



ELSEVIER

Available online at www.sciencedirect.com

SCIENCE @ DIRECT®

C. R. Biologies 327 (2004) 211–224



Biological modelling / Biomodélisation

Modeling operon dynamics: the tryptophan and lactose operons as paradigms

Michael C. Mackey^{a,*}, Moisés Santillán^b, Necmettin Yildirim^c

^a *Departments of Physiology, Physics and Mathematics, Centre for Nonlinear Dynamics, McGill University, 3655 Drummond Street, Montreal, Quebec, Canada H3G 1Y6*

^b *Esc. Sup. de Física y Matemáticas, Instituto Politécnico Nacional, 07738, México D.F., Mexico*

^c *Atatürk Üniversitesi, Bilgisayar Bilimleri Uygulama ve Araştırma Merkezi, 25240 Erzurum, Turkey*

Received 15 April 2003; accepted after revision 26 November 2003

Presented by Pierre Auger

Abstract

Understanding the regulation of gene control networks and their ensuing dynamics will be a critical component in the understanding of the mountain of genomic data being currently collected. This paper reviews recent mathematical modeling work on the tryptophan and lactose operons which are, respectively, the classical paradigms for repressible and inducible operons. **To cite this article:** *M.C. Mackey et al., C. R. Biologies 327 (2004).*

© 2004 Published by Elsevier SAS on behalf of Académie des sciences.

Résumé

Modélisation de la dynamique de l'opéron : les opérons tryptophane et lactose comme paradigmes. L'étude de la régulation des réseaux de contrôles des gènes et des dynamiques qui en découlent sera une composante critique de la compréhension de la masse de données génomiques collectées. Cet article fait le bilan des récents travaux de modélisation mathématique sur les opérons tryptophanes et lactoses, qui sont respectivement les paradigmes classiques pour les opérons répressibles et inductibles. **Pour citer cet article :** *M.C. Mackey et al., C. R. Biologies 327 (2004).*

© 2004 Published by Elsevier SAS on behalf of Académie des sciences.

Keywords: gene regulatory networks; nonlinear dynamics; computational molecular biology

Mots-clés: circuits de régulation des gènes ; dynamique non linéaire ; biologie moléculaire informatique

1. Introduction

Molecular biology as a field would probably not exist in its present form were it not for the impact

of the beautiful little book *What is Life?* [1] written by Ernst Schrödinger (one of the fathers of quantum mechanics). *What is Life?* was the written account of a series of lectures given by Schrödinger at the Dublin Institute for Advanced Studies in 1943. *What is Life?* was partially inspired by the work of the physicist Max Delbrück and was probably instrumental in recruiting

* Corresponding author.

E-mail address: mackey@cnd.mcgill.ca (M.C. Mackey).

a whole generation of *theoretical physicists* like Walter Gilbert, Leo Szilard, Seymour Benzer, and Francis Crick away from physics *per se* and into the exciting new area of ‘molecular biology’ [2]. The results are history, but it has not ended there. For example, the current director of the Whitehead Center for Genomic Research at the MIT is Eric Lander, whose D. Phil. from Oxford was not in molecular biology but rather in mathematics (algebraic coding theory)! The mathematical and physical sciences continue to play a large role in the whole area of molecular biology and some of the potential areas of impact are detailed in the March 2001 issue of *Chaos* [3].

The operon concept [4], introduced by Jacob and coworkers in 1960 [5], has had a profound and lasting effect on the biological sciences. Not long after the operon concept was developed, Goodwin [6] gave the first mathematical analysis of operon dynamics. Griffith then put forward a more complete analysis of simple repressible (negative feedback [7]) and inducible (positive feedback [8]) gene control networks, and Tyson and Othmer [9] have summarized these results. Extensions considering the stability of inducible operons were published by Selgrade [10, 11] and Ji-Fa [12], but none of these treatments considered the role of the DNA transcription and mRNA translation delays though Tyson and Othmer pointed out that both should be considered.

In this paper we present recent mathematical modeling work on operon dynamics. In Section 2 we treat the (repressible) tryptophan operon, and in Section 3 the (inducible) lactose operon is considered. In both cases we have made every effort to construct biologically realistic mathematical models and to make accurate parameter estimation from the biological literature. In Section 4, we conclude the paper with some general comments related to our philosophical approach to the modeling of these systems.

2. The repressible tryptophan operon

After Jacob et al. introduced the operon concept [5], the development of more refined experimental techniques and the acquisition of more data made it clear that in addition to repression there are many regulatory mechanisms involved in the control of operons. These mechanisms vary from operon to

operon. One of the most intensively studied of these systems from an experimental point of view is the tryptophan operon. An excellent review of this system and its regulatory mechanisms is to be found in [13].

2.1. The Bliss model

One of the first mathematical models of the tryptophan operon was introduced by Bliss et al. in 1982 [14]. The independent variables of this model are the *trp mRNA* concentration (M), the concentration of anthranilate synthase (E) which, according to Bliss et al., of all the enzymes formed with the polypeptides of the tryptophan operon is the most important from a regulatory point of view, and the tryptophan concentration (T). The equations governing the dynamic evolution of these three quantities are

$$\frac{dM}{dt} = K_m O R(t - \tau_m) - K_1 M \quad (1)$$

$$\frac{dE}{dt} = K_p M(t - \tau_p) - K_2 E \quad (2)$$

and

$$\frac{dT}{dt} = K_t E I(T) - G(T) - K T \quad (3)$$

In Eq. (1), K_m is the intrinsic rate of transcription initiation of an operon that is not repressed, $R(T)$ is the probability that an operon is not repressed at time t , τ_m is the time delay between initiation of transcription and initiation of translation, O is the total operon concentration, and K_1 is a positive constant accounting for *mRNA* depletion due to dilution by growth and enzymatic degradation.

Bliss et al. assumed that R can be modeled by a Hill function

$$R(T) = \frac{K_r^m}{K_r^m + T^m}$$

with $m = 4$. In Eq. (2), K_p denotes the rate of translation initiation per *mRNA* molecule, τ_p is the time delay between initiation of translation and the appearance of functional enzyme, and K_2 is a positive constant accounting for enzyme depletion due to dilution by growth and enzymatic degradation. Finally, in Eq. (3), K_t is the tryptophan production rate per enzyme, $I(T)$ is the fraction of enzyme not inhibited by the end product, $G(T)$ is the tryptophan

consumption rate, and K is the growth rate of the bacterial culture.

This model clearly considers two different regulatory mechanisms, repression and enzymatic feedback inhibition by tryptophan. $R(T)$ accounts for repression, while $I(T)$ stands for feedback inhibition. Bliss et al. also assert that $I(T)$ can be modeled by a Hill function

$$I(T) = \frac{K_i^n}{K_i^n + T^n}$$

with $n = 2$, and that the demand for tryptophan obeys a Michaelis–Menten-like equation

$$G(T) = G_{\max} \frac{T}{T + K_g}$$

The authors made a careful estimation of the parameters in this model based on the available experimental data. They solved the model equations numerically and also performed an analytical stability analysis of the steady states. They were able to reproduce the results of derepression experiments with cultures of wild and mutant *E. coli* strains reported in [15]. From these experiments, the tryptophan operon loses stability in mutant *E. coli* cell cultures that have a partial loss of feedback inhibition. The previously stable steady state is replaced by an oscillatory production of tryptophan [15]. This mutation is modeled by increasing K_i to ten times its wild type value. However, to our knowledge, these experiments have never been successfully repeated. Moreover, more recent experimental evidence has demonstrated that the function $R(T)$, which Bliss et al. used to model repression, is in disagreement with the experimental facts about the interaction between the *trp* operon and repressor molecules. The issue of oscillations in tryptophan control was revisited by Xiu et al. [16] in their investigation of the role of growth and dilution rates on stability.

2.2. The Sinha model

In 1988, Sinha [17] introduced a different model for the tryptophan operon regulatory system in which the DNA-repressor interaction is modelled in a more detailed way. The independent variables of the Sinha model are the same as in the Bliss et al. model, i.e. the *trp* mRNA concentration (M), the anthranilate synthase enzyme concentration (E), and the tryptophan

concentration (T). The dynamic equations for these variables are

$$\frac{dM}{dt} = K_m \frac{K_o K_d + K_o T}{K_o K_d + K_o T + nRT} O - K_1 M \quad (4)$$

$$\frac{dE}{dt} = K_p M - K_2 E \quad (5)$$

and

$$\frac{dT}{dt} = K_t E - G_{\max} - K T \quad (6)$$

The first term in the right-hand side of Eq. (4) is the mRNA production rate, which is assumed to be proportional to the concentration of unrepressed operons, with a proportionality constant K_m . The concentration of unrepressed operons is calculated by taking into account the fact that active repressor molecules reversibly bind operons to block transcription initiation, and assuming that this reaction takes place sufficiently rapidly that it is in a quasi-steady state. From this, the concentration of unrepressed operons is $O K_o / (K_o + R_A)$, where K_o is the dissociation constant of the repressor-operon reaction, O is the total operon concentration, and R_A is the concentration of active repressor molecules.

When produced by the *trpR* operon, repressor molecules are unable to repress *trp* operons. For this repression to take place, they need to be activated by two tryptophan molecules which sequentially bind non-cooperatively in two independent places. From this, and assuming that this reaction takes place in a quasi-steady state, Sinha obtained the concentration of active repressor given by $R_A = nTR / (T + K_d)$, where $n = 2$, R is the total repressor concentration, and K_d is the repressor activation dissociation constant.

Compared with the Bliss et al. [14] model, the Sinha model considers the DNA-repressor and repressor-tryptophan interactions in a more detailed way. It also ignores the fact that anthranilate synthase is feedback inhibited by tryptophan as well as ignoring the time delays inherent to the system. These are both important features of the tryptophan operon regulatory system.

It is also important to note that Sinha assumes a constant tryptophan consumption rate G_{\max} . Observing this, Sen and Liu [18] modified the Sinha model to study the case of a non-linear tryptophan consump-

tion rate given by a Michaelis–Menten-like function $G(T) = G_{\max}T/(T + K_g)$. Sinha, as well as Sen and Liu, studied the stability of their models for various values of parameters K_0 and K_2 and found the corresponding stability regions in the (K_0, K_2) parameter space. However, they neither compare their models with experimental data nor do they discuss whether the values for K_0 and K_2 where stability is lost are phys-

ologically attainable in wild type or mutant bacterial strains.

2.3. A more detailed model

Recently Santillán and Mackey [19] introduced a more detailed mathematical model of the tryptophan operon regulatory system, which is shown schematically in Fig. 1. In this model, all of the known reg-

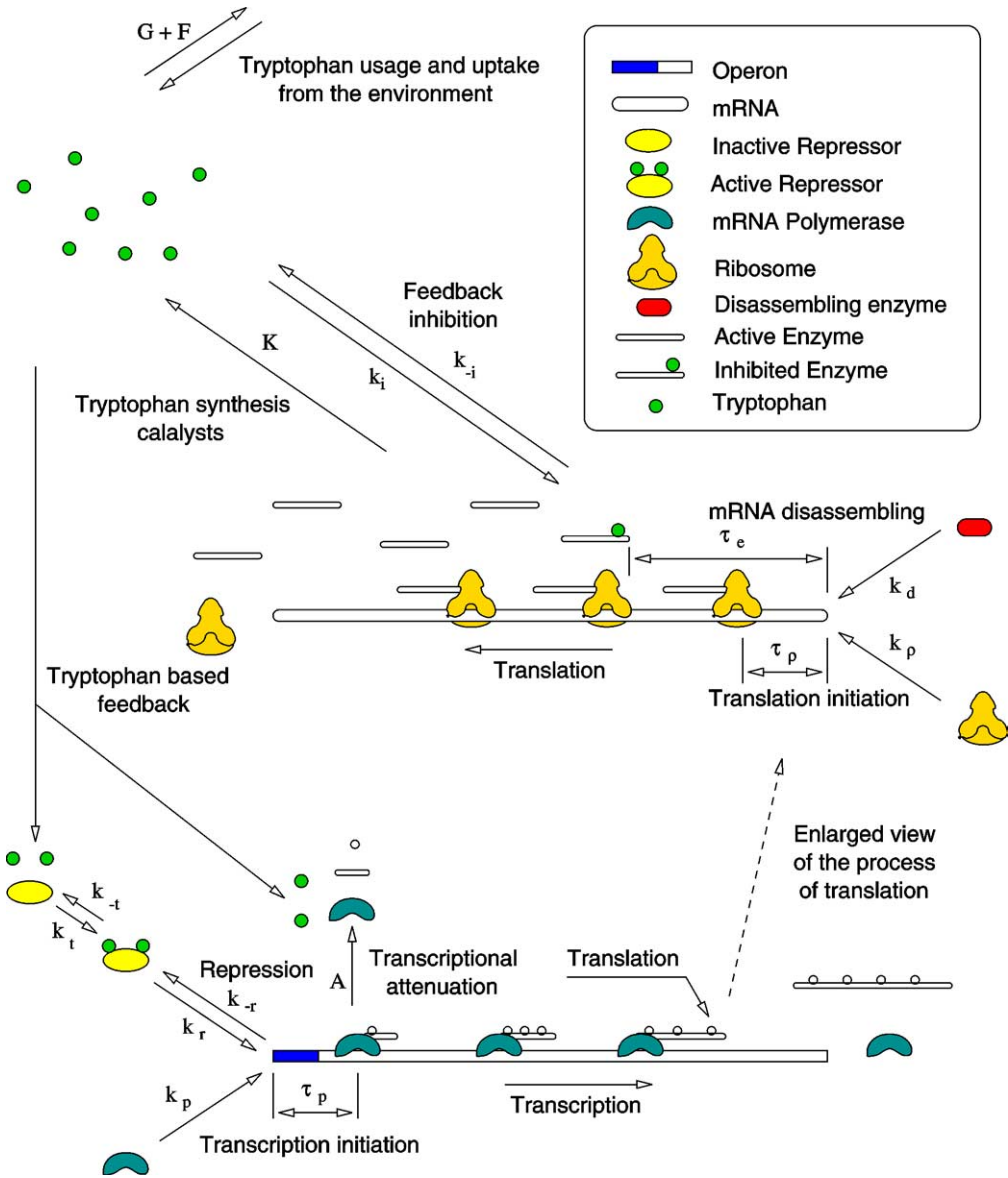


Fig. 1. Schematic representation of the tryptophan operon regulatory system. See the text for a detailed description.

ulatory mechanisms of the tryptophan operon (repression, transcriptional attenuation, and enzyme feedback inhibition) are taken into account. The system inherent time delays due to transcription and translation are also included, as well as the *mRNA* dynamics. The model considers four independent variables: The concentration of *trp* operons with operators free to be bound by *mRNA* polymerase molecules (O_F), the concentration of *trp mRNA* molecules with free *TrpE*-related ribosome binding sites (M_F), the concentration of the enzyme anthranilate synthase (E), and the tryptophan concentration (T).

Although five different polypeptides are synthesized during the expression of the *trp* operon and they combine to form the enzymes that catalyze the reaction pathway which synthesizes tryptophan from chorismic acid, this model concentrates on anthranilate synthase. The reason for this focus is that this enzyme is the most important from a regulatory point of view since it is subject to feedback inhibition by tryptophan and is the first to play its catalytic role in the reaction pathway [13].

Anthranilate synthase is a complex of two *TrpE* and two *TrpD* polypeptides. From this, Santillán and Mackey assume that the anthranilate synthase production rate is one half that of the *TrpE* polypeptide and look only at the *TrpE*-related ribosome binding sites of the *mRNA* chain. The equations governing the dynamics of these variables are:

$$\frac{dO_F}{dt} = \frac{K_r}{K_r + R_A(T)} \{ \mu O - k_p P [O_F(t) - O_F(t - \tau_p) e^{-\mu \tau_p}] \} - \mu O_F \quad (7)$$

$$\frac{dM_F}{dt} = k_p P O_F(t - \tau_m) e^{-\mu \tau_m} [1 - A(T)] - k_p \rho [M_F(t) - M_F(t - \tau_\rho) e^{-\mu \tau_\rho}] - (k_d D + \mu) M_F(t) \quad (8)$$

$$\frac{dE}{dt} = \frac{1}{2} k_p \rho M_F(t - \tau_e) e^{-\mu \tau_e} - (\gamma + \mu) E(t) \quad (9)$$

and

$$\frac{dT}{dt} = K E_A(E, T) - G(T) + F(T, T_{\text{ext}}) - \mu T \quad (10)$$

All the parameters in this model were estimated from reported experimental results, and the details of the parameter estimation process are in [19], which

also contains an extended description of the model derivation.

The factor

$$\frac{K_r}{K_r + R_A(T)}$$

in Eq. (7) is the fraction of unrepressed operons. This follows from the fact that when active repressor reversibly binds free operons, the experimentally reported reaction rates allow one to make a quasi-steady-state assumption. K_r is the dissociation constant of this reaction, and $R_A(T)$ is the active repressor concentration. The experimentally reported rates of the repressor activation reaction also support a quasi-steady-state assumption. Repressor molecules activate when two tryptophan molecules sequentially bind an inactive repressor in two independent sites non-cooperatively. From this

$$R_A(T) = \frac{RT(t)}{T(t) + K_t}$$

where R is the total repressor concentration and K_t is the dissociation constant of the repressor activation constant.

μ is the growth rate of the bacterial culture. The term μO stands for the operon production rate due to *DNA* replication, which was assumed to be such that it keeps the total operon concentration at a constant value O , balancing dilution by growth.

The term

$$-k_p P [O_F(t) - O_F(t - \tau_p) \exp(-\mu \tau_p)]$$

accounts for the rate of free operons binding by *mRNA* polymerase molecules and later freed when the polymerases have traveled for a distance along the operon in a time τ_p . k_p is the reaction rate of the *DNA*-polymerase binding reaction, assumed to be irreversible, and P is the *mRNA* polymerase concentration. The exponential factor $\exp(-\mu \tau_p)$ takes into account the dilution of O_F due to exponential growth during the time τ_p .

In Eq. (8), τ_m is the time it takes for a polymerase to produce a *mRNA* chain long enough to have an available *TrpE*-related ribosome binding site. $A(T)$ is the probability for transcription to be prematurely terminated due to transcriptional attenuation. Santillán and Mackey [19] take $A(T) = b[1 - \exp(-T/c)]$. The

term

$$-k_{\rho}\rho[M_F(t) - M_F(t - \tau_{\rho})\exp(-\mu\tau_{\rho})]$$

in Eq. (8) accounts for the rate of *mRNAs* being bound by ribosomes and liberated after they travel for a time τ_{ρ} along the *mRNA* chain. k_{ρ} is the *mRNA*-ribosome binding reaction rate and ρ is the ribosomal concentration.

The first term on the right-hand side of Eq. (9), is the anthranilate synthase production rate. The factor 1/2 is incorporated because, as mentioned earlier, the enzyme production rate is assumed to be one half that of *TrpE* polypeptide. The *TrpE* production rate equals the rate of ribosome binding free *TrpE*-related ribosome binding sites a time τ_e ago. τ_e is the time it takes for a ribosome to synthesize and release a *TrpE* polypeptide.

The tryptophan production rate, the first term on the right-hand side of Eq. (10), is assumed to be proportional to the concentration of active (uninhibited) enzymes (E_A) with a proportionality constant K . Anthranilate synthase is inhibited when two tryptophan molecules bind the *TrpE* subunits in a sequential cooperative reaction with a Hill coefficient of $n_H \simeq 1.2$. The enzyme feedback inhibition reaction rate constants support a quasi-steady state assumption, from which

$$E_A(E, T) = E(T) \frac{K_i^{n_H}}{T^{n_H} + K_i^{n_H}}$$

K_i is the dissociation constant of this reaction. The tryptophan consumption rate constant is modeled as a Michaelis–Menten-type function

$$G(T) = G_{\max} \frac{T}{T + K_g}$$

Finally, *E. coli* can incorporate tryptophan from the environment. This can be managed by different tryptophan permeases. The rate of tryptophan uptake is modeled by Santillán and Mackey as

$$F(T, T_{\text{ext}}) = d \frac{T_{\text{ext}}}{e + T_{\text{ext}}[1 + T/f]}$$

T_{ext} is the external tryptophan concentration.

All of the model parameters were estimated from the available experimental data. The details are given in Reference [19] supplementary web material. In Table 1 the value of all these parameters is shown.

Table 1

The tryptophan model parameters as estimated in [19]

$R \simeq 0.8 \mu\text{M}$	$O \simeq 3.32 \times 10^{-3} \mu\text{M}$
$P \simeq 2.6 \mu\text{M}$	$k_{-r} \simeq 1.2 \text{ min}^{-1}$
$\rho \simeq 2.9 \mu\text{M}$	$k_r \simeq 460 \mu\text{M}^{-1} \text{ min}^{-1}$
$\tau_p \simeq 0.1 \text{ min}$	$k_{-i} \simeq 720 \text{ min}^{-1}$
$\tau_m \simeq 0.1 \text{ min}$	$k_i \simeq 176 \mu\text{M}^{-1} \text{ min}^{-1}$
$\tau_{\rho} \simeq 0.05 \text{ min}$	$k_{-t} \simeq 2.1 \times 10^4 \text{ min}^{-1}$
$\tau_e \simeq 0.66 \text{ min}$	$k_t \simeq 348 \mu\text{M}^{-1} \text{ min}^{-1}$
$\gamma \simeq 0 \text{ min}^{-1}$	$k_p \simeq 3.9 \mu\text{M}^{-1} \text{ min}^{-1}$
$k_d D \simeq 0.6 \text{ min}^{-1}$	$k_{\rho} \simeq 6.9 \mu\text{M}^{-1} \text{ min}^{-1}$
$n_H \simeq 1.2$	$\mu \simeq 1.0 \times 10^{-2} \text{ min}^{-1}$
$b \simeq 0.85$	$c \simeq 4.0 \times 10^{-2} \mu\text{M}$
$K_g \simeq 0.2 \mu\text{M}$	$g \simeq 25 \mu\text{M} \text{ min}^{-1}$
$e \simeq 0.9 \mu\text{M}$	$d \simeq 23.5 \mu\text{M} \text{ min}^{-1}$
$f \simeq 380 \mu\text{M}$	$K \simeq 126.4 \text{ min}^{-1}$

The model was solved numerically, and the results compared with derepression experiments performed by Yanofsky and Horn [20] on wild-type bacterial cultures as well as *trpL29* and *trpL75* mutant strains of *E. coli*. In these experiments, a bacterial culture is first grown in a medium with a high tryptophan concentration during a period of time long enough for the *trp* operon to reach a steady state. Then the bacteria are washed and shifted to a medium without tryptophan, and the response of the anthranilate synthase activity is measured as a function of time after the nutritional shift.

The *trpL29* mutant strain has a mutation A to G at bp 29 in the leader region of the *trp* operon. This change replaces the leader peptide start codon by GUG , and decreases operon expression in cells growing in the presence or absence of tryptophan. This mutation is simulated by Santillán and Mackey by decreasing the rate constant k_p to 0.004 times its normal value.

The strain *trpL75* of *E. coli* has a mutation of G to A at bp 75 in the leader region of the *trp* operon. This change decreases the stability of the transcription antiterminator structure, and increases transcription termination at the attenuator. Consequently, it decreases operon expression in cells growing in the presence or absence of tryptophan. This mutation was simulated by increasing the value of parameter b , which determines the probability of transcriptional attenuation at high tryptophan levels. The normal value of this parameter is $b \simeq 0.85$, and to simulate the mutation it was increased to $b \simeq 0.9996$.

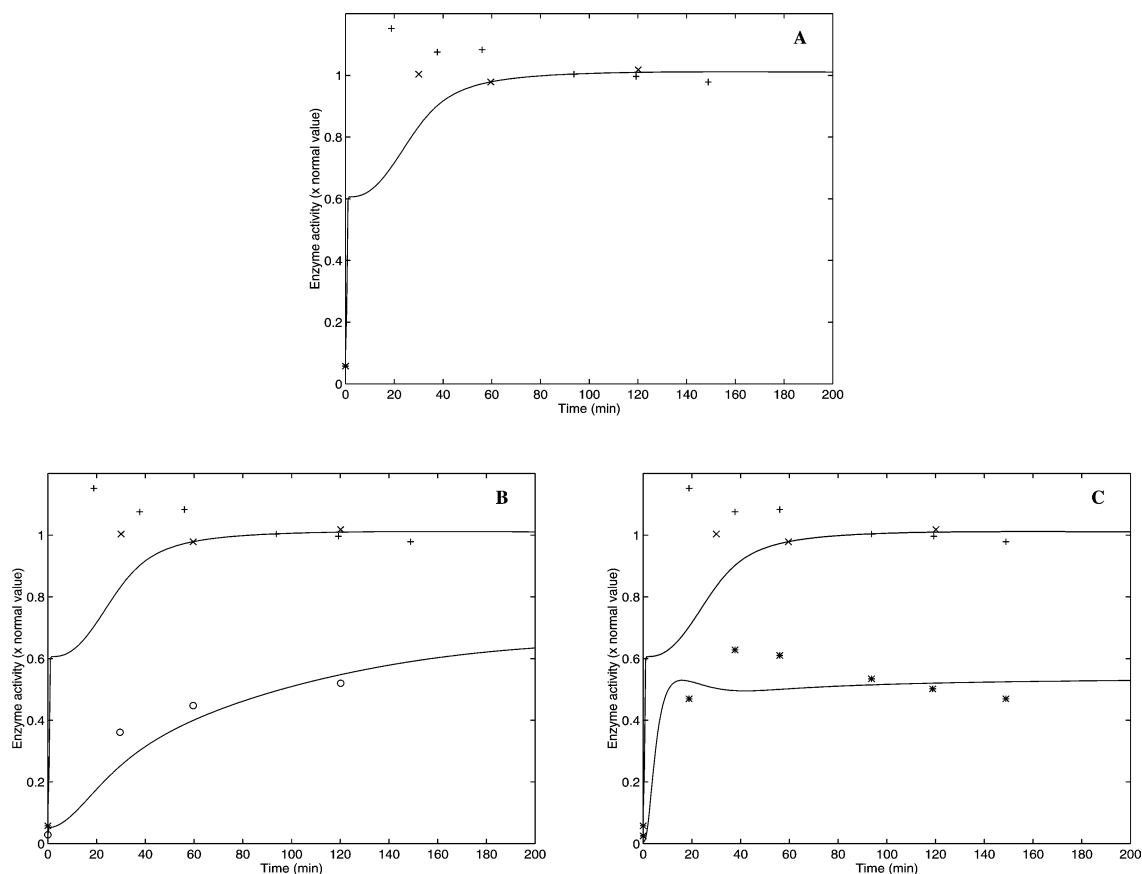


Fig. 2. Experimentally measured points [20] and simulated curves of anthranilate synthase activity vs. time after a nutritional shift (minimal + tryptophan medium to minimal medium), with (A) wild, (B) *trpL29* mutant, and (C) *trpL75* mutant strain cultures of *E. coli*. Two different sets of experimental results (crosses and pluses) for the wild-type strain are presented in A. One set of experimental points corresponding to the *trpL29* mutant strain is shown in B with circles. Another set of experimental points for the *trpL75* mutant strain is shown in C with asterisks. The corresponding simulations are shown with solid lines in all of the three graphics. The wild-type experimental points and simulation are repeated in B and C for comparison.

The experiments of Yanofsky and Horn were simulated numerically. In all of the three cases (wild type and two mutants), the model results show a reasonable qualitative agreement with the experiments, given the simplifying assumptions inherent to the model. The results of these simulations are shown in Fig. 2.

3. The inducible *lac* operon

3.1. Previous models

The *lac* operon is the classical bacterial example of an inducible system which encodes the genes for

the transport of external lactose into the cell and its conversion to glucose and galactose. Many experimental studies have examined the control of the genes involved in lactose metabolism in *E. coli*. One of the most recent mathematical modeling studies of the *lac* operon was carried out by Wong and coworkers [21]. Although many of the relevant biological details were taken into account in their model, they failed to treat the transcriptional and translational delays. Conversely, Mahaffy and Savev [22] modeled *lac* operon dynamics and included transcriptional and translational delays, but ignored the conversion of internal lactose to allolactose by β -galactoside and the spontaneous production rate of mRNA together with

the individual degradation terms for the proteins. In this section we give a brief summary of the work on a recent model developed by Yildirim et al. [23] on the *lac* operon including much of the relevant biological detail considered by Wong and coworkers [21] as well as the transcriptional and translational delays considered by Mahaffy and Savev [22]. A more complete mathematical analysis of the model presented here can be found in [23] and [24].

3.2. The mechanism

The *lac* operon consists of a promoter/operator region and three larger structural genes *lacZ*, *lacY*, and *lacA* together with a preceding regulatory operon responsible for producing a repressor (*R*) protein. To understand the mechanism under this control network easily, it is helpful to refer to Fig. 3. In the absence of glucose available for cellular metabolism, but in the presence of external lactose (L_e), lactose is transported into the cell by a permease (*P*). Intracellular lactose (*L*) is then broken down into allolactose (*A*) first and then glucose and galactose by the enzyme

β -galactosidase (*B*). The allolactose feeds back to bind with the lactose repressor and enable the transcription process to proceed. In the absence of allolactose (*A*) the repressor *R* binds to the operator region and prevents the RNA polymerase from transcribing the structural genes. However, if allolactose is present, a complex is formed between allolactose and the repressor that makes binding of the repressor to the operator region impossible. In that case, the RNA polymerase bound to the promoter is able to initiate transcription of the structural genes to produce mRNA. When a mRNA chain, long enough for a ribosome to bind it, has been produced, the process of translation is initiated. The *lacZ* gene encodes for the mRNA responsible for the production of β -galactosidase (*B*) and transcription of the *lacY* gene produces mRNA ultimately responsible for the production of a membrane permease (*P*). The mRNA produced by transcription of the *lacA* gene encodes for the production of thio-galactoside transacetylase which is thought to not play a role in the regulation of the *lac* operon [25] and will not be further considered here.

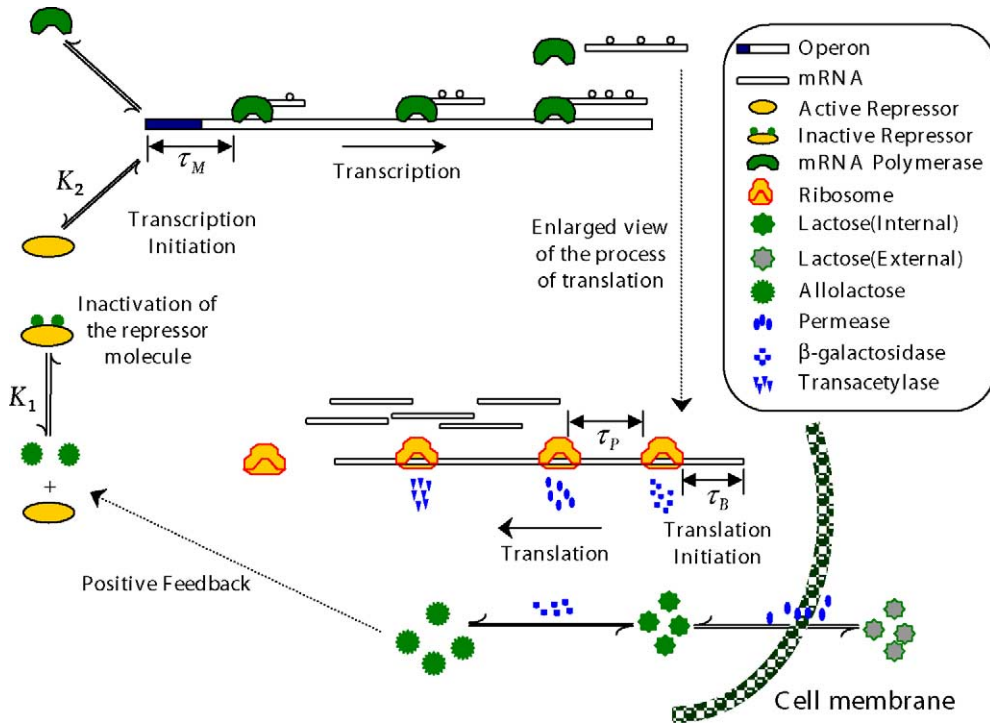


Fig. 3. Schematic representation of the lactose operon regulatory system. See the text for a detailed description.

There is another control mechanism known as catabolite repression. The presence of glucose in the external medium inhibits the production of cAMP in the bacterium. cAMP is necessary for transcription initiation to take place efficiently. In the present work we ignore these regulatory mechanism since we compare with experimental in which there is no glucose in the culture medium and thus, cAMP is always produced at a high rate.

3.3. A new mathematical model

The *lac* operon regulatory pathway is fairly complex: it involves three repressor binding sites (operators), and two CAP binding sites. Moreover, one repressor can bind two different operators simultaneously [26]. A detailed model of all these regulatory elements is beyond the scope of this work. Nonetheless, according to Yagil and Yagil [27], the fraction of free operators as a function of allolactose (A) can be modeled via a Hill-type function of the form:

$$f(A) = \frac{1 + K_1 A^n}{K + K_1 A^n} \quad (11)$$

where n , K and K_1 are parameters to be estimated from experimental data. The dynamics of mRNA production are given by Eq. (12), which is derived as follows. First, note that the production of mRNA from DNA via transcription is not an instantaneous process but requires a period of time τ_M before we have a mRNA long enough to be bound by a ribosome. The rate of change of the ribosome binding sites' concentration (M) is a balance between a production term $\alpha_M f$ and a loss term $(\gamma_M + \mu)M$. The argument of f in the production term is $e^{-\mu\tau_M} A_{\tau_M}$, where $A_{\tau_M} \equiv A(t - \tau_M)$, to account for the time τ_M required to produce the ribosome binding site. The factor $e^{-\mu\tau_M}$ accounts for the growth dependent allolactose dilution during the transcriptional period. The loss term in Eq. (12) is made up of an mRNA degradation term $(\gamma_M M)$ and an effective loss due to dilution (μM) .

$$\frac{dM}{dt} = \alpha_M \frac{1 + K_1 (e^{-\mu\tau_M} A_{\tau_M})^n}{K + K_1 (e^{-\mu\tau_M} A_{\tau_M})^n} - (\gamma_M + \mu)M \quad (12)$$

The dynamics of the concentration (B) of β -galactosidase are described by Eq. (13) in which τ_B is the time required for β -galactosidase production through mRNA translation. We assume that the rate

of production of B is proportional to the concentration of M a time τ_B ago times a dilution factor due to cell growth during the translation process, $(\alpha_B e^{-\mu\tau_B} M_{\tau_B})$. The loss rate of B is given by $(\gamma_B + \mu)B$.

$$\frac{dB}{dt} = \alpha_B e^{-\mu\tau_B} M_{\tau_B} - (\gamma_B + \mu)B \quad (13)$$

For the allolactose concentration (A) dynamics, the first term in Eq. (14) gives the β -galactosidase mediated gain in allolactose from the conversion of lactose. The second term accounts for allolactose loss via conversion to glucose and galactose (again mediated by β -galactosidase). Since they are enzymatic reactions, both of these expressions are taken to be in Michaelis–Menten form. The last term takes into account the degradation and dilution of allolactose.

$$\frac{dA}{dt} = \alpha_A B \frac{L}{K_L + L} - \beta_A B \frac{A}{K_A + A} - (\gamma_A + \mu)A \quad (14)$$

The lactose concentration (L) dynamics are more complicated and given by Eq. (15). The first term in Eq. (15) accounts for the increase of intracellular lactose through the permease facilitated transport of external lactose (L_e). The second term deals with intracellular lactose loss to the extracellular fluid because of the reversible nature of the permease mediated transport [28–31]. The third term accounts for the conversion of lactose to allolactose as well as the hydrolysis of lactose to glucose and galactose via β -galactosidase (B). The fourth term is the decrease in internal lactose concentration due to degradation and dilution.

$$\begin{aligned} \frac{dL}{dt} = & \alpha_L P \frac{L_e}{K_{L_e} + L_e} - \beta_{L_1} P \frac{L}{K_{L_1} + L} \\ & - \alpha_A B \frac{L}{K_L + L} - (\gamma_L + \mu)L \end{aligned} \quad (15)$$

The permease concentration (P) dynamics are given by Eq. (16). The first term reflects the assumption that permease production is directly proportional to the mRNA concentration a time τ_P in the past, where τ_P is the time it takes to translate a permease polypeptide. The exponential factor $e^{-\mu\tau_P}$ accounts for dilution of mRNA concentration due to cell growth. The second term accounts for the degradation and dilution of permease.

$$\frac{dP}{dt} = \alpha_P e^{-\mu\tau_P} M_{\tau_P} - (\gamma_P + \mu)P \quad (16)$$

Table 2
The estimated parameters for the *lac* operon model, taken from [23, 24]

n	2	μ_{\max}	$3.47 \times 10^{-2} \text{ min}^{-1}$
γ_M	0.411 min^{-1}	γ_B	$8.33 \times 10^{-4} \text{ min}^{-1}$
γ_A	0.52 min^{-1}	K	1000
α_M	$9.97 \times 10^{-4} \text{ mM min}^{-1}$	τ_B	2.0 min
α_A	$1.76 \times 10^4 \text{ min}^{-1}$	K_{L_1}	1.81 mM
α_B	$1.66 \times 10^{-2} \text{ min}^{-1}$	K_A	1.95 mM
β_A	$2.15 \times 10^4 \text{ min}^{-1}$	τ_M	0.10 min
K_L	$9.7 \times 10^{-4} \text{ M}$	γ_L	0.0 min^{-1}
γ_P	0.65 min^{-1}	α_L	2880 min^{-1}
α_P	10.0 min^{-1}	τ_P	0.83 min
β_{L_1}	$2.65 \times 10^3 \text{ min}^{-1}$	K_{L_c}	0.26 mM
K_1	$2.52 \times 10^{-2} (\mu\text{M})^{-2}$		

3.4. Analysis of the model and comparison with experimental data

We have carried out an extensive search of the existing literature for data that would allow us to estimate the model parameters in Eqs. (12)–(16) which are summarized in Table 2, and the details of how we arrived at these parameter values can be found in [24].

The model as formulated in Eqs. (12)–(16) can have one, two, or three steady states depending on the values of the parameters (μ, L_e) . The details of how these steady states are determined are contained in [23] and [24]. The results of these considerations are presented in Fig. 4. There we show in (L_e, \bar{A}) space the region where a non-negative steady state can exist. Note in particular that for a range of L_e values there may be coexisting two stable and one unstable steady state values of the intracellular allolactose levels, \bar{A} and, consequently, of $(\bar{M}, \bar{B}, \bar{A}, \bar{L}, \bar{P})$.

In other words, analytical and numerical studies of the model predict that for physiologically realistic values of external lactose and the bacterial growth rate, a regime exists where there may be a bistable steady state behavior, and that this corresponds to a cusp bifurcation in the model dynamics. This means that, as the external lactose concentration L_e is increased, the bacterial culture must switch from a low-internal-allolactose-concentration uninduced steady state to a high-internal-allolactose-concentration induced steady state. For very low L_e values, only the uninduced state exists. Conversely, for very high L_e concentrations, only the induced state is available for the system. All these facts are in agree-

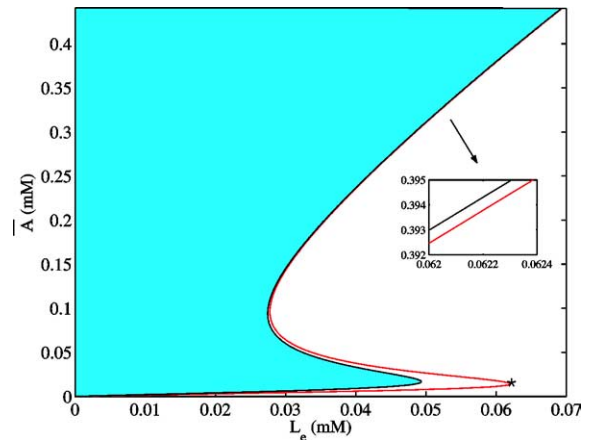


Fig. 4. The region in the (L_e, \bar{A}) space where a non-negative steady state can exist as a function of external lactose levels L_e for the model when all parameters are held at the estimated values in Table 2 and when $\bar{\mu} = 2.26 \times 10^{-2} \text{ min}^{-1}$. The shaded area shows the region where a steady state is not defined, while the solid line is the locus of (L_e, \bar{A}) values satisfying the steady state. The inset box shows that at large values of L_e there is still a separation of the line for the steady state from the region where steady states are not defined. Notice that for these values of the parameters there is a range of L_e values for which there are three coexisting steady-state values of allolactose A . The point marked with a * corresponds to the minimum lactose level for induction at this growth rate, and predicts that induction should occur for a lactose concentration of about 62.0 μM .

ment with the experimentally-observed dynamic behavior of *E. coli* cultures. On the other hand, the model also predicts a range of L_e values in which both the uninduced and the induced stable steady states coexist. If this bistability behavior is real, an hysteresis phenomenon must be observed in inducing and uninducing the *lac* operon of a given bacterial culture by changing the external lactose concentration. This prediction is qualitatively confirmed by the observations in [32] and [33].

Although a full analysis of the stability properties of the model is not possible due to its complexity, we have found that the basic properties contained in a reduced version of this model [23] (which considered only the dynamics of M , B , and A) are apparently retained in this much more complicated model as far as we are able to ascertain analytically and numerically.

Given the parameter values determined in Table 2, we numerically solved the model equations to compare the predicted behavior with three distinct experimental data sets. The first data set is from [34]. In

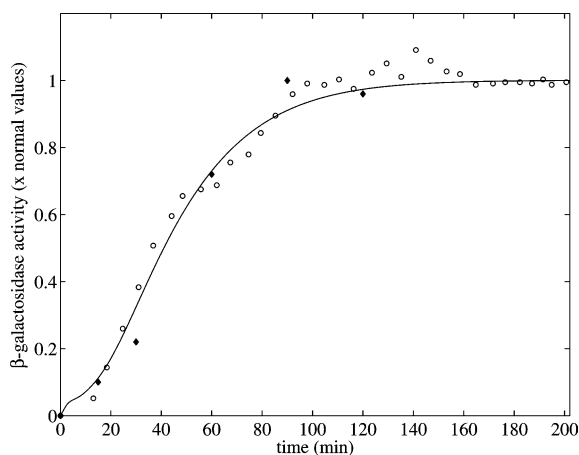


Fig. 5. β -galactosidase activity vs. time when $L_e = 8.0 \times 10^{-2}$ mM. The experimental datasets were taken from [34] for *E. coli* ML30 (o) and from [35] for *E. coli* 294 (◆). The model simulation (solid line) was obtained using the parameters of Table 2 with a growth rate $\bar{\mu} = 2.26 \times 10^{-2} \text{ min}^{-1}$. For details see, [24].

there, the activity of β -galactosidase in an *E. coli* culture (strain ML 30) was measured after the bacteria were switched from a glucose-rich non-lactose to a non-glucose lactose-rich medium. The second data set is from [35]. In this paper, Pestka et al. studied specific inhibition of translation of single mRNA molecules and gave data for the specific activity of β -galactosidase versus time for *E. coli* 294 in the presence of IPTG. These two data sets and the model simulation determined using Matlab's `dde23` [36] routine are shown in Fig. 5.

For this simulation, initial values for the variable were chosen close to the steady state values when $L_e = 8.0 \times 10^{-2}$ mM. (With this value of L_e there is a single uninduced steady state.) To compare these two sets of experimental data with the model simulation predictions, the data were scaled so the steady-state values of measured β -galactosidase activities and those produced by the simulation were equal. As seen in Fig. 5, there is relatively good agreement between both sets of experimental data and the model predicted temporal approach of β -galactosidase activity to its steady-state value.

In the experiments of Pestka et al. [35], the bacterial culture medium has no glucose all of the time. This means that catabolite repression is not a factor in the regulation of the *lac* operon and thus, the comparison with our simulations is straightforward. This is not

the case with the experiments of Knorre [34] since, in such case, the bacterial were originally growing in a medium rich in glucose. Notice however that the time course of both experimental data is quite similar. This does not necessarily justifies the comparison of such experimental data with our model results, but highlights the necessity of including catabolite repression in further mathematical models of the *lac* operon in order to test its influence on the system dynamic behavior.

As a third test of the model, a data set from Goodwin [37] was used. In this paper, the dynamic behavior of β -galactosidase was studied in chemostat cultures of *E. coli* synchronized with respect to cell division by periodic phosphate feeding at a period equal to the bacterial doubling time. Experimentally, oscillations in β -galactosidase concentration were observed with a period equal to the feeding period. Since in these experiments, there is no glucose present in the environment, we can attempt to simulate them with our model.

To mimic the periodic phosphate feeding in our simulation, we assumed that the bacterial growth rate varies as a function of time in manner given by

$$\mu(t) = \bar{\mu} - \alpha \text{mod}(t, T) \quad (17)$$

Here, $\bar{\mu}$ is the maximal growth rate for the bacteria, T is the period of the feeding and α is a positive parameter with dimension min^{-2} . $\text{mod}(t, T)$ is a function that gives the remainder on division of t by T . Selection of this type of function is motivated by the observation that growth rates decrease as nutrient levels fall and sharply increase after the addition of nutrient.

The maximal value of $\mu(t)$ corresponds the times that phosphate was added. The minimal value of this function represents the minimal amount of nutrient left in the vessel. Assuming no nutrient is left ϵ minutes before the addition of phosphate and letting $\epsilon \rightarrow 0$, α can be estimated if the doubling time (T) is known: $\alpha \simeq \bar{\mu}/T$. In Fig. 6, we compare the data on β -galactosidase activity from the forced culture as a function of time with the model predictions for a feeding period of $T = 100$ min which is the population doubling time. Again, there is a reasonable good agreement between the model predicted curve and this dataset.

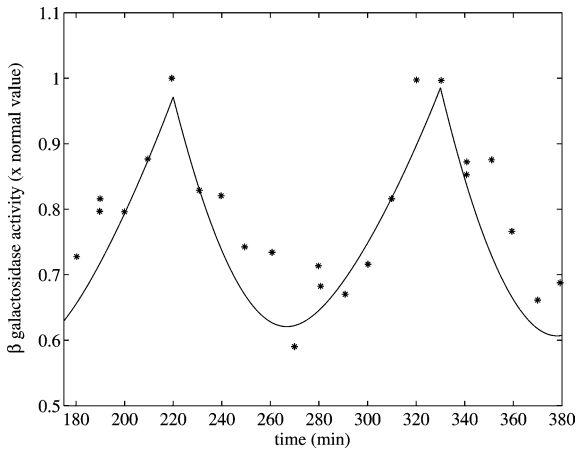


Fig. 6. Oscillation in β -galactosidase activity in response to periodic phosphate feeding with period $T = 100$ min, which is the culture doubling time, and $\bar{\mu} = 2.26 \times 10^{-2} \text{ min}^{-1}$. The experimental data (*) together with the model simulation (solid line) using the parameters of Table 2 are presented. The experimental data are taken from [37]. In the numerical simulation, periodic phosphate feeding was imitated by choosing a periodic function given by Eq. (17). The simulation was calculated by numerically solving the system of delay differential equations given by Eqs. (12)–(16). For details, see [24].

4. Discussion and conclusions

Given the virtual flood of information that is emerging from the current initiatives in molecular biology, we feel that mathematical modeling of the dynamics of molecular control systems will assume an essential role in the coming years. The reasons for this are threefold:

1. Mathematical modeling offers a concise summary of many biological facts and insights in an easily manipulable language.
2. The consequences of analyzing these models can reveal potentially new dynamical behaviors that can be then searched for experimentally.
3. Failures of *realistic* mathematical models to explain experimentally observed behavior often point to the existence of unknown biological details and/or pathways, and thereby can act as a guide to experimentalists in their search for biological reality.

These mathematical models will, however, only be of use to biologists if they are made biologically realis-

tic, and if every effort is made to identify the values of parameters appearing in the models from published data or from experiments explicitly designed to measure these parameters.

The paucity of *dynamic* data (as opposed to steady-state data) with which we were able to compare the predictions of our models for the tryptophan and lactose operons highlights a serious problem in the interaction between experimentalists and mathematical modelers. Namely, we feel that there must be a close and cooperative relationship between these two groups if maximal use is to be made of the experimental data that is being so laboriously collected and if the mathematical models constructed to explain these data.

The last area that we want to touch on is the nature of the mathematical models that are constructed. Those that we have discussed in this paper are all written with the explicit assumption that one is dealing with large numbers cells (and hence of large numbers of molecules) so that the law of large numbers is operative. However, the situation is quite different if one is interested in the dynamics of small numbers (or single) prokaryotic or eukaryotic cells, for then the numbers of molecules are small. Adequate means to analytically treat such problems do not exist in a satisfactory form as of now [38], and one is often reduced to mostly numerical studies [39–41]. The situation is analogous to examining the interactions between small numbers of interacting particles (where the laws of mechanics or electrodynamics hold), and then deriving from these formulations the behavior of large numbers of identical units as is done (not completely successfully even at this point) in statistical mechanics. We view this connection between the ‘micro’ and ‘macro’ levels as one of the major mathematical challenges facing those interested in the understanding of gene control networks.

Acknowledgements

We thank Dr Photini Pitsikas for her initial suggestion of these problems, Prof. Claire Cupples for her advice in parameter estimation section and her general comments, and Mr. Daisuke Horike for many helpful discussions. This work was supported by CGEPI–IPN and COFAA–IPN (México), TUBITAK (Turkey), MITACS (Canada), the Natural Sciences and Engineer-

ing Research Council (NSERC grant OGP-0036920, Canada), the ‘Alexander von Humboldt Stiftung’, the ‘Fonds pour la formation de chercheurs et l’aide à la recherche’ (FCAR grant 98ER1057, Québec), and the Leverhulme Trust (UK).

References

- [1] E. Schrödinger, *What Is Life?*, Cambridge University Press, Cambridge, UK, 1943.
- [2] R.C. Olby, *The Path to the Double Helix*, University of Washington Press, Seattle, WA, 1974.
- [3] J.J. Tyson, M.C. Mackey, Molecular, metabolic and genetic control: an introduction, *Chaos* 11 (2001) 81–83.
- [4] J. Beckwith, The operon: an historical account, in: F.C. Neidhart et al. (Eds.), *Escherichia coli and Salmonella thyphymurium: Cellular and Molecular Biology*, vol. 2, Am. Soc. Microbiology, Washington, DC, 1996, pp. 1553–1569.
- [5] F. Jacob, D. Perrin, C. Sanchez, J. Monod, L’opéron : groupe de gène à expression par un opérateur, *C. R. Acad. Sci. Paris* 250 (1960) 1727–1729.
- [6] B. Goodwin, Oscillatory behaviour in enzymatic control process, *Adv. Enz. Regul.* 3 (1965) 425–438.
- [7] J.S. Griffith, Mathematics of cellular control processes. I. Negative feedback to one gene, *J. Theor. Biol.* 2 (1968) 202–208.
- [8] J.S. Griffith, Mathematics of cellular control processes. II. Positive feedback to one gene, *J. Theor. Biol.* 2 (1968) 209–216.
- [9] J.J. Tyson, H.G. Othmer, The dynamics of feedback control circuits in biochemical pathways, in: R. Rosen (Ed.), in: *Prog. Biophys.*, vol. 5, Academic Press, New York, 1978, pp. 1–62.
- [10] J.F. Selgrade, Mathematical analysis of a cellular control process with positive feedback, *SIAM J. Appl. Math.* 36 (1979) 219–229.
- [11] J.F. Selgrade, A Hopf bifurcation in single loop positive feedback systems, *Q. Appl. Math.* 40 (1982) 347–351.
- [12] J. Ji-Fa, A Liapunov function for four dimensional positive feedback systems, *Q. Appl. Math.* 52 (1994) 601–614.
- [13] C. Yanofsky, I.P. Crawford, The tryptophan operon, in: F.C. Neidhart et al. (Eds.), *Escherichia coli and Salmonella thyphymurium: Cellular and Molecular Biology*, vol. 2, Am. Soc. Microbiology, Washington DC, 1987, pp. 1454–1472.
- [14] R.D. Bliss, R.P. Painter, A.G. Marr, Role of feedback inhibition in stabilizing the classical operon, *J. Theor. Biol.* 97 (1982) 177–193.
- [15] R.D. Bliss, A specific method for determination of free tryptophan and endogenous tryptophan in *Escherichia coli*, *Anal. Biochem.* 93 (1979) 390–398.
- [16] Z.L. Xiu, A.P. Zeng, W.D. Deckwer, Model analysis concerning the effects of growth rate and intracellular tryptophan level on the stability and dynamics of tryptophan biosynthesis in bacteria, *J. Biotechnol.* 58 (1997) 125–140.
- [17] S. Sinha, Theoretical study of tryptophan operon: application in microbial technology, *Biotechnol. Bioeng.* 31 (1988) 117–124.
- [18] A.K. Sen, W. Liu, Dynamic analysis of genetic control and regulation of amino acid synthesis: the tryptophan operon in *Escherichia coli*, *Biotechnol. Bioeng.* 35 (1989) 185–194.
- [19] M. Santillán, M.C. Mackey, Dynamic regulation of the tryptophan operon: a modeling study and comparison with experimental data, *Proc. Natl Acad. Sci. USA* 98 (2001) 1364–1369.
- [20] C. Yanofsky, V. Horn, Role of regulatory features of the *trp* operons of *E. coli* in mediating a response to a nutritional shift, *J. Bacteriol.* 176 (1994) 6245–6254.
- [21] P. Wong, S. Gladney, J.D. Keasling, Mathematical model of the *lac* operon: inducer exclusion, catabolite repression, and diauxic growth on glucose and lactose, *Biotechnol. Prog.* 13 (1997) 132–143.
- [22] J.M. Maffah, E. Simeonov, Stability analysis for a mathematical model of the *lac* operon, *Q. Appl. Math.* 57 (1999) 37–53.
- [23] N. Yildirim, M. Santillán, D. Horike, M.C. Mackey, Dynamics and bistability in a reduced model of the lactose operon: comparison with experimental data, *Chaos* (2004) in press.
- [24] N. Yildirim, M.C. Mackey, Feedback regulation in the lactose operon: a mathematical modeling study and comparison with experimental data, *Biophys. J.* 84 (2003) 2841–2851.
- [25] J. Beckwith, The lactose operon, in: F.C. Neidhart et al. (Eds.), *Escherichia coli and Salmonella thyphymurium: Cellular and Molecular Biology*, vol. 2, Am. Soc. Microbiology, Washington, DC, 1996.
- [26] W.S. Reznikoff, The lactose operon-controlling elements: a complex paradigm, *Mol. Microbiol.* 6 (1992) 2419–2422.
- [27] G. Yagil, E. Yagil, On the relation between effector concentration and the rate of induced enzyme synthesis, *Biophys. J.* 11 (1971) 1–27.
- [28] M.H. Saier, Inducer exclusion and regulation of the melibiose, maltose, glycerol, and lactose transport systems by the phosphoenolpyruvate: sugar phosphotransferase system, *J. Biol. Chem.* 251 (1976) 6606–6615.
- [29] T. Osumi, M.H. Saier, Regulation of lactose permease activity by the phosphoenolpyruvate:sugar phosphotransferase system: evidence for direct binding of the glucose specific enzyme III to the lactose permease, *Proc. Natl Acad. Sci. USA* 79 (1982) 1457–1461.
- [30] P.W. Postma, J.W. Lengeler, G.R. Jacobson, Phosphoenolpyruvate-carbohydrate phosphotransferase systems, in: F.C. Neidhart et al. (Eds.), *Escherichia coli and Salmonella thyphymurium: Cellular and Molecular Biology*, vol. 2, Am. Soc. Microbiology, Washington, DC, 1996.
- [31] M.H. Saier, T.M. Ramseier, J. Reizer, Regulation of carbon utilization, in: F.C. Neidhart et al. (Eds.), *Escherichia coli and Salmonella thyphymurium: Cellular and Molecular Biology*, vol. 2, Am. Soc. Microbiology, Washington, DC, 1996.
- [32] A. Novick, M. Wiener, Enzyme induction as an all-or-none phenomenon, *Proc. Natl Acad. Sci. USA* 43 (1957) 553–566.
- [33] M. Cohn, K. Horibata, Inhibition by glucose of the induced synthesis of the β -galactosidase-enzyme system of *Escherichia coli*: analysis of maintenance, *J. Bacteriol.* 78 (1959) 613–623.

- [34] W.A. Knorre, Oscillation of the rate of β -galactosidase in *Escherichia coli* ML 30 and ML 308, *Biochim. Biophys. Res. Commun.* 31 (1968) 812–817.
- [35] S. Pestka, B.L. Daugherty, V. Jung, K. Hotta, R.K. Pestka, Anti-mRNA: specific inhibition of translation of single mRNA molecules, *Proc. Natl Acad. Sci. USA* 81 (1984) 7525–7528.
- [36] L.F. Shampine, S. Thompson, Solving DDEs with MATLAB, www.radford.edu/~thompson/webdeds/.
- [37] B.C. Goodwin, Control dynamics of β -galactosidase in relation to the bacterial cell cycle, *Eur. J. Biochem.* 10 (1969) 515–522.
- [38] T.B. Kepler, T.C. Elston, Stochasticity in transcriptional regulation: origins, consequences, and mathematical representations, *Biophys. J.* 81 (2001) 3116–3136.
- [39] D.T. Gillespie, Exact stochastic simulation of coupled chemical reactions, *J. Phys. Chem.* 81 (1977) 2340–2361.
- [40] H.H. McAdams, L. Shapiro, Circuit simulation of genetic networks, *Science* 269 (1995) 650–656.
- [41] A. Arkin, J. Ross, H.H. McAdams, Stochastic kinetic analysis of developmental pathway bifurcation in phage λ -infected *Escherichia coli* cells, *Genetics* 149 (1998) 1633–1648.