



Available online at www.sciencedirect.com

SCIENCE @ DIRECT®

C. R. Biologies 328 (2005) 429–444



<http://france.elsevier.com/direct/CRASS3/>

Biological modelling / Biomodélisation

A simple framework to describe the regulation of gene expression in prokaryotes

Filipa Alves, Rui Dilão *

Non-Linear Dynamics Group, Instituto Superior Técnico, Department of Physics, Av. Rovisco Pais, 1049-001 Lisboa, Portugal

Received 23 November 2004; accepted after revision 25 January 2005

Available online 19 February 2005

Presented by Pierre Auger

Abstract

Based on the bimolecular mass action law and the derived mass conservation laws, we propose a mathematical framework in order to describe the regulation of gene expression in prokaryotes. It is shown that the derived models have all the qualitative properties of the activation and inhibition regulatory mechanisms observed in experiments. The basic construction considers genes as templates for protein production, where regulation processes result from activators or repressors connecting to DNA binding sites. All the parameters in the models have a straightforward biological meaning. After describing the general properties of the basic mechanisms of positive and negative gene regulation, we apply this framework to the self-regulation of the *trp* operon and to the genetic switch involved in the regulation of the *lac* operon. One of the consequences of this approach is the existence of conserved quantities depending on the initial conditions that tune bifurcations of fixed points. This leads naturally to a simple explanation of threshold effects as observed in some experiments. **To cite this article:** F. Alves, R. Dilão, C. R. *Biologies* 328 (2005).

© 2005 Académie des sciences. Published by Elsevier SAS. All rights reserved.

Résumé

Un cadre simple décrivant l'expression génique chez les procaryotes. En partant de la loi d'action de masse bimoléculaire et des lois dérivées de conservation de masse, nous proposons un cadre mathématique permettant de décrire la régulation de l'expression génique chez les procaryotes. On montre que les modèles dérivés possèdent tous les propriétés qualitatives des mécanismes régulateurs d'activation et d'inhibition observés expérimentalement. La construction basique considère les gènes comme des modèles types pour la production de protéines, pour lesquels les processus de régulation résultent d'activateurs or de répresseurs connectant les sites de liaison de l'ADN. Tous les paramètres de ces modèles ont une signification biologique manifeste. Après avoir décrit les propriétés générales des mécanismes basiques de la régulation positive et négative des gènes, nous appliquons ce cadre à l'autorégulation de l'opéron *trp* et à la mutation génétique impliquée dans la régulation de l'opéron *lac*. Une des conséquences de cette approche est l'existence de quantités conservées dépendant des conditions initiales qui

* Corresponding author.

E-mail address: rui@sd.ist.utl.pt (R. Dilão).

commandent les bifurcations de points fixes. Ceci conduit naturellement à une explication des effets de seuil observés dans quelques expériences. *Pour citer cet article* : F. Alves, R. Dilão, C. R. Biologies 328 (2005).

© 2005 Académie des sciences. Published by Elsevier SAS. All rights reserved.

Keywords: Genetic regulatory networks; Mathematical modelling of genetic regulatory networks; *trp* operon; *lac* operon

Mots-clés : Réseaux de régulation ; Modélisation mathématique de réseaux de régulation génétique ; Opéron *trp* ; Opéron *lac*

1. Introduction

Genome sequencing is producing huge amounts of data and modern experimental techniques are enabling large scale gene expression analysis (see, for example, [1–5]).

Mathematical modelling of gene expression can provide interpretation of experimental results, help to focus new experiments on critical missing data, and predictions can be made about results that are difficult to obtain *in vivo*. Mathematical modelling can still be used to explore and predict properties of a given system, testing the behaviour of different network structures and parameter values. On the other hand, the inferential power of modelling can be used to deal with ambiguous or unknown information in experiments.

Several theoretical strategies are being used to approach the modelling of gene regulation networks [6–9]. For example, several authors have developed simple representations of real genetic networks and have explored its basic properties [10,11]. Some models are built starting from simplified views of known regulatory networks and then the model parameters are adjusted in order to find results close to the patterns of gene expression observed *in vivo* [12,13]. Another approach is to infer unknown regulatory mechanisms from known experimental data [14–18]. A different strategy is the design of simple *in vivo* artificial genetic networks that have specific properties to be tested under several conditions [19–25].

From the mathematical point of view, there are several methodologies to describe genetic regulatory networks. Some authors use discrete models, building, for example, boolean or bayesian networks [26–32]. Others, use several types of continuous models, as differential equations, difference equations and reaction-diffusion equations [33–38]. All these models aim to describe the real system with different levels of detail.

Here we propose a simple mathematical framework in order to build models of gene regulatory interactions

using differential equations. Our approach is based on the operon model of Jacob and Monod [39], as a paradigm of the mechanisms of gene regulation in bacteria. The description of these mechanisms starts by establishing a basic model for the quantitative description of protein production. The action of the activators and repressors regulating the expression of each gene are the building blocks of these mathematical models.

We consider that genes are templates for protein production, and regulation occurs due to the inter-conversion between the activated and the repressed states of the genes. We also assume that both transcriptional and translational mechanisms are described by one overall rate constant. As a consequence, we are able to describe genetic regulatory systems where the control of the protein production is associated with ‘on’ and ‘off’ states of the gene expression. The resulting gene expression patterns predicted by these mathematical models can be compared with the real experimental system in order to be validated and calibrated. A similar strategy can also be used to infer some unknown gene interactions from known expression patterns.

One of our concerns has been to maintain this mathematical framework as simple as possible, enabling the derivation of the qualitative properties of the models, as well as the possibility of constructing simple computer programs for the derivation and integration of the equations for the temporal evolution of protein concentrations. From the methodological point of view, this construction is based on the bi-molecular mass action law and the basic mechanisms do not introduce ad hoc regulatory functions, as is in general used in the mathematical modelling of gene expression [7,13,40–42]. As a consequence, the modelling strategy proposed here is close to the basic principles of chemical kinetics, and do not introduces neither ad hoc threshold effects, nor Michaelis-Menten type functional forms. However, threshold effects will be

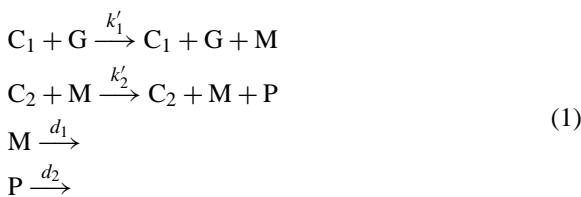
found as a consequence of the conservation laws associated to the mass action law.

In Section 2, we introduce the basic simplifying assumptions for the construction of the differential equations describing genetic regulatory networks, and we analyze the basic mechanism of protein production resulting from gene expression, without any regulatory mechanism. From Sections 3 to 6, we analyze the simplest cases of positive and negative regulation. It is shown that the proposed models have all the qualitative properties of the regulatory mechanisms analyzed in experiments. Therefore, if a biological mechanism in prokaryotes is described by a known genetic network, in principle, the formalism developed here is a good candidate to describe the most important qualitative features of the biological process under analysis.

In Section 7, as an illustrative example, we apply this framework to model of the negative self-regulation of the *trp* operon. In Section 8, we analyze the genetic switch involved in the regulation of the *lac* operon, responsible for the control of the lactose metabolism in *Escherichia coli*. Finally, in Section 9, we discuss the main conclusions of this paper.

2. The basic model for protein production from a DNA template

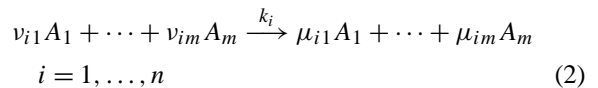
The simplest model of protein production from a DNA template without any regulation mechanism can be described by the kinetic diagram:



where G, M and P represent the gene, the mRNA and the protein, respectively, C_1 and C_2 are catalysts not consumed in the reactions; k'_1 , k'_2 , d_1 and d_2 are rate constants.

In chemical kinetics, diagrams of type (1) represent the collision between molecules and subsequent chemical transformation. In a chemical process of type (1), after the production of one molecule of protein P, the gene template, the mRNA and the catalysts become free, and new transcription and translation processes

can be initiated. If the collisional interactions that initiates these processes occur randomly in the media, the time evolution of the several components is described by the mass action law, [43,44]. To be more specific, we consider several substances, A_j , $j = 1, \dots, m$, involved in the n kinetic diagrams:



where v_{ij} and μ_{ij} are integer stoichiometric coefficients, and the constants k_i measure the rates of the transformations. From a master equation approach [43, pp. 166–172], it follows that the concentrations of the A_j evolve in time according to the equations:

$$\begin{aligned}
 \frac{dA_j}{dt} &= \sum_{i=1}^n k_i (\mu_{ij} - v_{ij}) A_1^{v_{i1}} \dots A_m^{v_{im}} \\
 j &= 1, \dots, m
 \end{aligned}
 \tag{3}$$

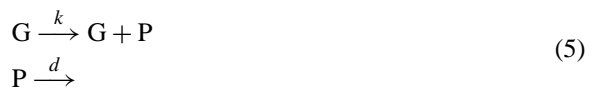
In the following, Eq. (3) calculated from diagrams (2), will be simply called mass action law.

Applying the mass action law to the kinetic diagram (1), the time evolution of the concentrations of G, M and P is described by the differential equations:

$$\begin{aligned}
 \frac{dG}{dt} &= 0 \\
 \frac{dM}{dt} &= k'_1 G C_1 - d_1 M \\
 \frac{dP}{dt} &= k'_2 M C_2 - d_2 P
 \end{aligned}
 \tag{4}$$

where we use the same notation to represent a specific substance or its concentration.

To simplify the model diagram (1) and Eq. (4), we assume that the concentrations of C_1 and C_2 are not limiting factors of the reactions in (1), that is, C_1 and C_2 do not degrade. We also assume that during transcription the mRNA (M) concentration is constant, implying that, by (4), $M = k'_1 G C_1 / d_1$. Under these conditions, we substitute the mechanism (1) by the simpler diagram:



In this case, the equations for the time evolution of the concentrations is:

$$\begin{aligned} \frac{dP}{dt} &= kG - dP \\ \frac{dG}{dt} &= 0 \end{aligned} \tag{6}$$

Comparing (4) with (6), we obtain $k = k'_1 k'_2 C_1 C_2 / d_1$. As we are concerned with the regulation of protein production, in the following we will use the simplified model form (5) to represent both transcriptional and translational mechanisms. Under these conditions, the concentrations of catalysts will affect the rate constant k .

The general solution of (6) is:

$$\begin{aligned} P(t) &= \left(P(0) - \frac{k}{d} G(0) \right) e^{-dt} + \frac{k}{d} G(0) \\ G(t) &= G(0) \end{aligned} \tag{7}$$

As far as protein is being produced at a constant rate k , by (7) and in the limit $t \rightarrow \infty$, the concentration level of protein attains the equilibrium value $P = kG(0)/d$. If protein production is interrupted at some time $t = t^*$, the level of protein production decays as $P(t) = P(t^*)e^{-dt}$, where $P(t^*)$ is the amount of protein at time t^* .

In this simple model for protein production, the use of the mass action law relies on the hypothesis that DNA templates, mRNA molecules and proteins are well distributed inside the cell and that the collisional interaction mechanisms are essentially random. The rate constants k and d are effective velocity constants multiplied by catalysts, nucleotides and aminoacids concentrations. In order to simplify the models, we will always avoid the explicit introduction of other species than gene templates (DNA) and proteins, being the gene concentration always a conserved quantity.

In the following, in order to describe the binding of transcription factors that account for activation and repression of proteins, we associate binding sites or operons to genes where transcription factors (transcriptional activators and repressors) can bind in order to promote or repress protein production, Fig. 1. This approach follows the operon model of Jacob and Monod [39].

We will now consider the basic cases of protein production by positive and negative regulation. The interaction between the regulators and inducer or corepressor molecules will not be taken into account in these models. We will only consider the interaction

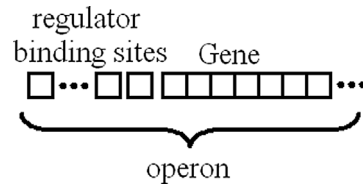


Fig. 1. Jacob and Monod operon model for the regulation of protein production in bacteria. The transcriptional regulators bind to the binding sites of the gene and transcription is activated or repressed. We can consider different binding sites for different transcription factors or, in the competitive case, the same binding site can be occupied by an activator or a repressor.

between activator or repressor molecules and the respective DNA binding sites. These interactions obey the mass action law (see, for example, [45]) and this approach will be followed in the next sections.

3. Positive regulation

3.1. The activator total concentration is constant

In the mechanism of protein production by positive regulation, an activator binds to an operator site in the DNA and transcription and consequent protein production (translation) is initiated. Representing the activator concentration in the cell by A , the concentration of template gene by G , and the concentration of binded DNA-activator complex by G_A , the basic mechanism for protein production is:



We are assuming that genes can be measured in units of concentration (number of molecules or gene templates per unit of volume). This applies to batch cultures where we have several copies of the same gene, and to tissues where we have several copies of identical cells. For cells with only one expressing gene, we can introduce an ergodic argument, and assume that for a time scale larger than the characteristic time of protein production, the protein production is also described by a mass action law mechanism. In fact, this ergodic argument is on the very foundation of the mass action law, in the derivation of Eqs. (3) from

molecular mechanisms (2) [46, pp. 89, 291–294]; [43, pp. 166–172].

Applying the mass action law to the above reactions, we obtain the following system of differential equations:

$$\begin{aligned} \frac{dA}{dt} &= -k_1 A G + k_2 G_A \\ \frac{dG}{dt} &= -k_1 A G + k_2 G_A \\ \frac{dG_A}{dt} &= k_1 A G - k_2 G_A \\ \frac{dP}{dt} &= a G_A - dP \end{aligned} \tag{9}$$

The evolution equations in (9) are not independent, obeying the conservation laws,

$$\begin{aligned} A + G_A &= A(0) + G_A(0) = c_1 \\ G + G_A &= G(0) + G_A(0) = c_2 \end{aligned} \tag{10}$$

where, $A(0)$, $G(0)$ and $G_A(0)$ denote initial concentrations, and c_1 and c_2 are positive constants. By (10), we can eliminate G and G_A from the system of Eqs. (9), obtaining the set of equations describing protein production:

$$\begin{aligned} \frac{dA}{dt} &= -k_1 A(c_2 - c_1 + A) + k_2(c_1 - A) \\ \frac{dP}{dt} &= a(c_1 - A) - dP \end{aligned} \tag{11}$$

Knowing the time solutions $A(t)$ and $P(t)$ of the system of Eqs. (11), by (10), we have, $G(t) = A(t) - c_1 + c_2$ and $G_A(t) = c_1 - A(t)$.

The system of Eqs. (11) has the two fixed points in phase space:

$$(A_p, P_p) \quad \text{and} \quad (A_n, P_n) \tag{12}$$

where

$$\begin{aligned} A_{p,n} &= \frac{1}{2} \left(c_1 - c_2 - \frac{k_2}{k_1} \right) \\ &\quad \pm \frac{1}{2} \sqrt{\left(c_1 - c_2 - \frac{k_2}{k_1} \right)^2 + 4c_1 \frac{k_2}{k_1}} \\ P_{p,n} &= \frac{a}{d} (A(0) + G_A(0) - A_{p,n}) \end{aligned} \tag{13}$$

and the indexes p and n stand for positive and negative signs, respectively. Note that $(0, 0)$ is a fixed point of Eq. (11) if and only if $c_1 = 0$. In this case, there is no

protein production because the activator A is absent. The only fixed point that corresponds to meaningful asymptotic solutions of Eq. (11) is (A_p, P_p) , which is a stable node provided $a > 0$, $d > 0$, $k_1 > 0$ and $k_2 > 0$. In this case, the Jacobian matrix of Eq. (11) calculated at (A_p, P_p) has eigenvalues $\lambda_1 = -k_1 s < 0$ and $\lambda_2 = -d < 0$, where s is square root term in (13). The same analysis shows that (A_n, P_n) is an (unstable) saddle point with eigenvalues $\lambda_1 = k_1 s > 0$ and $\lambda_2 = -d < 0$.

As the first equation in (11) is independent of P , integration by quadratures gives:

$$\begin{aligned} A(t) &= \frac{A_p(A_n - A_0)e^{(A_p k_1 t)} + A_n(A_0 t - A_p)e^{(A_n k_1 t)}}{(A_n - A_0)e^{(A_p k_1 t)} + (A_0 - A_p)e^{(A_n k_1 t)}} \\ P(t) &= P(0)e^{-dt} + e^{-dt} \int_0^t A(s) ds \end{aligned} \tag{14}$$

where $A_0 = A(0)$. The transient behaviour of protein depends on the reaction and degradation rates as well as on the conservation laws through (13) and (10). This is a typical behaviour in reaction kinetics where, in the presence of conservation laws, the parametric behaviour of transients, fixed points and phase space structure of orbits change with the initial concentration of the agents involved in the reactions. This effect has been pointed out by Travers [45], in the context of the control of the cell cycle by transcriptional switching.

In Fig. 2, we show the time solution (14) as a function of the time t , for a particular choice of parameters and initial conditions. For $t < 10$, the activator concentration is zero, and there is no protein production. At $t = 10$, the activator is introduced, and the concentration of protein increases with the usual sigmoidal growth form. The first conclusion we derive from this model is that the amount of produced protein (at equilibrium) is determined by the initial concentration of the activator, the initial concentration of the gene, the rates at which the activator binds and unbinds to the DNA template and the production and degradation rates. Comparing the quantitative behaviour of activated protein production with the non regulated production model derived in Section 2, depending of the activator concentration, in the asymptotic limit, positive regulation leads eventually to higher protein

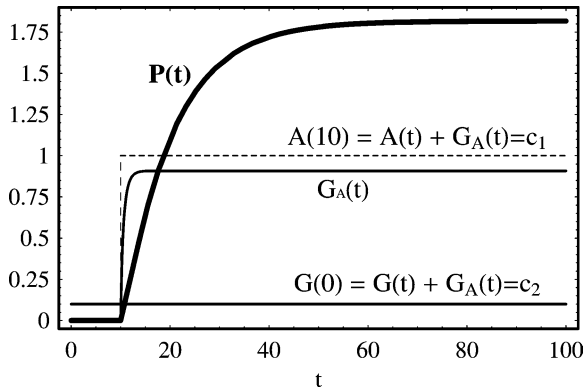
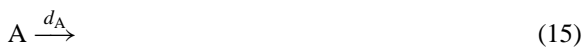


Fig. 2. Positive regulation of protein production with constant activator total concentration. We show the temporal evolution of protein concentration calculated from (11). From $t = 0$ to $t = 10$, the activator is absent, $A(t) = 0.0$. In this case, by (10), $c_1 = 0$ and there is no protein production. At time $t = 10$, we have changed the activator concentration for $A(10) = 1.0 = c_1$. Other parameter values are: $k_1 = 1.0$, $k_2 = 0.1$, $a = 0.2$ and $d = 0.1$. Initial concentrations are: $A(0) = 0.0$, $A(10) = 1.0 = c_1$, $P(0) = 0.0$, $G(0) = 0.1 = c_2$ and $G_A(0) = 0.0$. After the introduction of the activator at $t = 10$, the equilibrium solutions of the free activator and protein are $(A_p, P_p) = (0.092, 1.817)$.

concentrations than the nonactivated simple model for protein production.

3.2. The activator degrades

In the simple model (8), the activator does not degrade in the process. If the activator A is itself a protein, we must consider an extra mechanism accounting for the activator degradation:



Therefore, the first equation in (9) is now:

$$\frac{dA}{dt} = -k_1AG + k_2G_A - d_AA \quad (16)$$

In this case, we have only the second conservation law in (10), and the equations for protein production are:

$$\begin{aligned} \frac{dA}{dt} &= -k_1AG + k_2(c_2 - G) - d_AA \\ \frac{dG}{dt} &= -k_1AG + k_2(c_2 - G) \\ \frac{dP}{dt} &= a(c_2 - G) - dP \end{aligned} \quad (17)$$

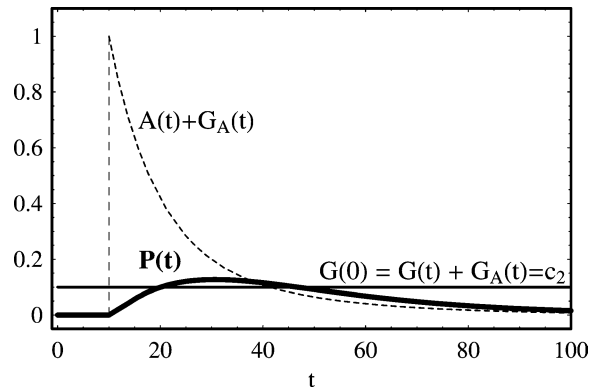


Fig. 3. Positive regulation of protein production with a degrading activator. We show the temporal evolution of protein concentration calculated from (17), for the same parameter values and initial concentrations of Fig. 1. Initially, the activator is absent and there is no protein production. At $t = 10$, we have introduced the activator with the initial value $A(10) = 1.0$, and the degradation rate of the activator protein has been set to $d_A = 0.1$. The equilibrium values of the activator, gene template and protein are $(A^*, G^*, P^*) = (0, 0.1, 0)$.

where we have eliminated G_A owing to the second conservation law in (10).

Eq. (17) has one stable fixed point with coordinates $(A^*, G^*, P^*) = (0, G(0) + G_A(0), 0) = (0, c_2, 0)$, and the asymptotic time solution converge to this fixed point. In Fig. 3, we show the time solution (17) as a function of t for the same parameter values of Fig. 1, with a small degradation rate d_A . In this case, the concentration of the free activator protein decreases very fast in time and the produced protein, after attaining a maximum value, has a slow decaying rate. Numerical analysis shows that the maximum value of protein concentrations depends on the magnitude of the degradation rate d_A . If $d_A \rightarrow 0$, Eq. (17) converges to Eq. (11), with an additional conservation law. At the asymptotic time limit the system ends up with the same amount of initial gene template, without protein and activator.

3.3. Self-activation

An important case occurs when the produced protein activates its own production. For this situation, the activation mechanism is:



and by the mass action law, we obtain the following system of differential equations:

$$\begin{aligned} \frac{dG}{dt} &= -k_1PG + k_2G_P \\ \frac{dG_P}{dt} &= k_1PG - k_2G_P \\ \frac{dP}{dt} &= aG_P - dP - k_1PG + k_2G_P \end{aligned} \tag{19}$$

As we have the conservation law, $G + G_P = G(0) + G_P(0) = c_2$, the independent set of equations describing the time evolution of the process (18) is:

$$\begin{aligned} \frac{dG}{dt} &= -k_1PG + k_2(c_2 - G) \\ \frac{dP}{dt} &= a(c_2 - G) - dP - k_1PG + k_2(c_2 - G) \end{aligned} \tag{20}$$

Eqs. (20) have two steady states with coordinates,

$$(G^{(1)}, P^{(1)}) = (c_2, 0)$$

and

$$(G^{(2)}, P^{(2)}) = \left(\frac{dk_2}{ak_1}, \frac{ac_2k_1 - dk_2}{dk_1} \right)$$

Calculating the Jacobian matrix of Eq. (20) at the fixed points, if $k_2d < k_1c_2a$, the fixed point $(G^{(1)}, P^{(1)})$ is of saddle type, and $(G^{(2)}, P^{(2)})$ is a stable node. As the determinant of the Jacobian matrices calculated at $(G^{(1)}, P^{(1)})$ and $(G^{(2)}, P^{(2)})$ have opposite signs, if $k_2d > k_1c_2a$, $(G^{(1)}, P^{(1)})$ is stable and $(G^{(2)}, P^{(2)})$ is unstable. Note that, in this last case, $P^{(2)}$ is negative.

In Fig. 4, we show the temporal evolution of the concentration of a self-activating protein, for two different initial protein concentrations.

If $k_1c_2a > k_2d$, a self-activating protein mechanism leads to a nonzero asymptotic protein concentration. As $c_2 = G(0) + G_P(0)$, having initially enough gene templates, the equilibrium values of the produced protein depend on the reaction rates of the processes involved in the transcriptional mechanisms, and on the concentrations of the gene template. If the initial concentration of the gene template is below a certain threshold, asymptotically in time, the protein concentration decays to zero. In this case, the conservation law derived from the mass action law implies the existence of a threshold effect. In Fig. 5, we have plotted the steady state value of the protein concentration as a function of the conserved quantity $c_2 = G(0) + G_P(0)$. The threshold occurs for $c_2 = dk_2/ak_1$.

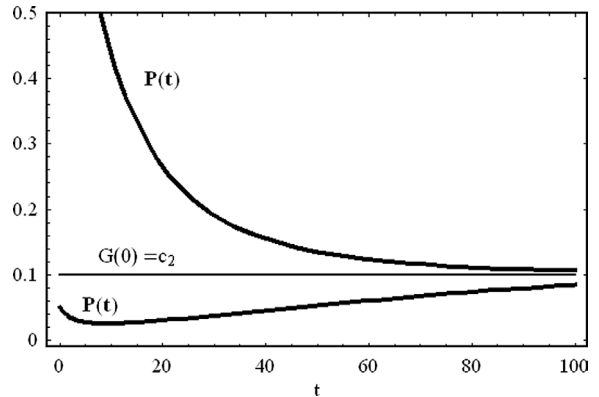


Fig. 4. Self-activation of protein production. We show the temporal evolution of protein concentration calculated from (20), for protein initial concentrations $P(0) = 1.0$ and $P(0) = 0.05$. At $t = 0$, the protein concentration is larger than zero, and the gene has no additional activating mechanisms. If the initial concentration of gene templates available for protein production is above some definite threshold value defined by $k_1c_2a > k_2d$, where $c_2 = G(0) + G_P(0)$, in the limit $t \rightarrow \infty$, the protein attains a constant value. The parameter values of the simulation are: $k_1 = 1.0$, $k_2 = 0.1$, $a = 0.2$ and $d = 0.1$, and the initial concentrations of the gene template is $G(0) = 0.1 = c_2$. If $k_1c_2a < k_2d$, the final protein concentration goes to zero.

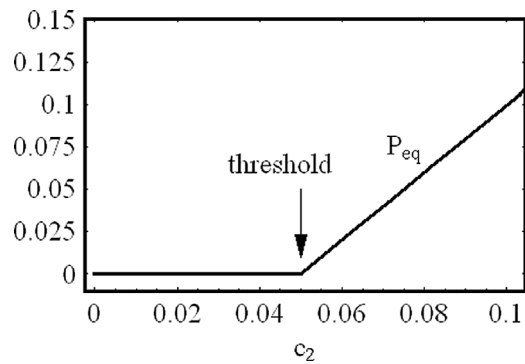


Fig. 5. Threshold effect for positive regulation with a self-activator, for the parameter values $k_1 = 1.0$, $k_2 = 0.1$, $a = 0.2$ and $d = 0.1$. When the conserved quantity c_2 is below dk_2/ak_1 , the equilibrium value ($t \rightarrow \infty$) of the protein concentration is $P_{eq} = P^{(1)} = 0$. If, $c_2 \geq dk_2/ak_1$, then $P_{eq} = P^{(2)} = (ac_2k_1 - dk_2)/dk_1$, and the threshold occurs for $c_2 = dk_2/ak_1$.

4. Negative regulation

4.1. The repressor total concentration is constant

In negative regulation of protein production, an inhibitor binds to an operator site and mRNA transcrip-

tion is not initiated. Representing by R the concentration of the repressor, the template gene by G, and the binded complex by G_R , the basic mechanism of protein production is now:



Applying the mass action law to the above reactions, we obtain the following set of differential equations:

$$\begin{aligned}
 \frac{dR}{dt} &= -k_3 R G + k_4 G_R \\
 \frac{dG}{dt} &= -k_3 R G + k_4 G_R \\
 \frac{dG_R}{dt} &= k_3 R G - k_4 G_R \\
 \frac{dP}{dt} &= a G - d P
 \end{aligned}
 \tag{22}$$

and the conservation laws:

$$\begin{aligned}
 R + G_R &= R(0) + G_R(0) = e_1 \\
 G + G_R &= G(0) + G_R(0) = e_2
 \end{aligned}
 \tag{23}$$

where $R(0)$ is the initial concentration of the protein inhibitor, and e_1 and e_2 are positive constants. Introducing the conservation laws (23) into (22), the independent set of equations describing protein inhibition is:

$$\begin{aligned}
 \frac{dR}{dt} &= -k_3 R(e_2 - e_1 + R) + k_4(e_1 - R) \\
 \frac{dP}{dt} &= a(e_2 - e_1 + R) - d P
 \end{aligned}
 \tag{24}$$

As in the case of activation, the system of Eqs. (24) has two fixed points with coordinates:

$$(R_p, P_p) \quad \text{and} \quad (R_n, P_n)
 \tag{25}$$

where,

$$\begin{aligned}
 R_{p,n} &= \frac{1}{2} \left(e_1 - e_2 - \frac{k_4}{k_3} \right) \\
 &\quad \pm \frac{1}{2} \sqrt{\left(e_1 - e_2 - \frac{k_4}{k_3} \right)^2 + 4e_1 \frac{k_4}{k_3}} \\
 P_{p,n} &= \frac{a}{d} (G(0) - R(0) + R_{p,n})
 \end{aligned}
 \tag{26}$$

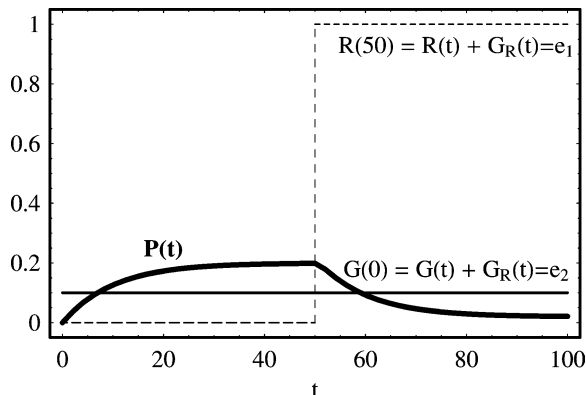


Fig. 6. Negative regulation of protein production with constant repressor concentration. We show the temporal evolution of protein concentration calculated from (27). From $t = 0$ to $t = 50$, the repressor is absent ($R(0) = 0.0$) and, therefore, protein production is initiated. At time $t = 50$, we have introduced a repressor $R(50) = 1.0$ and protein decay is initiated. After the introduction of the repressor, asymptotically in time, the protein concentration attains a low equilibrium value, $(R_p, P_p) = (0.910, 0.020)$. Parameter values of the simulation are: $k_3 = 1.0$, $k_4 = 0.1$, $a = 0.2$ and $d = 0.1$. Initial concentrations are: $P(0) = 0.0$, $G(0) = 0.1$ and $G_R(0) = 0.0$.

and the indexes p and n stand for positive and negative signs, respectively. The fixed point (R_p, P_p) is a stable node and (R_n, P_n) is a (unstable) saddle point.

The solution of the system of Eqs. (24) is:

$$\begin{aligned}
 R(t) &= \frac{R_p(R_n - R_0)e^{(R_p k_3 t)} + R_n(R_0 - R_p)e^{(R_n k_3 t)}}{(R_n - R_0)e^{(R_p k_3 t)} + (R_0 - R_p)e^{(R_n k_3 t)}} \\
 P(t) &= P(0)e^{-dt} + e^{-dt} \int_0^t e^{ds} R(s) ds
 \end{aligned}
 \tag{27}$$

where $R_0 = R(0)$. Comparing (24) with (13), for the inhibition and the activation mechanisms the time solutions are analytically similar.

In Fig. 6, we plot the time evolution of protein production as a function of time. In order to compare the effect of activation and repression agents, we have taken the same parameter values as in the case of Fig. 1. In this case, there exists protein production in the absence of repressor, but when we connect the repression mechanisms, the production process stops and asymptotically in time, the steady state of the protein reaches a low value but away from zero.

4.2. The repressor degrades

In the simple model (21), the repressor does not degrade during the process. If the repressor R is itself a protein, we must consider the extra mechanism accounting for repressor degradation,



Therefore, the first equation in (22) is now:

$$\frac{dR}{dt} = -k_3RG + k_4G_R - d_RR \quad (29)$$

In this case, we have only the second conservation law in (23), and the equations for protein production are:

$$\begin{aligned} \frac{dR}{dt} &= -k_3RG + k_4(e_2 - G) - d_RR \\ \frac{dG}{dt} &= -k_3RG + k_4(e_2 - G) \\ \frac{dP}{dt} &= aG - dP \end{aligned} \quad (30)$$

where we have eliminated G_R due to the second conservation law in (23). The system of Eqs. (30) has one stable fixed point with nonnegative coordinate $(R^*, G^*, P^*) = (0, e_2, ae_2/d) = (0, G(0) + G_R(0), a(G(0) + G_R(0))/d)$, where the constant e_2 is given by (23).

In the case of a repressor agent that degrades during time, the asymptotic protein concentration assumes a positive value, $P^* = a(G(0) + G_R(0))/d$. This contrasts with the case of a degrading activator, where protein concentration goes to zero, asymptotically in time.

4.3. Self-repression

An important case is when the produced protein represses its own production. For these situations, the mechanism is:



and by the mass action law, we obtain the following system of differential equations,

$$\frac{dG}{dt} = -k_3PG + k_4G_P$$

$$\frac{dG_P}{dt} = k_3PG - k_4G_P \quad (32)$$

$$\frac{dP}{dt} = aG - dP - k_3PG + k_4G_P$$

As we have the conservation law, $G(t) + G_P(t) = G(0) + G_P(0) = c_1$, the independent set of equations describing the time evolution of the process (31) is:

$$\begin{aligned} \frac{dG}{dt} &= -k_3PG + k_4(c_1 - G) \\ \frac{dP}{dt} &= aG - dP - k_3PG + k_4(c_1 - G) \end{aligned} \quad (33)$$

If $c_1 > 0$, the system of Eqs. (33) has one stable steady state, with coordinates:

$$(G^*, P^*) = \left(\frac{-dk_4 + \sqrt{d^2k_4^2 + 4ac_1dk_4}}{2ak_3}, \frac{-dk_4 + \sqrt{d^2k_4^2 + 4ac_1dk_4}}{2dk_3} \right) \quad (34)$$

For a zero initial concentration of protein and $c_1 > 0$, asymptotically in time, the protein concentration attains the value P^* , given by (34), Fig. 7. This case can be compared with the solution of nonregulated protein production in (7). With $k_3 \rightarrow 0$ and $c_1 = G(0)$, we obtain the same asymptotic state as in (7). As expected, protein self-repression makes protein production less efficient. For the same parameter values, the equilibrium value of protein concentration is lower in the case

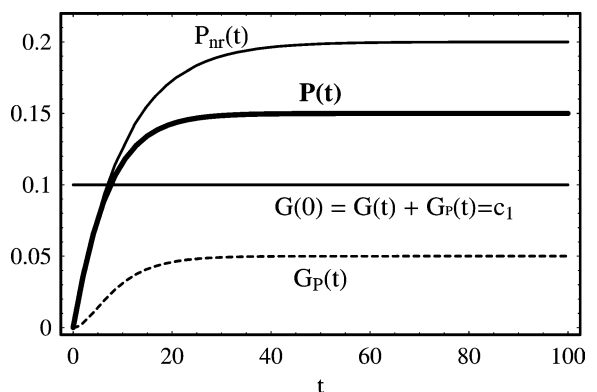


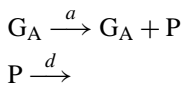
Fig. 7. Self-repression of protein production. $P(t)$ represents the temporal evolution of a self-repressing protein as described by mechanism (31). $P_{nr}(t)$ represents the temporal evolution of a non-regulated protein according to mechanism (5). Parameter values are: $k_3 = 1.0$, $k_4 = 0.1$, $a = 0.2$ and $d = 0.1$. Initial concentrations are: $P(0) = 0.0$ and $G(0) = 0.1$.

of self-repression, but this equilibrium state is reached faster than for the nonregulated case of protein production.

5. Positive and negative regulation with the regulators binding the same operator

We now extend this model to the case where the template DNA has overlapping binding site where an activator or an inhibitor can bind, leading to the activation or repression of protein production. This can be seen as the case of competition by the same binding site.

Adopting the same strategy of the preceding sections, we now have:



where protein production only occurs when the operon binding site of the DNA template is occupied by an activator. We can now proceed by applying mass action law to diagram (35), obtaining:

$$\begin{aligned} \frac{dA}{dt} &= -k_1AG + k_2G_A \\ \frac{dR}{dt} &= -k_3RG + k_4G_R \\ \frac{dG}{dt} &= -k_1AG + k_2G_A - k_3RG + k_4G_R \\ \frac{dG_A}{dt} &= k_1AG - k_2G_A \\ \frac{dG_R}{dt} &= k_3RG - k_4G_R \\ \frac{dP}{dt} &= aG_A - dP \end{aligned} \quad (36)$$

Eqs. (36) have the conservation laws:

$$\begin{aligned} G + G_A + G_R &= G(0) + G_A(0) + G_R(0) = f_1 \\ A + G_A &= A(0) + G_A(0) = f_2 \\ R + G_R &= R(0) + G_R(0) = f_3 \end{aligned} \quad (37)$$

Solving Eqs. (37) in order to get G , G_A and G_R , and eliminating these dependent variables from (36), we

obtain:

$$\begin{aligned} \frac{dA}{dt} &= -k_1A(A + R + f_1 - f_2 - f_3) + k_2(f_2 - A) \\ \frac{dR}{dt} &= -k_3R(A + R + f_1 - f_2 - f_3) + k_4(f_3 - R) \\ \frac{dP}{dt} &= a(f_2 - A) - dP \end{aligned} \quad (38)$$

The first two equations in (38) are independent of P , and we can determine their fixed points. Equating $\frac{dA}{dt}$ and $\frac{dR}{dt}$ to zero, and solving in order to get R and A , respectively, we obtain the equations for the nullclines:

$$\begin{aligned} R &= -A - f_1 + f_2 + f_3 - \frac{k_2}{k_1} + \frac{k_2f_2}{Ak_1} \\ \text{A nullcline} \\ A &= -R - f_1 + f_2 + f_3 - \frac{k_4}{k_3} + \frac{k_4f_3}{Rk_3} \\ \text{R nullcline} \end{aligned} \quad (39)$$

The nullclines (39) are monotonic functions in the first quadrant of the (A, R) phase space and intersect at only one point, say, (A^*, R^*) . The phase space point (A^*, R^*) is the only fixed point with positive coordinates of system (38). Calculating the Jacobian matrix

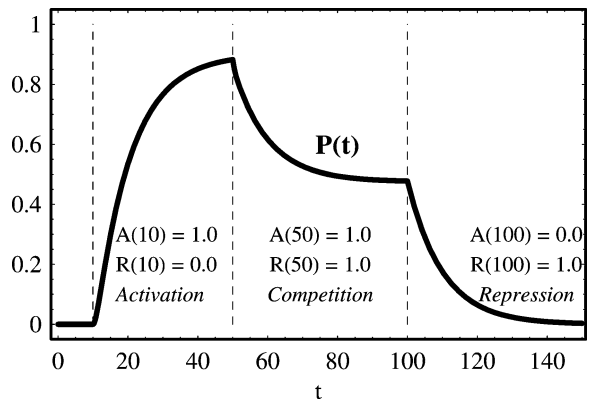


Fig. 8. Competition of a positive and a negative regulator for the same binding site. We show the temporal evolution of protein concentration calculated from (38). From $t = 0$ to $t = 10$, both the activator and the repressor are absent. From $t = 10$ to $t = 50$, the activator A is present. From $t = 50$ to $t = 100$, both the activator A and the repressor R are present. From $t = 100$ to $t = 150$, only the repressor R is present. Other parameter values of the simulation are: $k_1 = 1.0$, $k_2 = 0.1$, $k_3 = 1.0$, $k_4 = 0.1$, $a = 1.0$ and $d = 0.1$. Initial concentrations are: $P(0) = 0.0$, $G(0) = 0.1$ and $G_A(0) = G_R(0) = 0.0$.

at (A^*, R^*) of the first two equations in (39), we obtain:

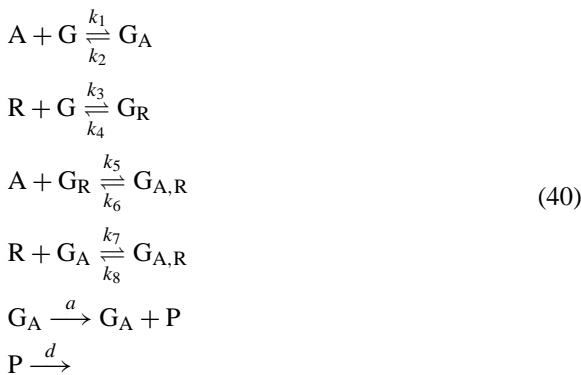
$$M = \begin{pmatrix} -k_1 A^* - \frac{k_2 f_2}{A^*} & -k_1 A^* \\ -k_3 R^* & -k_3 R^* - \frac{k_4 f_3}{R^*} \end{pmatrix}$$

As, $\text{Det } M = Ak_1 k_4 f_3 / R + k_2 k_4 f_2 f_3 / AR + Rk_2 k_3 f_2 / A > 0$, and $\text{Tr } M < 0$, the fixed point (A^*, R^*) is always asymptotically stable and the system of Eqs. (38) has a unique stable steady state with positive coordinates.

In the simulation of Fig. 8, we show that when a repressor is present, even in small concentrations, and competes for the same binding site as the activator, its effect is to decrease the equilibrium value of the produced protein.

6. Positive and negative regulation with the regulators binding different operators

For the case where the template DNA has two operon sites, one for the activator and another for the inhibitor, the reaction mechanisms are:



where protein production only occurs when the operon binding site of the DNA template is occupied by an activator and the repressor binding site is free. We can now proceed by applying mass action law to diagram (40), obtaining:

$$\begin{aligned} \frac{dA}{dt} &= -k_1 A G + k_2 G_A - k_5 A G_R + k_6 G_{A,R} \\ \frac{dR}{dt} &= -k_3 R G + k_4 G_R - k_7 R G_A + k_8 G_{A,R} \\ \frac{dG}{dt} &= -k_1 A G + k_2 G_A - k_3 R G + k_4 G_R \\ \frac{dG_A}{dt} &= k_1 A G - k_2 G_A - k_7 R G_A + k_8 G_{A,R} \end{aligned} \tag{41}$$

$$\begin{aligned} \frac{dG_R}{dt} &= k_3 R G - k_4 G_R - k_5 A G_R + k_6 G_{A,R} \\ \frac{dG_{A,R}}{dt} &= k_5 A G_R - k_6 G_{A,R} + k_7 R G_A - k_8 G_{A,R} \\ \frac{dP}{dt} &= a G_A - dP \end{aligned}$$

Eqs. (41) obey the conservation laws:

$$\begin{aligned} G + G_A + G_R + G_{A,R} &= g_1 \\ A + G_A + G_{A,R} &= g_2 \\ R + G_R + G_{A,R} &= g_3 \end{aligned} \tag{42}$$

Solving Eqs. (42) in order to get G_A , G_R and $G_{A,R}$, and eliminating the dependent variables from (41), we obtain:

$$\begin{aligned} \frac{dA}{dt} &= -k_1 A G + k_2 G_A - k_5 A G_R + k_6 G_{A,R} \\ \frac{dR}{dt} &= -k_3 R G + k_4 G_R - k_7 R G_A + k_8 G_{A,R} \\ \frac{dG}{dt} &= -k_1 A G + k_2 G_A - k_3 R G + k_4 G_R \\ \frac{dP}{dt} &= a G_A - dP \end{aligned} \tag{43}$$

where $G_A = g_1 - g_3 - G + R$, $G_R = g_1 - g_2 + A - G$ and $G_{A,R} = g_2 - g_1 + g_3 - A - R$.

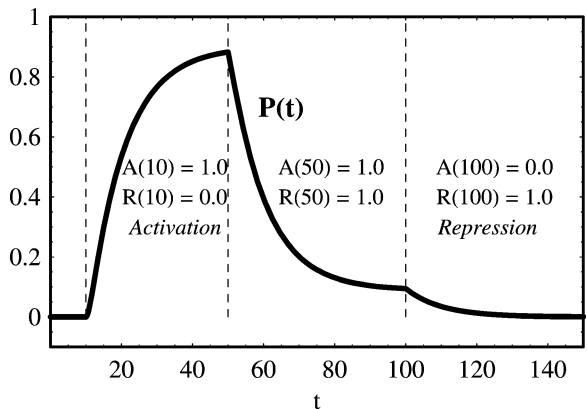
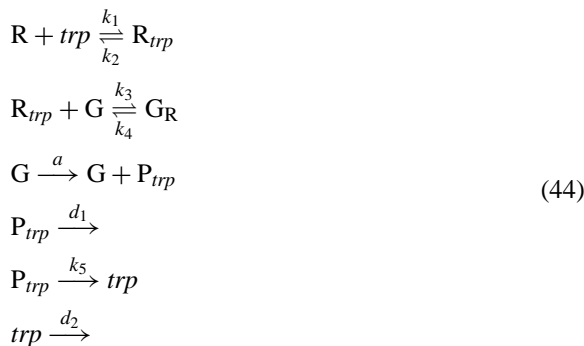


Fig. 9. Positive and negative regulation of protein production with the regulators binding different operators. We show the temporal evolution of protein concentration calculated from (43). From $t = 0$ to $t = 10$ both the activator and the repressor are absent. From $t = 10$ to $t = 50$ the activator A is present. From $t = 50$ to $t = 100$ both an activator A and a repressor R are present. From $t = 100$ to $t = 150$ only the repressor R is present. Parameter values are: $k_1 = k_3 = k_5 = k_7 = 1.0$, $k_2 = k_4 = k_6 = k_8 = 0.1$, $a = 1.0$ and $d = 0.1$. Initial concentrations are: $P(0) = 0.0$, $G(0) = 0.1$, $G_A(0) = G_R(0) = G_{A,R}(0) = 0.0$.

In the simulation of Fig. 9, we show that when the activator and the repressor have different binding sites, the effect of the repressor is more pronounced when compared with the competitive case of Fig. 8.

7. The negative regulation of the *trp* operon

The tryptophan (*trp*) operon is an example of a repressible system. The mRNA coded by the *trp* operon is translated into five polypeptides that assemble the enzymes responsible for the biosynthesis of the aminoacid tryptophan. Both the synthesis and the activity of these enzymes are controlled by the level of tryptophan in the cell. The cell also produces a repressor protein (R) that is only active if bounded to a tryptophan molecule (R_{trp}). If the levels of tryptophan are low inside the cell, the gene of the *trp* operon (G) is expressed by default. When the tryptophan accumulates in the cell, the repressor protein is activated and the expression of the genes controlled by the *trp* operon are inhibited. There are other mechanisms involved in the regulation of the tryptophan synthesis, as transcriptional attenuation and feedback regulation of the enzymes activity. For another model, using ad hoc regulatory functions, see [13]. Here we are focusing only on the mechanism of negative regulation. This situation is similar to the general case presented in Section 4.3 with two additional mechanism. One mechanism accounts for the induction of the repressor, and the other mechanism is a simplified representation of the biosynthesis of the aminoacid *trp*. The simplified kinetics of this regulation mechanism is,



where G represents the gene of the *trp* operon, G_R is the repressed *trp* operon (R_{trp} bounded to the *trp*

operon), and P_{trp} represents the enzymes and precursors responsible for *trp* biosynthesis. By the mass action law, we obtain the following system of differential equations,

$$\begin{aligned}
 \frac{dR}{dt} &= -k_1 R_{trp} + k_2 R_{trp} \\
 \frac{dR_{trp}}{dt} &= k_1 R_{trp} - k_2 R_{trp} - k_3 R_{trp} G + k_4 G_R \\
 \frac{dP_{trp}}{dt} &= a G - d_1 P_{trp} - k_5 P_{trp} \\
 \frac{dG}{dt} &= -k_3 R_{trp} G + k_4 G_R \\
 \frac{dG_R}{dt} &= k_3 R_{trp} G - k_4 G_R \\
 \frac{dtrp}{dt} &= -k_1 R_{trp} + k_2 R_{trp} + k_5 P_{trp} - d_2 trp
 \end{aligned} \tag{45}$$

with the conservation laws,

$$\begin{aligned}
 R + R_{trp} + G_R &= h_1 \\
 G + G_R &= h_2
 \end{aligned} \tag{46}$$

Solving Eqs. (46) in order to get G_R and R_{trp} , and eliminating the dependent variables from (45), we obtain,

$$\begin{aligned}
 \frac{dR}{dt} &= -k_1 R_{trp} + k_2 (h_1 - h_2 + G - R) \\
 \frac{dP_{trp}}{dt} &= a G - d_1 P_{trp} - k_5 P_{trp} \\
 \frac{dG}{dt} &= -k_3 (h_1 - h_2 + G - R) G + k_4 (h_2 - G) \\
 \frac{dtrp}{dt} &= -k_1 R_{trp} + k_2 (h_1 - h_2 + G - R) \\
 &\quad + k_5 P_{trp} - d_2 trp
 \end{aligned} \tag{47}$$

where $G_R = h_2 - G$, and $R_{trp} = h_1 - h_2 + G - R$.

The system of Eqs. (47) has two fixed points in phase space. If $h_1 > 0$ and $h_2 > 0$, one of the fixed points has positive coordinates and is stable. The other fixed point has negative coordinates. At the steady state, the level of the aminoacid *trp* depends on the initial concentration of the nonrepressed gene, and on the rate at which *trp* is transcribed and degraded. In Fig. 10, we show the time evolution of the concentration of the aminoacid tryptophan from two different initial conditions. One initial condition is zero and the other is positive. For low *trp* initial concentration, the aminoacid is continuously produced, reaching a steady

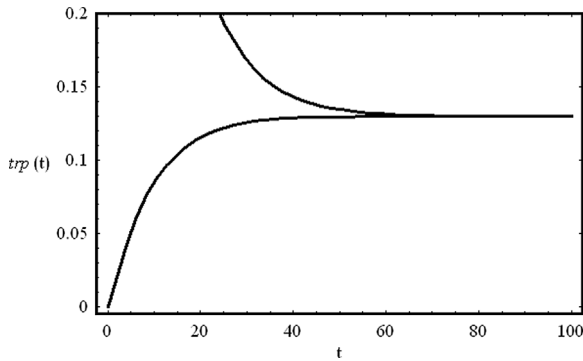


Fig. 10. The negative regulation of the *trp* operon. We show the temporal evolution of the concentrations of the amino acid *trp* from two different initial conditions ($trp(0) = 0.0$ and $trp(0) = 1.0$). The production of the amino acid tryptophan is self-repressed as described by the mechanism (44). Parameter values are: $k_1 = 1.0$, $k_2 = 0.1$, $k_3 = 1.0$, $k_4 = 0.1$, $k_5 = 1.0$, $a = 0.2$ and $d_1 = d_2 = 0.1$. Initial concentrations are: $G(0) = 0.1$, $R(0) = 0.1$ and $P_{trp}(0) = 0.0$.

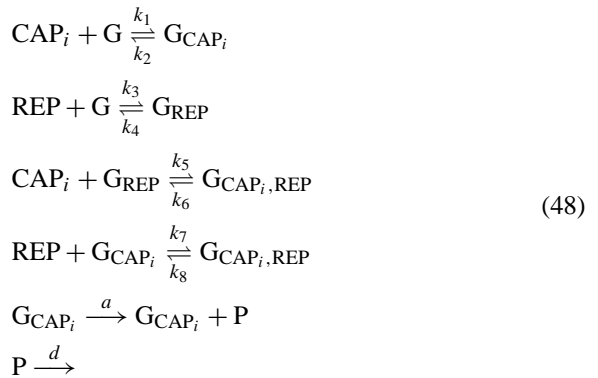
state. For higher concentrations, the repressor protein inhibits its production, and the same steady state is reached. The proposed mechanism shows how self-repression maintains stable levels of *trp* inside de cell.

8. The genetic switch in the regulation of the *lac* operon

In *Escherichia coli* the main source of carbon is glucose. Lactose is only used as an alternative in the absence of glucose. The *lac* operon controls the expression of a set of genes that code for proteins involved in the metabolism of the disaccharide lactose. In the presence of glucose, the *lac* operon is off and the genes under its control are not transcribed. The *lac* operon (G) is controlled by a transcriptional activator, the ‘catabolite activator protein’ (CAP), and a transcriptional repressor, the ‘*lac* repressor’ (REP), binding different operators. A decrease in glucose concentration inside the cell leads to the rise of CAP induced by cAMP (CAP_i). In this form, CAP_i is able to bind the DNA and the *lac* operon can then be turned on. If the glucose concentration increases, cAMP decreases and CAP cannot bind to the DNA, turning the *lac* operon off. The ‘*lac* repressor’ binds to the *lac* operon in the absence of lactose, turning off the operon, even if CAP_i is bound. The presence of lactose induces the increase in allolactose concentration that

binds to the *lac* repressor protein, removing it from the DNA. CAP_i can only turn on the *lac* operon if the *lac* repressor is not repressing it. This means that *lac* operon is turned on only if glucose is absent and lactose is present. Under the control of the *lac* operon the main genes that are turned on encode permease and β -galactosidase enzymes, responsible for the lactose transport and hydrolysis.

In this framework, the regulation of the *lac* operon is described by the mechanism:



Protein production only occurs when the operon has a binded activator, but not a repressor (G_{CAP_i}). Comparing (48) with (40), with the identification, $CAP_i = A$,

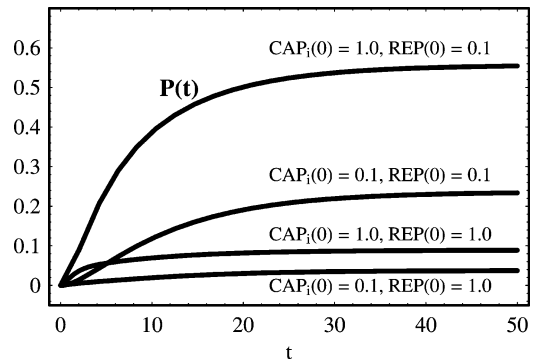


Fig. 11. The regulation of the *lac* operon. This is a case of positive and negative regulation of protein production, with the regulators binding different operators. We show the temporal evolution of protein concentration for different initial concentrations of the activator CAP_i and the repressor REP. Parameter values are: $k_1 = k_3 = k_5 = k_7 = 1.0$, $k_2 = k_4 = k_6 = k_8 = 0.1$, $a = 1.0$ and $d = 0.1$. Initial concentrations are: $P(0) = 0.0$, $G(0) = 0.1$ and $G_{CAP_i}(0) = G_{REP}(0) = G_{CAP_i,REP}(0) = 0.0$. It is clear that the *lac* operon is significantly activated only when the concentration of the activator is at least about ten times higher than the concentration of the repressor.

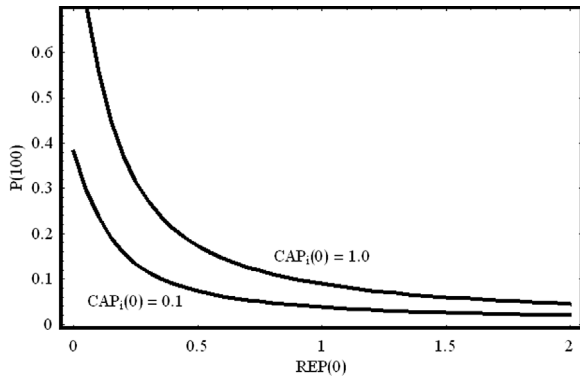


Fig. 12. On/off switch of the *lac* operon. We have plotted the protein concentration for $t = 100$ as a function of the initial concentration of the repressor REP. For fixed activator initial concentration, as the repressor concentration increases linearly, the equilibrium value of the produced protein decreases exponentially. We have chosen two different initial conditions for the activator CAP_i . The other parameter values are the same as in Fig. 11.

REP = R, the evolution equations of concentrations are the same as Eqs. (41), with the conservation laws:

$$\begin{aligned} G + G_{CAP_i} + G_{REP} + G_{CAP_i,REP} &= g_1 \\ CAP_i + G_{CAP_i} + G_{CAP_i,REP} &= g_2 \\ REP + G_{REP} + G_{CAP_i,REP} &= g_3 \end{aligned} \quad (49)$$

The simulations in Fig. 11 show that the *lac* operon behaves like an ‘on/off soft switch’: the operon is turned on only when the concentration of CAP_i is high and the concentration of REP is low. These conditions only occur if glucose is absent and lactose is present. In Fig. 12 we show the steady state value of protein concentration as a function of the initial concentration of the repressor REP.

9. Discussion and conclusions

We have developed a class of differential equation models describing the basic biological mechanisms of regulation of gene expression. The models are easy to handle, enabling a straightforward comparison with experiments in order to be validated and calibrated. The basic assumptions in the construction of this simple framework are the following:

- (1) genes are considered as templates for protein production;

- (2) gene transcription is activated or repressed by transcriptional regulators binding to the regulator sites of the gene;
- (3) all the mechanisms are bi-molecular and derived from the mass action law;
- (4) both transcriptional and translational mechanisms are described in time by the same rate constants.

In this approach, the interactions between regulators and inducer or co-repressor molecules, cooperative effects, and multimerization of regulatory proteins are not taken into account. However, these mechanisms can also be described using this mathematical framework.

We have explored and explained the quantitative behaviour of basic simple gene regulatory mechanisms. For example, when we compare nonregulated protein production with activated protein production, the equilibrium protein values become dependent of the activator initial concentration, and therefore, are not only limited by the rate constants of the production mechanisms. In the case of nonregulated protein production, the equilibrium values are limited by the rate at which production and degradation mechanisms occur.

If the protein activator concentration degrades during time, the produced protein, after reaching a maximum concentration value, decays to zero. This contrasts with the case of protein self-activation, where the equilibrium values of the produced protein depend on the reaction rates of the processes involved in the production mechanisms and on the concentrations of the gene template. In this last case, if the concentration of the gene template is below a certain threshold, protein concentration decays to zero.

When the protein is produced by default and exists a negative regulation mechanism, the equilibrium value of the protein concentration is lower in the presence of the repressor and depends on the repressor initial concentration. If in addition the repressor molecules degrade, the resulting protein concentration reaches a higher asymptotic positive value. In the case of self-repression, the produced protein reaches lower equilibrium values when compared with the nonregulated case. Without regulation, protein concentration at equilibrium only depends on the reaction rates of the reactions.

It follows from these simple models that the threshold effect observed in some experimental systems is associated with the existence of conserved quantities derived from the mass action law. In general, the conserved quantities change the position of fixed points in phase space and, in some cases, leads to bifurcation of equilibria. In particular, in the case of protein production by self-activation, depending on the initial concentration of the gene, different asymptotic states of the protein are reached (Fig. 5).

We have also analyzed the two cases where both an activator and a repressor are regulating the production of the same protein. In the simulation of Fig. 8, we can see that when a repressor is present, and competes for the same binding site of the activator, its effect is to decrease the equilibrium value of the produced protein. When the activator and the repressor have different binding sites, Fig. 9, the effect of the repressor is more pronounced, leading to even lower protein concentrations.

The models of positive and negative regulation of protein production are the building blocks for more complex mechanisms of gene regulatory interactions, as we have shown in the cases of the *trp* and the *lac* operon. Analyzing the mechanism of self-repression of the tryptophan synthesis, we have shown how this process contributes for the regulation of the aminoacid concentration inside the cell. It is shown that the regulation mechanism of the *lac* operon behaves like a soft on/off switch, and we have analyzed the dependence on the concentrations of the activator and the repressor.

These models depend on biologically relevant parameters, are robust to the variation of the parameters and have low computational cost. For complex networks of gene regulation, the approach developed here enables the construction of general computer programs for the derivation and integration of the equations describing the temporal (and spatial) evolution of the protein concentrations. The predicted gene expression patterns can then be compared with the real experimental system in order to be validated and calibrated.

Acknowledgements

This work has been partially supported by Fundação para a Ciência e a Tecnologia (Portugal), under

the framework of POCTI project C/FIS/13161/1998, and by the grant PRAXIS XXI/BD/18379/98.

References

- [1] S. Fields, Y. Kohara, D.J. Lockhart, Functional genomics, Proc. Natl Acad. Sci. USA 96 (1999) 8825–8826.
- [2] A.M. Huerta, H. Salgado, D. Thieffry, J. Collado-Vides, RegulonDB: a database on transcriptional regulation in *Escherichia coli*, Nucl. Acids Res. 26 (1998) 55–59.
- [3] M. Kanehisa, S. Goto, KEGG: Kyoto encyclopedia of genes and genomes, Nucl. Acids Res. 28 (2000) 27–30.
- [4] P.D. Karp, M. Riley, M. Saier, I.T. Paulsen, J. Collado-Vides, S.M. Paley, A. Pellegrini-Toole, C. Bonavides, S. Gama-Castro, The EcoCyc database, Nucl. Acids Res. 30 (2002) 56–58.
- [5] A.V. Spirov, T. Bowler, J. Reinitz, HOX Pro: a specialized database for clusters and networks of homeobox genes, Nucl. Acids Res. 28 (2000) 337–340.
- [6] J. Hastay, D. McMillen, F. Isaacs, J.J. Collins, Computational studies of gene regulatory networks: in numero molecular biology, Nat. Rev. – Genetics 2 (2001) 268–279.
- [7] H.D. Jong, Modelling and simulations of genetic regulatory systems: a literature review, J. Comput. Biol. 9 (2002) 67–103.
- [8] P. Smolen, D.A. Baxter, J.H. Byrne, Mathematical modelling of gene networks, Neuron 26 (2000) 567–580.
- [9] M.A. Savageau, Design principles for elementary gene circuits: elements, methods, and examples, Chaos 11 (2001) 142–159.
- [10] D.E. Koshland Jr., A. Goldbeter, J.B. Stock, Amplification and adaptation in regulatory and sensory systems, Science 217 (1982) 220–225.
- [11] M.A. Savageau, Design of gene circuitry by natural selection: analysis of the lactose catabolic system in *Escherichia coli*, Biochem. Soc. Trans. 27 (1999) 264–270.
- [12] M. Ronen, R. Rosenberg, B.I. Shraiman, U. Alon, Assigning numbers to the arrows: parameterizing a gene regulation network by using accurate expression kinetics, Proc. Natl Acad. Sci. USA 99 (2002) 10555–10560.
- [13] M. Santillán, M.C. Mackey, Dynamic regulation of the tryptophan operon: a modelling study and comparison with experimental data, Proc. Natl Acad. Sci. USA 98 (2001) 1364–1369.
- [14] J.-M. Claverie, Computational methods for the identification of differential and coordinated gene expression, Human Mol. Genet. 8 (1999) 1821–1832.
- [15] H. Herzel, D. Beule, S. Kielbasa, J. Korbel, C. Sers, A. Malik, H. Eickhoff, H. Lehrach, J. Schuchhardt, Extracting information from cDNA arrays, Chaos 11 (2001) 98–107.
- [16] T.E. Ideker, V. Thorsson, R.M. Karp, Discovery of regulatory interactions through perturbation: inference and experimental design, Pac. Symp. Biocomput. 5 (2000) 302–313.
- [17] D.M. Mutch, A. Berger, R. Mansourian, A. Rytz, M.A. Roberts, The limit fold change model: a practical approach for selecting differentially expressed genes from microarray data, BMC Bioinformatics 3 (2002) 17–27.

- [18] D. Thieffry, H. Salgado, A.M. Huerta, J. Collado-Vides, Prediction of transcriptional regulatory sites in the complete genome sequence of *Escherichia coli* K-12, *Bioinformatics* 14 (1998) 391–400.
- [19] A. Becskei, B. Seraphin, L. Serrano, Positive feedback in eukaryotic gene networks: cell differentiation by graded to binary response conversion, *EMBO J.* 20 (2001) 2528–2535.
- [20] M.B. Elowitz, S. Leibler, A synthetic oscillatory network of transcriptional regulators, *Nature* 403 (2000) 335–338.
- [21] T.S. Gardner, C.R. Cantor, J.J. Collins, Construction of a genetic toggle switch in *Escherichia coli*, *Nature* 403 (2000) 339–342.
- [22] J. Hasty, F. Isaacs, M. Dolnik, D. McMillen, J.J. Collins, Designer gene networks: towards fundamental cellular control, *Chaos* 11 (2001) 207–220.
- [23] J. Hasty, D. McMillen, J.J. Collins, Engineered gene circuits, *Nature* 420 (2002) 224–230.
- [24] H.H. McAdams, A. Arkin, Gene regulation: towards a circuit engineering discipline, *Curr. Biol.* 10 (2000) R318–R320.
- [25] M.A. Savageau, Rules for the evolution of gene circuitry, *Pac. Symp. Biocomput.* 3 (1998) 54–65.
- [26] T. Akutsu, S. Miyano, S. Kuhara, Inferring qualitative relations in genetic networks and metabolic pathways, *Bioinformatics* 16 (2000) 727–734.
- [27] T. Akutsu, S. Miyano, S. Kuhara, Algorithms for inferring qualitative models of biological networks, *Pac. Symp. Biocomput.* 5 (2000) 290–301.
- [28] H.H. McAdams, L. Shapiro, Circuit simulation of genetic networks, *Science* 269 (1995) 650–656.
- [29] D. Thieffry, From global expression data to gene networks, *BioEssays* 21 (1999) 895–899.
- [30] D. Thieffry, D. Romero, The modularity of biological regulatory networks, *BioSystems* 50 (1999) 49–59.
- [31] D. Thieffry, R. Thomas, Qualitative analysis of gene networks, *Pac. Symp. Biocomput.* 3 (1998) 77–88.
- [32] A. Wuensche, Genomic regulation modeled as a network with basins of attraction, *Pac. Symp. Biocomput.* 3 (1998) 89–102.
- [33] A. Arkin, P. Shen, J. Ross, A test case of correlation metric construction of a reaction pathway from measurements, *Science* 277 (1997) 1275–1279.
- [34] T. Chen, H.L. He, G.M. Church, Modelling gene expression with differential equations, *Pac. Symp. Biocomput.* 4 (1999) 29–40.
- [35] P. D’Haeseleer, Linear modelling of mRNA expression levels during CNS development and injury, *Pac. Symp. Biocomput.* 4 (1999) 41–52.
- [36] E. Mjolsness, T. Mann, R. Castaño, B. Wold, From coexpression to coregulation: an approach to inferring transcriptional regulation among gene classes from large-scale expression data, *Adv. Neural Inform. Process. Syst.* 12 (2000) 928–934.
- [37] D.C. Weaver, C.T. Workman, G.D. Stormo, Modelling regulatory networks with weight matrices, *Pac. Symp. Biocomput.* 4 (1999) 112–123.
- [38] D.A. Drew, A mathematical model for prokaryotic protein synthesis, *Bull. Math. Biol.* 63 (2001) 329–351.
- [39] F. Jacob, J. Monod, Genetic regulatory mechanisms in the synthesis of proteins, *J. Mol. Biol.* 3 (1961) 318–356.
- [40] R.D. Bliss, P.R. Painter, A.G. Marr, Role of feedback inhibition in stabilizing the classical operon, *J. Theor. Biol.* 97 (1982) 177–193.
- [41] D.M. Wolf, F.H. Eeckman, On the relationship between genomic regulatory element organization and gene regulatory dynamics, *J. Theor. Biol.* 195 (1998) 167–186.
- [42] T.A. Carrier, J.D. Keasling, Investigating autocatalytic gene expression systems through mechanistic modeling, *J. Theor. Biol.* 201 (1999) 25–36.
- [43] N.G. Van Kampen, *Stochastic Processes in Physics and Chemistry*, Elsevier, Amsterdam, 1992.
- [44] A.I. Volpert, V.A. Volpert, V.A. Volpert, *Traveling Wave Solutions of Parabolic Systems*, American Mathematical Society, Providence, RI, 1994.
- [45] A. Travers, Transcriptional switches: the role of the mass action, *Phys. Life Rev.* 1 (2004) 57–69.
- [46] H. Haken, *Synergetics. An Introduction*, Springer-Verlag, Berlin, 1983.