in $[0, 1]$ and $z \in [0, \infty)$). For an initial condition $a \in \Omega$, let $[x(t, a) \ z(t, a)]'$ denote the solution of the RFMNP at time t . It is straightforward to show that the solution remains in Ω for all $t \geq 0$.

4.1. Stability

For $s \geq 0$, let L_s denote a level set of the first integral H :

$$L_s := \{y \in \Omega: 1_d' y = s\}.$$

In other words, $y \in L_s$ means that y is a condition corresponding to a total occupancy of s ribosomes in the system.

Theorem 4.1. *Every level set L_s , $s \geq 0$, contains a unique equilibrium point e_{L_s} of the RFMNP, and for any initial condition $a \in L_s$, the solution of the RFMNP converges to e_{L_s} . Furthermore, for any $0 \leq s < p$,*

$$e_{L_s} \ll e_{L_p}. \quad (4.1)$$

In particular, this means that every trajectory converges to an equilibrium point, representing steady-state ribosomal densities in the RFMIOs and the pool. Equation (4.1) means that the continuum of equilibrium points, namely, $\{e_{L_s}: s \in [0, \infty)\}$, are linearly ordered. In other word, given any two equilibrium points, e_{L_s} and e_{L_p} , with $p > s$, then every entry in e_{L_s} is strictly smaller than the corresponding entry in e_{L_p} . In other words, if the RFMNP is initiated with two initial conditions, with the first one corresponding to a higher number of ribosomes in the network than the second then the steady-state corresponding to the first condition will have a higher density than the

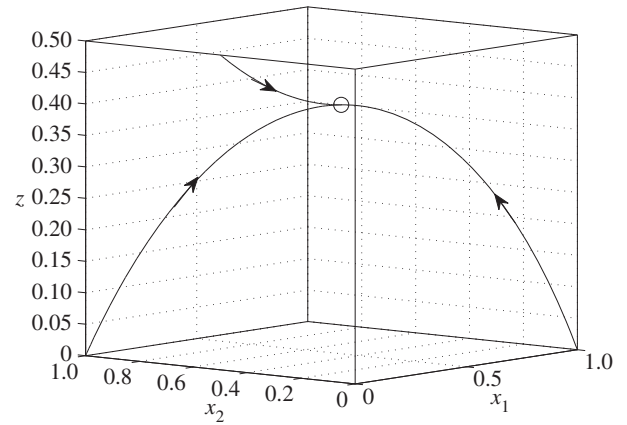


Figure 10. Trajectories of the system in example 4.2 for three initial conditions in L_1 . The equilibrium point (4.3) is marked by a circle.

steady state corresponding to the second in each site in each mRNA and in the pool.

It follows from proposition 4.8, stated in §4.3 below, that for any $s > 0$, $e_{L_s} \in \text{Int}(\Omega)$. In other words, the steady state densities will never include a density that is either zero or one, and the steady-state pool occupancy is always strictly positive.

Example 4.2. Consider an RFMNP with $m = 2$ RFMIOs with dimensions $n_1 = n_2 = 1$, and $G_i(z) = z$, $i = 1, 2$, i.e.

$$\left. \begin{aligned} \dot{x}_1^1 &= \lambda_0^1(1 - x_1^1)z - \lambda_1^1 x_1^1, \\ \dot{x}_1^2 &= \lambda_0^2(1 - x_1^2)z - \lambda_1^2 x_1^2, \\ \text{and } \dot{z} &= \lambda_1^1 x_1^1 + \lambda_1^2 x_1^2 - \lambda_0^1(1 - x_1^1)z - \lambda_0^2(1 - x_1^2)z. \end{aligned} \right\} \quad (4.2)$$

Note that even in this simple case the RFMNP is a non-linear system. Assume that $\lambda_0^1 = \lambda_0^2 = 1$, and that $\lambda_1^1 = \lambda_1^2$, and denote this value simply by λ . Pick an initial condition in Ω , and let $s := x_1^1(0) + x_1^2(0) + z(0)$, so that the trajectory belong to L_s for all $t \geq 0$. Any equilibrium point $e = [e_1 \ e_2 \ e_z]'$ $\in L_s$ satisfies

$$(1 - e_1)e_z = \lambda e_1,$$

$$(1 - e_2)e_z = \lambda e_2$$

and

$$e_1 + e_2 + e_z = s.$$

This yields two solutions

$$\left. \begin{aligned} e_1 = e_2 &= \frac{(s + 2 + \lambda - \sqrt{(s + 2 + \lambda)^2 - 8s})}{4} \\ \text{and } e_z &= \frac{(s - 2 - \lambda + \sqrt{(s + 2 + \lambda)^2 - 8s})}{2} \end{aligned} \right\} \quad (4.3)$$

and

$$\left. \begin{aligned} e_1 = e_2 &= \frac{(s + 2 + \lambda + \sqrt{(s + 2 + \lambda)^2 - 8s})}{4} \\ \text{and } e_z &= \frac{(s - 2 - \lambda - \sqrt{(s + 2 + \lambda)^2 - 8s})}{2}. \end{aligned} \right\}$$

It is straightforward to verify that in the latter solution $e_z < 0$, so this is not a feasible solution. The solution (4.3) does belong to L_s , so the system admits a unique equilibrium in L_s . Figure 10 depicts trajectories of (4.2) for three initial conditions

in L_1 , namely $[1 \ 0 \ 0]$, $[0 \ 1 \ 0]$ and $[0 \ 1/2 \ 1/2]$, and the equilibrium point (4.3) for $s = \lambda = 1$. It may be seen that every one of these trajectories converges to e . ■

Various possible intracellular mechanisms may affect any of the transition rates. For example, synonymous mutations/changes (in endogenous or heterologous) genes inside the coding region may modify the adaptation of codons to the tRNA pool (codons that are recognized by tRNAs with higher intracellular abundance usually tend to be translated more quickly [34]); the local folding of the mRNA (stronger folding tends to decrease elongation rates [35]); the interaction/hybridization between the ribosomal RNA and the mRNA [36] (there are nucleotide sub-sequences that tend to interact with the ribosomal RNA, causing transient pausing of the ribosome and thus reducing the translation elongation rate). Non-synonymous mutations or changes inside the coding region may also affect the elongation rates, for example via the interaction between the nascent peptide and the exit tunnel of the ribosome [37,38]. In addition, intracellular changes in various translation factors (e.g. tRNA levels, translation elongation factors, concentrations of amino acids, concentrations of aminoacyl tRNA synthetase) and, as explained above, the mRNA levels can also affect elongation rates. Furthermore, various recent studies have demonstrated that manipulating the codons of a heterologous gene tend to result in significant changes in the translation rates and protein levels of the gene [2,40,56]. Because of the competition for the finite pool of ribosomes a change in one molecule will also affect the decoding of other mRNA molecules. The next subsection describes the main result in the paper, namely, the local and global effect of mutations in one mRNA on the others resulting from the competition for the finite resources.

4.2. Competition

A natural question is how will a change in the parameters (that is, the transition rates) of the RFMNP affect the equilibrium point of the network (and, in particular, the steady-state production rates). For example, if we increase some transition rate λ_i^j in RFMIO # j , how will this affect the steady-state production rate in the other RFMIOs? Of course, this effect is due to the competition for the finite resources. Without loss of generality, we assume that the change is in a transition rate of RFMIO #1.

Theorem 4.3. Consider an RFMNP with m RFMIOs with dimensions n_1, \dots, n_m . Let $\lambda := [\lambda_0^1 \dots \lambda_{n_m}^m]^T$ denote the set of all parameters of the RFMNP, and let

$$e = [e_1^1 \ \dots \ e_{n_1}^1 \ e_1^2 \ \dots \ e_{n_2}^2 \ \dots \ e_1^m \ \dots \ e_{n_m}^m \ e_z]^T \\ \in (0, 1)^{n_1 + \dots + n_m} \times \mathbb{R}_{++}$$

denote the equilibrium point of the RFMNP on some fixed level set of H . Pick $i \in \{0, \dots, n_1\}$. Consider the RFMNP obtained by modifying λ_i^1 to $\bar{\lambda}_i^1$, with $\bar{\lambda}_i^1 > \lambda_i^1$. Let \bar{e} denote the equilibrium point in the new RFMNP and let $\bar{e} := \bar{e} - e$. Then

$$\bar{e}_i^1 < 0, \quad (4.4)$$

$$\bar{e}_j^1 > 0, \quad \text{for all } j \in \{i+1, \dots, n_1\} \quad (4.5)$$

$$\text{and } \text{sign}(\bar{e}_i^j) = \text{sign}(\bar{e}_z), \quad \text{for all } i \neq 1 \text{ and all } j. \quad (4.6)$$

(In the case $i = 0$, condition (4.4) is vacuous.)

Increasing λ_i^1 means that ribosomes flow ‘more easily’ from site i to site $i+1$ in RFMIO #1. Equation (4.4) means that the effect on the density in this RFMIO is that the number of ribosomes in site i decreases, whereas (4.5) implies that the number of ribosomes in all the sites to the right of site i increases. Equation (4.6) describes the effect on the steady-state densities in all the other RFMIOs and the pool: either all these steady-state values increase or they all decrease. The first case agrees with the results in example 3.3 above.

Note that the theorem does not provide any information on the change in e_j^1 , $j < i$. Our simulations show that any of these values may either increase or decrease, with the outcome depending on the various parameter values. Thus, the amount of information provided by (4.5) depends on i . In particular, when $\lambda_{n_1}^1$ is changed to $\bar{\lambda}_{n_1}^1 > \lambda_{n_1}^1$, then the information provided by (4.5) is only that

$$\bar{e}_{n_1}^1 < 0.$$

Much more information is available when $i = 0$.

Corollary 4.4. Suppose that λ_0^1 is changed to $\bar{\lambda}_0^1 > \lambda_0^1$. Then

$$\bar{e}_j^1 > 0, \quad \text{for all } j \in \{1, \dots, n_1\} \quad (4.7)$$

and

$$\bar{e}_i^j < 0, \quad \text{for all } i \neq 1 \text{ and all } j, \text{ and } \bar{e}_z < 0. \quad (4.8)$$

Indeed, for $i = 0$, (4.5) yields (4.7). Also, we know that the changes in the densities in all other RFMIOs and the pool have the same sign. This sign cannot be positive, as combining this with (4.7) contradicts the conservation of ribosomes, so (4.8) follows.

In other words, increasing λ_0^1 , the intrinsic initiation rate in RFMIO #1, yields an increase in all the densities in RFMIO #1, and a decrease in all the densities in all the other RFMIOs. This makes sense, as increasing λ_0^1 means that it is easier for ribosomes to bind to the mRNA molecule. This increases the total number of ribosomes along this molecule and, by competition, decreases all the densities in the other molecules and the pool. Note that this special case agrees well with the results described in [11] (see (1.1)). Thus, our model allows studying fundamental biological phenomena that are not covered by models that ignore the elongation dynamics.

Example 4.5. Consider the RFMNP in (3.7) with $G_i(z) = z$, $\lambda_0^1 = \lambda_0^2 = 1$, $\lambda_1^2 = \lambda_2^2 = 0.1$, $\lambda_3^2 = 1$ and initial condition $(1/4)1_6$. We consider a range of values for λ_2^1 . For each fixed value, we simulated the dynamics until steady state for two cases: $\lambda_1^1 = 1$ and $\bar{\lambda}_1^1 = 10$. Figure 11 depicts \bar{e}_1^1, \bar{e}_2^1 for the various fixed values of λ_2^1 . It may be seen that we always have $\bar{e}_1^1 < 0$ and $\bar{e}_2^1 > 0$. Figure 12 depicts \bar{e}_i^2 , $i = 1, 2, 3$, and \bar{e}_z for the various fixed values of λ_2^1 . It may be seen that for a small value of λ_2^1 all the \bar{e}_i^2 's and \bar{e}_z are negative, whereas for large values of λ_2^1 they all become positive. ■

Intuitively, this can be explained as follows. When λ_2^1 is small it is the bottleneck rate in RFMIO #1, and increasing λ_1^1 only generates more ‘traffic jams’ along RFMIO #1. This depletes the pool, and thus decreases the production rate in the second RFMIO. On the other hand, when λ_2^1 is large λ_1^1 becomes the bottleneck rate, and increasing it to $\bar{\lambda}_1^1$ allows ribosomes to traverse RFMIO #1 more quickly. This decreases

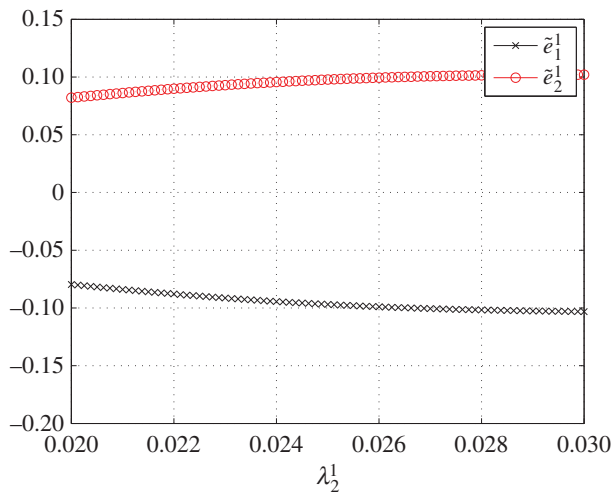


Figure 11. Steady-state perturbations in RFM #1 in the RFMNP in example 4.5 when changing $\lambda_1^1 = 1$ to $\bar{\lambda}_1^1 = 10$ for various values of λ_2^1 . (Online version in colour.)

the number of ribosomes on this RFMIO, increases the pool occupancy, and this leads to a higher initiation rate in the second RFMIO.

Theorem 4.3 implies that when the codons of a gene are modified into ‘faster’ codons (either via synthetic engineering or during evolution) then the translation rates of the other genes either all increase or all decrease. However, theorem 4.3 does not provide information on when each of these two cases actually happens. In order to address this, we need to calculate derivatives of the equilibrium point coordinates with respect to the rates. The next result shows that these derivatives are well defined. Denote the mapping from the parameters to the unique equilibrium point in $\text{Int}(\Omega)$ by α , that is, $e_j^i = \alpha_j^i(\lambda, H(0))$, $i = 1, \dots, m$, $j = 1, \dots, n_i$. (Recall that λ is the vector of all the parameters in the RFMIOs.)

Proposition 4.6. *The derivative $(\partial/\partial\lambda_q^p)\alpha_j^i(\lambda, H(0))$ exists for all i, j, p, q .*

The next example uses these derivatives to obtain information on the two cases that can take place as we change one of the rates.

Example 4.7. Consider an RFMNP with $m = 2$ RFMIOs with lengths n and ℓ . To simplify the notation, let $e = [e_1, \dots, e_n]$ [$v = [v_1, \dots, v_\ell]$] denote the equilibrium point of RFMIO #1 [RFMIO #2], and let λ_i , $i = 0, \dots, n$, denote the rates along RFMIO #1. Suppose that λ_1 is changed to $\bar{\lambda}_1$. Differentiating the steady-state equations

$$\lambda_0 G_1(e_z)(1 - e_1) = \lambda_1 e_1(1 - e_2) = \dots = \lambda_{n-1} e_{n-1}(1 - e_n) = \lambda_n e_n,$$

$$\sum_{i=1}^n e_i + \sum_{j=1}^{\ell} v_j + e_z = H(0),$$

w.r.t. λ_1 yields

$$\lambda_0 G_1'(e_z) e_z' (1 - e_1) - \lambda_0 G_1(e_z) e_1' = \lambda_n e_n',$$

$$e_z' + \sum_{j=1}^{\ell} v_j' = - \sum_{i=1}^n e_i',$$

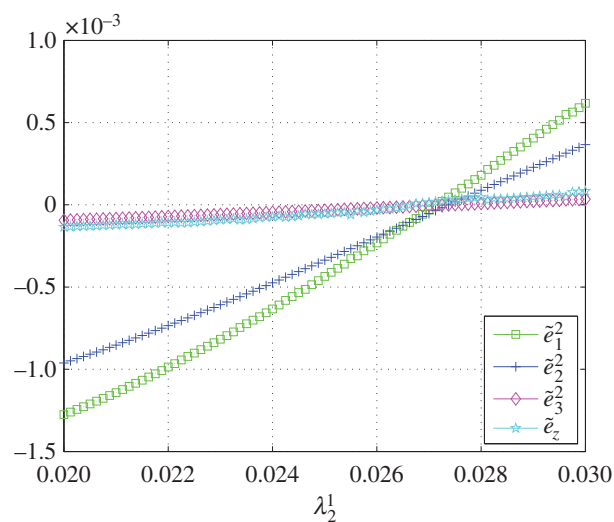


Figure 12. Steady-state perturbations in RFM #2 in the RFMNP in example 4.5 when changing $\lambda_1^1 = 1$ to $\bar{\lambda}_1^1 = 10$ for various values of λ_2^1 . (Online version in colour.)

where we use the notation $f' := (\partial/\partial\lambda_1)f$. These two equations yield

$$(\lambda_n + \lambda_0 G_1'(e_z)(1 - e_1)) e_z' + \lambda_n \sum_{j=1}^{\ell} v_j'$$

$$= (\lambda_0 G_1(e_z) - \lambda_n) e_1' - \lambda_n \sum_{i=2}^{n-1} e_i'.$$

Recall that $G_1(e_z) > 0$, $G_1'(e_z) > 0$, $\lambda_j > 0$ for all j , and $0 < e_p < 1$ for all p . Also, by theorem 4.3, $e_1' < 0$, $e_j' > 0$, for all $j \in 2, \dots, n$, and $\text{sign}(e_z') = \text{sign}(v_k')$, for all k . Thus,

$$\text{sign}(e_z') = \text{sign}(v_j')$$

$$= \text{sign} \left((\lambda_0 G_1(e_z) - \lambda_n) e_1' - \lambda_n \sum_{i=2}^{n-1} e_i' \right). \quad (4.9)$$

This means that the sign of the change in the densities in all the other RFMIOs and the pool depends on several steady-state quantities including terms related to the initiation rate $\lambda_0 G_1(e_z)$ and exit rate λ_n in RFMIO #1, and also the change in the total density $\sum_{i=1}^{n-1} e_i'$ in this RFMIO.

In the particular case $n = 2$ (i.e. a very short RFM), equation (4.9) becomes

$$\text{sign}(e_z') = \text{sign}(v_j') = \text{sign}(\lambda_2 - \lambda_0 G_1(e_z)). \quad (4.10)$$

Note that $\lambda_0 G_1(e_z)$ [λ_2] is the steady-state initiation [exit] rate in RFMIO #1. Thus, $\lambda_2 - \lambda_0 G_1(e_z) > 0$ means that it is ‘easier’ for ribosomes to exit than to enter RFMIO #1, and in this case (4.10) means that when λ_1 is increased the change in all other densities will be positive. This is intuitive, as more ribosomes will exit the modified molecule and this will improve the production rates in the other molecules. On the other hand, if $\lambda_2 - \lambda_0 G_1(e_z) < 0$, then it is ‘easier’ for ribosomes to enter than to exit RFMIO #1, so increasing λ_1 will lead to an increased number of ribosomes in RFMIO #1 and, by competition, to a decrease in the production rate in all the other RFMIOs. ■

Note that in the example above, increasing λ_1 always increases the steady-state production rate $R = \lambda_2 e_2$ in RFM #1 (recall that $e_2' > 0$). One may expect that this will always lead to an increase in the production rate in the second

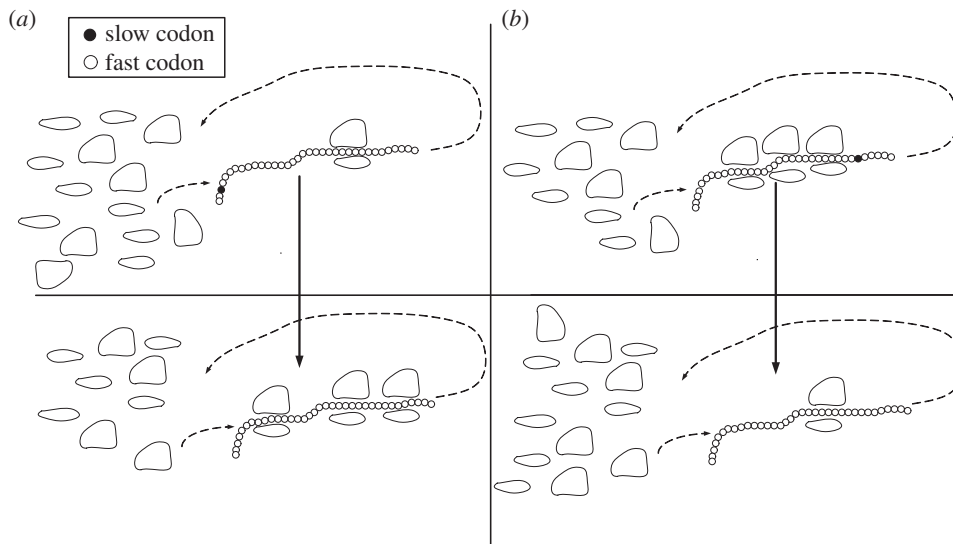


Figure 13. Increasing the transition rate of a site may have two different global effects. (a) Increasing the initiation rate of an mRNA molecule leads to an increase in the number of ribosomes along this mRNA; the number of free ribosomes in the pool decreases, and thus the initiation and translation rates of all the other mRNA molecules *decrease*. (b) Improving a slow bottleneck codon, alleviates the ‘traffic jam’ along this mRNA; more ribosomes are freed to the pool, and this improves the translation rate of all the other mRNAs.

RFMIO as well. However, the behaviour in the RFMNP is more complicated because the shared pool generates a feedback connection between the RFMIOs in the network. In particular, the effect on the other RFMIOs depends not only on the modified production rate of RFMIO #1, but also on other factors including the change in the *total* ribosome density in RFMIO #1 (see (4.9)).

This analysis of a very short RFM suggests that the steady-state initiation rate of the mRNA with the modified codon plays an important role in determining the effect of modifications in the network. If this initiation rate is relatively low (so it becomes the rate limiting factor), as believed to be the case in most endogenous genes [57], then the increase in the rate of one codon of the mRNA increases the translation rate in all the other mRNAs, whereas when this initiation rate is high then the opposite effect is obtained. This latter case may occur, for example, when a heterologous gene is highly expressed and thus ‘consumes’ some of the available elongation/termination factors making the elongation rates the rate limiting factors.

To recap, figure 13 describes the biological implications of our analysis. Making a codon ‘faster’ has two possible global outcomes. Figure 13a depicts the case of *increasing* the initiation rate of an mRNA molecule: the translation rate in the modified mRNA and also the number of ribosomes on the mRNA will increase; due to the increased ribosome density on the mRNA there are less free ribosomes in the pool, and thus the initiation and translation rates of all other mRNA molecules *decrease*. Figure 13b describes the second possible outcome. Improving a slow codon, that is a bottleneck rate in an mRNA, alleviates the ‘traffic jam’ along this mRNA, so the number of free ribosomes in the pool increases. This improves the translation rate of all other the mRNAs.

We now describe other mathematical properties of the RFMNP.

4.3. Persistence

The next result shows that for any initial condition, the following property holds. After an arbitrarily short time all the densities are larger than zero and smaller than one, and the pool occupancy is strictly positive.

Proposition 4.8. For any $\tau > 0$ there exists $\varepsilon = \varepsilon(\tau) > 0$, with $\varepsilon(\tau) \rightarrow 0$ when $\tau \rightarrow 0$, such that for all $t \geq \tau$, all $j = 1, \dots, m$, all $i = 1, \dots, n_j$ and all $a \in (\Omega \setminus \{0\})$,

$$\varepsilon \leq x_i^j(t, a) \leq 1 - \varepsilon$$

and

$$\varepsilon \leq z(t, a).$$

In other words, after any time $\tau > 0$ the solution is ε -separated from the boundary of Ω . This result is useful because on the boundary of Ω , denoted $\partial\Omega$, the RFMNP loses some desirable properties. For example, its Jacobian matrix may become reducible on $\partial\Omega$. Proposition 4.8 allows us to overcome this technical difficulty, as it implies that any trajectory is separated from the boundary after an arbitrarily short time.

4.4. Strong monotonicity

Recall that a cone $K \subseteq \mathbb{R}^n$ defines a partial order in \mathbb{R}^n as follows. For two vectors $a, b \in \mathbb{R}^n$, we write $a \leq b$ if $(b - a) \in K$; $a < b$ if $a \leq b$ and $a \neq b$; and $a \ll b$ if $(b - a) \in \text{Int}(K)$. A dynamical system $\dot{x} = f(x)$ is called *monotone* if $a \leq b$ implies that $x(t, a) \leq x(t, b)$ for all $t \geq 0$. In other words, monotonicity means that the flow preserves the partial ordering [58]. It is called *strongly monotone* if $a < b$ implies that $x(t, a) \ll x(t, b)$ for all $t > 0$.

From here on we consider the particular case where the cone is $K = \mathbb{R}_+^n$. Then $a \leq b$ if $a_i \leq b_i$ for all i , and $a \ll b$ if $a_i < b_i$ for all i . A system that is monotone with respect to this partial order is called *cooperative*.

The next result analyses the cooperativity of the RFMNP. Let $d := 1 + \sum_{i=1}^m n_i$ denote the dimension of the RFMNP.

Proposition 4.9. For any $a, b \in \Omega$ with $a \leq b$,

$$x(t, a) \leq x(t, b) \quad \text{and} \quad z(t, a) \leq z(t, b), \quad \text{for all } t \geq 0. \quad (4.11)$$

Furthermore, if $a < b$ then

$$x(t, a) \ll x(t, b) \quad \text{and} \quad z(t, a) < z(t, b), \quad \text{for all } t > 0. \quad (4.12)$$

This means the following. Consider the RFMNP initiated with two initial conditions such that the ribosomal densities in every site and the pool corresponding to the first initial condition are smaller or equal to the densities in the second initial condition. Then this correspondence between the densities remains true for all time $t \geq 0$.

4.5. Contraction

Contraction theory is a powerful tool for analysing nonlinear dynamical systems (e.g. [59]), with applications to many models from systems biology [60–62]. In a contractive system, the distance between any two trajectories decreases at an exponential rate. It is clear that the RFMNP is not a contractive system on Ω , with respect to any norm, as it admits more than a single equilibrium point. Nevertheless, the next result shows that the RFMNP is *non-expanding* with respect to the ℓ_1 norm defined by $|q|_1 = \sum_{i=1}^d |q_i|$.

Proposition 4.10. For any $a, b \in \Omega$,

$$\left| \begin{bmatrix} x(t, a) \\ z(t, a) \end{bmatrix} - \begin{bmatrix} x(t, b) \\ z(t, b) \end{bmatrix} \right|_1 \leq |a - b|_1, \quad \text{for all } t \geq 0. \quad (4.13)$$

In other words, the ℓ_1 distance between trajectories can never increase.

Pick $a \in \Omega$, and let $s := 1_d a$. Substituting $b = e_{L_s}$ in (4.13) yields

$$\left| \begin{bmatrix} x(t, a) \\ z(t, a) \end{bmatrix} - e_{L_s} \right|_1 \leq |a - e_{L_s}|_1, \quad \text{for all } t \geq 0. \quad (4.14)$$

This means that the convergence to the equilibrium point e_{L_s} is monotone in the sense that the ℓ_1 distance to e_{L_s} can never increase.

4.6. Entrainment

Many important biological processes are periodic. Examples include circadian clocks and the cell-cycle division process. Proper functioning requires certain biological systems to follow these periodic patterns, i.e. to *entrain* to the periodic excitation.

In the context of translation, it has been shown that both the RFM [26] and the RFMR [30] entrain to periodic translation rates, i.e. if all the transition rates are periodic time-varying functions, with a common (minimal) period $T > 0$ then each state variable converges to a periodic trajectory, with a period T . Here we show that the same property holds for the RFMNP.

We say that a function f is T -periodic if $f(t + T) = f(t)$ for all t . Assume that the λ_i^j 's in the RFMNP are time-varying functions satisfying

- there exist $0 < \delta_1 < \delta_2$ such that $\lambda_i^j(t) \in [\delta_1, \delta_2]$ for all $t \geq 0$ and all $j \in \{1, \dots, m\}$, $i \in \{1, \dots, n_j\}$.
- there exists a (minimal) $T > 0$ such that all the λ_i^j 's are T -periodic.

We refer to the model in this case as the *periodic ribosome flow model network with a pool* (PRFMNP).

Theorem 4.11. Consider the PRFMNP. Fix an arbitrary $s > 0$. There exists a unique function $\phi_s: \mathbb{R}_+ \rightarrow \text{Int}(\Omega)$, that is T -periodic, and for any $a \in L_s$ the solution of the PRFMNP converges to ϕ_s .

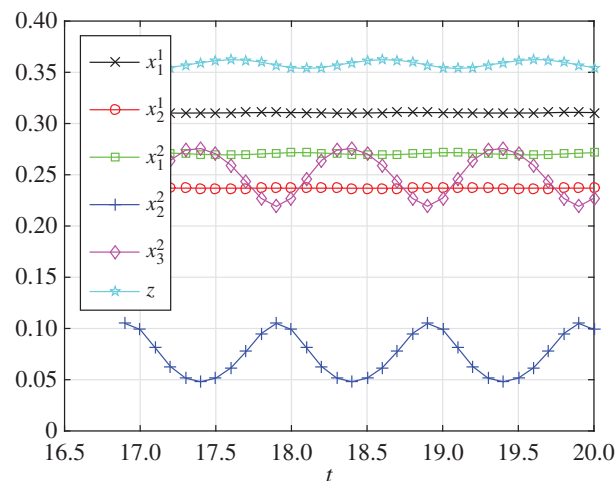


Figure 14. Trajectories of the PRFMNP in example 4.12 as a function of time. (Online version in colour.)

In other words, every level set L_s of H contains a unique periodic solution, and every solution of the PRFMNP emanating from L_s converges to this solution. Thus, the PRFMNP entrains (or phase locks) to the periodic excitation in λ_i^j 's. This implies in particular that all the protein production rates converge to a periodic pattern with period T .

Note that since a constant function is a periodic function for any T , theorem 4.11 implies entrainment to a periodic trajectory in the particular case where one of λ_i^j 's oscillates, and all the other are constant. Note also that the stability result in theorem 4.1 follows from theorem 4.11.

Example 4.12. Consider the RFMNP (3.7) with $G_i(z) = \tanh(z)$, and all rates equal to one except for $\lambda_2^2(t) = 5 + 4 \sin(2\pi t)$. In other words, there is a single time-varying periodic rate in RFMIO #2. Note that all these rates are periodic with a common minimal period $T = 1$. Figure 14 depicts the solution of this PRFMNP as a function of time t for $16.9 \leq t \leq 20$. The initial condition is $z(0) = x_i^j(0) = 1/4$ for all i, j . It may be seen that all the state variables converge to a periodic solution. In particular, all state variables $x_i^j(t)$ converge to a periodic solution with (minimal) period $T = 1$, and so does the pool occupancy $z(t)$. $x_1^j(t)$'s also converge to a periodic solution, but it is not possible to tell from the figure whether there are small oscillations with period $T = 1$ or the convergence is to a constant (of course, in both cases this is a periodic solution with period $T = 1$). However, it can be shown using the first two equations in (3.7) that if $z(t)$ converges to a periodic solution then so do $x_1^1(t)$ and $x_2^1(t)$. Note that the peaks in $x_3^2(t)$ are correlated with dips in $x_2^2(t)$, this is because when $\lambda_2^2(t)$ is high on some time interval, i.e. the transition rate from site 2 to site 3 is high, there is a high flow of ribosomes from site 2 to site 3 during this interval. ■

From the biophysical point of view, this means that the competition for free ribosomes induces a coupling between the mRNA molecules. This can induce periodic oscillations in *all* the protein production rates even when all the transition rates in the molecules are constant, except for a single rate in a single molecule that oscillates periodically. The translation rate of codons is affected among others by the tRNA supply (i.e. the intracellular abundance of the different tRNA species) and demand (i.e. total number of codons

from each type on all the mRNA molecules) [63]. Thus, the translation rate of a codon(s) is affected by changes in the demand (e.g. oscillations in mRNA levels) or by changes in the supply (e.g. oscillations in tRNA levels). The results reported here may suggest that oscillations in the mRNA levels of some genes or in the concentration of some tRNA species (that occur, for example, during the cell cycle [64,65]), can induce oscillations in the translation rates of the rest of the genes.

5. Discussion

We introduced a new model, the RFMNP, for large-scale simultaneous translation and competition for ribosomes that combines several RFMIOs interconnected via a dynamic pool of free ribosomes. To the best of our knowledge, this is the first model of a network composed of interconnected RFMIOs. The RFMNP is amenable to analysis because it is a monotone dynamical system that admits a non-trivial first integral. The fact that the total number of ribosomes in the network is conserved means that local properties of any mRNA molecule (e.g. the abundance of the corresponding tRNA molecules) affects its own translation rate, and via competition, also *globally* affects the translation rates of all the other mRNAs in the network. We provide several examples demonstrating the behaviour of the RFMNP, and rigorously analyse some of its mathematical properties. In particular, the RFMNP is an irreducible cooperative dynamical system admitting a continuum of linearly ordered equilibrium points, and every trajectory converges to an equilibrium point. The RFMNP is also on the ‘verge of contraction’ with respect to the ℓ_1 norm, and it entrains to periodic transition rates with a common period.

An important implication of our analysis and simulation results is that there are regimes and parameter values where there is a strong coupling between the different ‘translation components’ (ribosomes and mRNAs) in the cell. Such regimes cannot be studied using models for translation of a single isolated mRNA molecule. The RFMNP is specifically important when studying highly expressed genes with many mRNA molecules and ribosomes translating them because the dynamics of such genes strongly affects the ribosomal pool. For example, changes in the translation dynamics of a heterologous gene which is expressed with a very strong promoter, resulting in very high mRNA copy number should affect the entire tRNA pool, and thus the translation of other endogenous genes. Highly expressed endogenous genes ‘consume’ many ribosomes. Thus, a mutation that affects their (local) translation rate is expected to affect also the translation dynamics of other mRNA molecules. Studying the evolution of such genes should be based on understanding the global effect of such mutations using a computational model such as the RFMNP.

On the other hand, we can approximate the dynamics of genes that are not highly expressed (e.g. a gene with mRNA levels that are 0.01% of the mRNA levels in the cell) using a single RFM. In this case, the relative effect of the mRNA on all other mRNAs is expected to be limited.

Our analysis shows that increasing the translation initiation rate of a heterologous gene will always have a negative effect on the translation rate of the other genes (i.e. their translation rates decrease) and vice versa. The effect of

increasing [decreasing] the translation rate of a codon of the heterologous gene on the translation rate of other genes is more complicated: while it always increases [decreases] the translation rate of the heterologous gene it may either increase or decrease the translation rate of all other genes. The specific outcome of such a manipulation can be predicted using the RFMNP with parameter values that are based on the biophysical properties of the heterologous genes and the host genome.

Some of our analytical results may be tested experimentally. This can be done by designing and expressing a library of heterologous reporter genes [2,40,56]. The initiation rates can be manipulated based on the engineering of the nucleotides surrounding the START codon (e.g. [40–42]). The elongation rates of codons (and thus the transition rates) can be estimated using ribosome profiling [66] in *in vivo* experiments [34,40,67]. We can simulate a model such as the RFM (or TASEP) based on the inferred initiation and elongation rates (e.g. [23]), and use the models described here to design and demonstrate examples of heterologous genes where increasing [decreasing] the rate of a codon increases [decreases] the initiation (and translation) rate of all other mRNAs.

The effect of the manipulation of a codon (i.e. increasing or decreasing its rate) of the heterologous reporter gene on the ribosomal densities and translation rates of all the mRNAs (endogenous and heterologous) can be performed via ribosome profiling [66] in addition to measurements of mRNA levels, translation rates and protein levels [68]. One can also measure the fluorescence level of the protein related to an *additional* reporter gene following the expression of our library, as was recently done in [69]. Variants in the library that decrease/increase the free ribosomal pool are expected to decrease/increase the fluorescence level of the protein related to the additional reporter gene. Furthermore, we can measure the growth rate of the host and examine whether indeed a global increase [decrease] in the initiation rates yields a positive [negative] effect on the host growth rate.

Our analysis suggests that the effect of improving the translation rate of a codon in an mRNA molecule on the production rate of other genes and the pool of ribosomes depends on the initiation rate in the modified mRNA. When the initiation rate is very low the effect is expected to be positive (all other production rates increase). However, if the initiation rate is high the effect may be negative. This may partially explain the selection for slower codons in highly expressed genes that practically decrease the initiation rate [1,4]. This relation may also suggest a new factor that contributes to the evolution of highly expressed genes towards higher elongation and termination rates (i.e. the tendency of highly expressed genes to include ‘fast’ codons). Indeed, lower elongation rates (and thus a relatively high initiation rate) may decrease the production rates of other mRNAs that are needed for proper functioning of the organism.

The RFM, and thus also the model described here, does not capture certain aspects of mRNA translation. For example, eukaryotic ribosomes may translate mRNAs in multiple cycles before entering the free ribosomal pool [18,44,70]. This phenomenon may perhaps be modelled by adding positive feedback [44] in the RFMNP. In addition, different genes are transcribed at different rates, resulting in a different number of (identical) mRNA copies for different genes. This can be modelled using a set of identical RFMs for each gene. Such a model can help in understanding how changes in mRNA levels of one gene

affect the translation rates of all the mRNAs. The analysis here suggests that modifying the mRNA levels of a gene will affect the translation rates of all other genes in the same way. These and other aspects of biological translation may be integrated in our model in future studies. Hamadeh & Del Vecchio [71] develop the notion of the realizable region for steady-state gene expression under resource limitations, and methods for mitigating the effects of ribosome competition. Another interesting research direction is studying these topics in the context of the RFMNP.

We believe that networks of interconnected RFMIOs may also prove to be powerful modelling and analysis tools for

other natural and artificial systems. These include communication networks, intracellular trafficking in the cell, coordination of large groups of organisms (e.g. ants), traffic control and more.

Competing interests. We declare we have no competing interests.

Funding. The research of M.M. and T.T. is partially supported by a research grant from the Israeli Ministry of Science, Technology, and Space. The work of E.D.S. is supported in part by grants AFOSR FA9550-14-1-0060 and ONR 5710003367.

Acknowledgement. We thank Yoram Zarai for helpful comments. We are grateful to the anonymous reviewers for comments that helped us to improve this paper.

References

- Tuller T *et al.* 2010 An evolutionarily conserved mechanism for controlling the efficiency of protein translation. *Cell* **141**, 344–354. (doi:10.1016/j.cell.2010.03.031)
- Kudla G, Murray AW, Tollervey D, Plotkin JB. 2009 Coding-sequence determinants of gene expression in *Escherichia coli*. *Science* **324**, 255–258. (doi:10.1126/science.1170160)
- Sharp PM, Emery LR, Zeng K. 2010 Forces that influence the evolution of codon bias. *Phil. Trans. R. Soc. B* **365**, 1203–1212. (doi:10.1098/rstb.2009.0305)
- Tuller T, Zur H. 2015 Multiple roles of the coding sequence 5' end in gene expression regulation. *Nucleic Acids Res.* **43**, 13–28. (doi:10.1093/nar/gku1313)
- Richter JD, Smith LD. 1981 Differential capacity for translation and lack of competition between mRNAs that segregate to free and membrane-bound polysomes. *Cell* **27**, 183–191. (doi:10.1016/0092-8674(81)90372-X)
- Vind J, Sorensen MA, Rasmussen MD, Pedersen S. 1993 Synthesis of proteins in *Escherichia coli* is limited by the concentration of free ribosomes: expression from reporter genes does not always reflect functional mRNA levels. *J. Mol. Biol.* **231**, 678–688. (doi:10.1006/jmbi.1993.1319)
- Jens M, Rajewsky N. 2015 Competition between target sites of regulators shapes post-transcriptional gene regulation. *Nat. Rev. Genet.* **16**, 113–126. (doi:10.1038/nrg3853)
- Zenkhusen D, Larson DR, Singer RH. 2008 Single-RNA counting reveals alternative modes of gene expression in yeast. *Nat. Struct. Mol. Biol.* **15**, 1263–1271. (doi:10.1038/nsmb.1514)
- Warner J. 1999 The economics of ribosome biosynthesis in yeast. *Trends Biochem. Sci.* **24**, 437–440. (doi:10.1016/S0968-0004(99)01460-7)
- Gyorgy A, Del Vecchio D. 2014 Limitations and trade-offs in gene expression due to competition for shared cellular resources. In *Proc. 53rd IEEE Conf. on Decision and Control*, pp. 5431–5436.
- Gyorgy A, Jimenez JI, Yazbek J, Huang H, Chung H, Weiss R, Del Vecchio D. 2015 Isocost lines describe the cellular economy of genetic circuits. *Biophys. J.* **109**, 639–646. (doi:10.1016/j.bpj.2015.06.034)
- Mather W, Hasty J, Tsimring LS, Williams RJ. 2013 Translational cross talk in gene networks. *Biophys. J.* **104**, 2564–2572. (doi:10.1016/j.bpj.2013.04.049)
- Brewster RC, Weinert FM, Garcia HG, Song D, Rydenfelt M, Phillips R. 2014 The transcription factor titration effect dictates level of gene expression. *Cell* **156**, 1312–1323. (doi:10.1016/j.cell.2014.02.022)
- Dana A, Tuller T. 2012 Efficient manipulations of synonymous mutations for controlling translation rate—an analytical approach. *J. Comput. Biol.* **19**, 200–231. (doi:10.1089/cmb.2011.0275)
- Brackley CA, Romano MC, Thiel M. 2011 The dynamics of supply and demand in mRNA translation. *PLoS Comput. Biol.* **7**, e1002203. (doi:10.1371/journal.pcbi.1002203)
- MacDonald CT, Gibbs JH, Pipkin AC. 1968 Kinetics of biopolymerization on nucleic acid templates. *Biopolymers* **6**, 1–25. (doi:10.1002/bip.1968.360060102)
- Shaw LB, Zia RKP, Lee KH. 2003 Totally asymmetric exclusion process with extended objects: a model for protein synthesis. *Phys. Rev. E* **68**, 021910. (doi:10.1103/PhysRevE.68.021910)
- von der Haar T. 2012 Mathematical and computational modelling of ribosomal movement and protein synthesis: an overview. *Comput. Struct. Biotechnol. J.* **1**, 1–7. (doi:10.5936/csbj.201204002)
- Ha M, den Nijs M. 2002 Macroscopic car condensation in a parking garage. *Phys. Rev. E* **66**, 036118. (doi:10.1103/PhysRevE.66.036118)
- Cook LJ, Zia RKP. 2009 Feedback and fluctuations in a totally asymmetric simple exclusion process with finite resources. *J. Stat. Mech.* **2009**, P02012. (doi:10.1088/1742-5468/2009/02/P02012)
- Cook LJ, Zia RKP. 2012 Competition for finite resources. *J. Stat. Mech.* **2012**, P05008. (<http://arxiv.org/abs/1203.1602>)
- Greulich P, Ciandrini L, Allen RJ, Romano MC. 2012 Mixed population of competing totally asymmetric simple exclusion processes with a shared reservoir of particles. *Phys. Rev. E* **85**, 011142. (doi:10.1103/PhysRevE.85.011142)
- Reuveni S, Meilijson I, Kupiec M, Ruppim E, Tuller T. 2011 Genome-scale analysis of translation elongation with a ribosome flow model. *PLoS Comput. Biol.* **7**, e1002127. (doi:10.1371/journal.pcbi.1002127)
- Blythe RA, Evans MR. 2007 Nonequilibrium steady states of matrix-product form: a solver's guide. *J. Phys. A* **40**, R333–R441. (doi:10.1088/1751-8113/40/46/R01)
- Margaliot M, Tuller T. 2012 Stability analysis of the ribosome flow model. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **9**, 1545–1552. (doi:10.1109/TCBB.2012.88)
- Margaliot M, Sontag ED, Tuller T. 2014 Entrainment to periodic initiation and transition rates in a computational model for gene translation. *PLoS ONE* **9**, e96039. (doi:10.1371/journal.pone.0096039)
- Poker G, Zarai Y, Margaliot M, Tuller T. 2014 Maximizing protein translation rate in the non-homogeneous ribosome flow model: a convex optimization approach. *J. R. Soc. Interface* **11**, 20140713. (doi:10.1098/rsif.2014.0713)
- Zarai Y, Margaliot M, Tuller T. 2014 Maximizing protein translation rate in the ribosome flow model: the homogeneous case. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **11**, 1184–1195. (doi:10.1109/TCBB.2014.2330621)
- Poker G, Margaliot M, Tuller T. 2015 Sensitivity of mRNA translation. *Sci. Rep.* **5**, 12795. (doi:10.1038/srep12795)
- Raveh A, Zarai Y, Margaliot M, Tuller T. 2015 Ribosome flow model on a ring. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **12**, 1429–1439. (doi:10.1109/TCBB.2015.2418782)
- Del Vecchio D, Murray RM. 2014 *Biomolecular feedback systems*. Princeton, NJ: Princeton University Press.
- Qian Y, Del Vecchio D. 2015 Effective interaction graphs arising from resource limitations in gene networks. In *Proc. Am. Control Conf. (ACC2015)*, Chicago, IL, 1–3 July. AACC.
- Ala U *et al.* 2013 Integrated transcriptional and competitive endogenous RNA networks are cross-regulated in permissive molecular environments. *Proc. Natl Acad. Sci. USA* **110**, 7154–7159. (doi:10.1073/pnas.1222509110)
- Dana A, Tuller T. 2014 The effect of tRNA levels on decoding times of mRNA codons. *Nucleic Acids Res.* **42**, 9171–9181. (doi:10.1093/nar/gku646)
- Tuller T, Veksler I, Gazit N, Kupiec M, Ruppim E, Ziv M. 2011 Composite effects of the coding sequences

- determinants on the speed and density of ribosomes. *Genome Biol.* **12**, R110. (doi:10.1186/gb-2011-12-11-r110)
36. Li G-W, Oh E, Weissman J. 2012 The anti-Shine-Dalgarno sequence drives translational pausing and codon choice in bacteria. *Nature* **484**, 538–541. (doi:10.1038/nature10965)
 37. Lu J, Deutsch C. 2008 Electrostatics in the ribosomal tunnel modulate chain elongation rates. *J. Mol. Biol.* **384**, 73–86. (doi:10.1016/j.jmb.2008.08.089)
 38. Sabi R, Tuller T. 2015 A comparative genomics study on the effect of individual amino acids on ribosome stalling. *BMC Genomics* **16**, S5. (doi:10.1186/1471-2164-16-S10-S5)
 39. Shine J, Dalgarno L. 1975 Determinant of cistron specificity in bacterial ribosomes. *Nature* **254**, 34–38. (doi:10.1038/254034a0)
 40. Ben-Yehzekel T *et al.* 2015 Rationally designed, heterologous *S. cerevisiae* transcripts expose novel expression determinants. *RNA Biol.* **12**, 972–984. (doi:10.1080/15476286.2015.1071762)
 41. Zur H, Tuller T. 2013 New universal rules of eukaryotic translation initiation fidelity. *PLoS Comput. Biol.* **9**, e1003136. (doi:10.1371/journal.pcbi.1003136)
 42. Kozak M. 1986 Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell* **44**, 283–292. (doi:10.1016/0092-8674(86)90762-2)
 43. Tripathy G, Barma M. 1998 Driven lattice gases with quenched disorder: exact results and different macroscopic regimes. *Phys. Rev. E* **58**, 1911–1926. (doi:10.1103/PhysRevE.58.1911)
 44. Margaliot M, Tuller T. 2013 Ribosome flow model with positive feedback. *J. R. Soc. Interface* **10**, 20130267. (doi:10.1098/rsif.2013.0267)
 45. Angeli D, Sontag ED. 2003 Monotone control systems. *IEEE Trans. Automat. Control* **48**, 1684–1698. (doi:10.1109/TAC.2003.817920)
 46. Adams D, Schmittmann B, Zia R. 2008 Far-from-equilibrium transport with constrained resources. *J. Stat. Mech.* **2008**, P06009.
 47. Galla T. 2011 Optimizing evacuation flow in a two-channel exclusion process. *J. Stat. Mech.* **2011**, P09004.
 48. Margaliot M, Tuller T. 2012 On the steady-state distribution in the homogeneous ribosome flow model. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **9**, 1724–1736. (doi:10.1109/TCBB.2012.120)
 49. Bremer H, Dennis PP. 1996 Modulation of chemical composition and other parameters of the cell by growth rate. In *Escherichia coli and Salmonella typhimurium: cellular and molecular biology*, 2nd edn. ch. 97, 1559.
 50. Tuller T, Waldman YY, Kupiec M, Ruppin E. 2010 Translation efficiency is determined by both codon bias and folding energy. *Proc. Natl Acad. Sci. USA* **107**, 3645–3650. (doi:10.1073/pnas.0909910107)
 51. Castillo-Davis C, Mekhedov S, Hartl D, Koonin E, Kondrashov F. 2002 Selection for short introns in highly expressed genes. *Nat. Genet.* **31**, 415–418. (doi:10.1038/ng940)
 52. Eisenberg E, Levanon E. 2003 Human housekeeping genes are compact. *Trends Genet.* **19**, 362–365. (doi:10.1016/S0168-9525(03)00140-9)
 53. Li S, Feng L, Niu D. 2007 Selection for the miniaturization of highly expressed genes. *Biochem. Biophys. Res. Commun.* **360**, 586–592. (doi:10.1016/j.bbrc.2007.06.085)
 54. Tuller T, Ruppin E, Kupiec M. 2009 Properties of untranslated regions of the *S. cerevisiae* genome. *BMC Genomics* **10**, 329. (doi:10.1186/1471-2164-10-391)
 55. Carmel L, Koonin E. 2009 A universal nonmonotonic relationship between gene compactness and expression levels in multicellular eukaryotes. *Genome Biol. Evol.* **1**, 382–390. (doi:10.1093/gbe/evp038)
 56. Welch M, Govindarajan S, Ness JE, Villalobos A, Gurney A, Minshull J, Gustafsson C. 2009 Design parameters to control synthetic gene expression in *Escherichia coli*. *PLoS ONE* **4**, e7002. (doi:10.1371/journal.pone.0007002)
 57. Jacques N, Dreyfus M. 1990 Translation initiation in *Escherichia coli*: old and new questions. *Mol. Microbiol.* **4**, 1063–1067. (doi:10.1111/j.1365-2958.1990.tb00679.x)
 58. Smith HL. 1995 *Monotone dynamical systems: an introduction to the theory of competitive and cooperative systems*. Mathematical Surveys and Monographs 41. Providence, RI: American Mathematical Society.
 59. Lohmiller W, Slotine J-JE. 1998 On contraction analysis for non-linear systems. *Automatica* **34**, 683–696. (doi:10.1016/S0005-1098(98)00019-3)
 60. Aminzare Z, Sontag ED. 2014 Contraction methods for nonlinear systems: a brief introduction and some open problems. In *Proc. 53rd IEEE Conf. on Decision and Control*, pp. 3835–3847. Los Angeles, CA.
 61. Russo G, di Bernardo M, Sontag ED. 2010 Global entrainment of transcriptional systems to periodic inputs. *PLoS Comput. Biol.* **6**, e1000739. (doi:10.1371/journal.pcbi.1000739)
 62. Margaliot M, Sontag ED, Tuller T. 2016 Contraction after small transients. *Automatica* **67**, 178–184. (<http://arxiv.org/abs/1506.06613>)
 63. Pechmann S, Frydman J. 2013 Evolutionary conservation of codon optimality reveals hidden signatures of cotranslational folding. *Nat. Struct. Mol. Biol.* **20**, 237–243. (doi:10.1038/nsmb.2466)
 64. Spellman P, Sherlock G, Zhang M, Iyer V, Anders K, Eisen M, Brown P, Botstein D, Futcher B. 1998 Comprehensive identification of cell cycle-regulated genes of the yeast *Saccharomyces cerevisiae* by microarray hybridization. *Mol. Biol. Cell.* **9**, 3273–3297. (doi:10.1091/mbc.9.12.3273)
 65. Frenkel-Morgenstern M, Danon T, Christian T, Igarashi T, Cohen L, Hou YM, Jensen LJ. 2012 Genes adopt non-optimal codon usage to generate cell cycle-dependent oscillations in protein levels. *Mol. Syst. Biol.* **8**, 572. (doi:10.1038/msb.2012.3)
 66. Ingolia NT, Ghaemmaghami S, Newman JR, Weissman JS. 2009 Genome-wide analysis *in vivo* of translation with nucleotide resolution using ribosome profiling. *Science* **324**, 218–223. (doi:10.1126/science.1168978)
 67. Dana A, Tuller T. 2014 Mean of the typical decoding rates: a new translation efficiency index based on the analysis of ribosome profiling data. *G3 (Bethesda)* **5**, 73–80. (doi:10.1534/g3.114.015099)
 68. Schwanhauser B, Busse D, Li N, Dittmar G, Schuchhardt J, Wolf J, Chen W, Selbach M. 2011 Global quantification of mammalian gene expression control. *Nature* **473**, 337–342. (doi:10.1038/nature10098)
 69. Ceroni F, Algar R, Stan G-B, Ellis T. 2015 Quantifying cellular capacity identifies gene expression designs with reduced burden. *Nat. Methods* **12**, 415–418. (doi:10.1038/nmeth.3339)
 70. Kopeina G, Afonina ZA, Gromova K, Shirokov V, Vasiliev V, Spirin A. 2008 Step-wise formation of eukaryotic double-row polyribosomes and circular translation of polysomal mRNA. *Nucleic Acids Res.* **36**, 2476–2488. (doi:10.1093/nar/gkm1177)
 71. Hamadeh A, Del Vecchio D. 2014 Mitigation of resource competition in synthetic genetic circuits through feedback regulation. In *Proc. 53rd IEEE Conf. on Decision and Control*, pp. 3829–3834.