Review

The utility of simple mathematical models in understanding gene regulatory dynamics

Michael C. Mackey[a,*], Moisés Santillán[b], Marta Tyran-Kamińska[c] and Eduardo S. Zeron[d]
[a]*Departments of Physiology, Physics & Mathematics, McGill University, Montreal, Quebec, Canada*
[b]*Centro de Investigación y de Estudios Avanzados del IPN, Unidad Monterrey, Parque
de Investigación e Innovación Tecnológica, Apodaca NL, México*
[c]*Institute of Mathematics, University of Silesia, Bankowa, Katowice, Poland*
[d]*Departamento de Matemáticas, Centro de Investigación y de Estudios Avanzados del IPN,
Apartado Postal, México DF, México*

Abstract. In this review, we survey work that has been carried out in the attempts of biomathematicians to understand the dynamic behaviour of simple bacterial operons starting with the initial work of the 1960's. We concentrate on the simplest of situations, discussing both repressible and inducible systems and then turning to concrete examples related to the biology of the lactose and tryptophan operons. We conclude with a brief discussion of the role of both extrinsic noise and so-called intrinsic noise in the form of translational and/or transcriptional bursting.

1. Introduction

The operon concept for the regulation of bacterial genes, first put forward in [1], has had an astonishing and revolutionary effect on the development of understanding in molecular biology. It is a testimony to the strength of the theoretical and mathematical biology community that modeling efforts aimed at clarifying the implications of the operon concept appeared so rapidly after the concept was embraced by biologists. Thus, to the best of our knowledge the first analysis of operon dynamics appeared in [2] and in [3]. These first attempts were swiftly followed by Griffith's analysis of a simple repressible operon [4] and an inducible operon [5], and these and other results were beautifully summarized in [6].

Since these modeling efforts in the early days of development in molecular biology, both our biological knowledge and level of sophistication in modeling

---

*Corresponding author: Michael C. Mackey, Departments of Physiology, Physics & Mathematics, McGill University, 3655 Promenade Sir William Osler, Montreal, Quebec H3G 1Y6, Canada. E-mail: michael.mackey@mcgill.ca.

have proceeded apace to the point where new knowledge of the biology is actually driving the development of new mathematics. This is an extremely exciting development and one which many have expected–that biology would act as a driver for mathematics in the 21st century much as physics was the driver for mathematics in the 19th and 20th centuries. However, as this explosion of biological knowledge has proceeded hand in hand with the development of mathematical modeling efforts to understand and explain it, the difficulty in comprehending the nature of the field becomes ever deeper due to the sheer volume of work being published.

In this highly idiosyncratic review we discuss work from our group over the past few years directed at the understanding of *really simple* operon control dynamics. We start this review in Section 2 by discussing transcription and translation kinetics (Section 2.1) and then pass to general dynamics considerations in Section 2.2 which is largely a recap of earlier work with additional insights derived from the field of nonlinear dynamics. We then pass to the role of transcriptional and translational delays in Section 2.3 and finish with a short consideration of fast and slow variables in Section 2.4. Following this, in Sections 3.1 and 3.2 we pass from the realm of mathematical nicety to biological reality by looking at *realistic* models for the lactose and tryptophan operons respectively. These two examples, two of the most extensively experimentally studied systems in molecular biology, and for which we have vast amounts of data, illustrate the reality of dealing with experimental biology and the difficulties of applying realistic modeling efforts to understand that biology.

Finally, in Section 4 we turn to one of the more interesting and challenging aspects of understanding operon dynamics. In the last few years with the advent of ever improved imaging techniques combined with rapid data acquisition techniques experimentalists have acquired the ability to peer ever more closely into the details of these dynamics at virtually the single molecule level. This means, therefore, that all manner of interesting statistical behaviours are being uncovered–behaviours that reveal many interesting types of 'random' behaviour not well understood from a mathematical perspective. We explore aspects of this in Section 4.1 where we consider the effects of transcriptional and/or translational bursting, and in Section 4.2 where we look at the effects of fluctuations in degradation rates. The review ends with a brief discussion in Section 5.

## 2. Generic deterministic models of prokaryotic gene regulation

The central tenet of molecular biology was put forward some half century ago, and though modified in detail still stands in its basic form. Transcription of DNA produces messenger RNA (mRNA, denoted M here). Then through the process of translation of mRNA, intermediate protein (I) is produced which is capable of controlling metabolite (E) levels that in turn can feedback and affect transcription and/or translation. A typical example would be in the lactose operon of Section 3.1 where the intermediate is β-galactosidase and the metabolite is allolactose. These metabolites are often referred to as effectors, and their effects can, in the simplest case, be either stimulatory (so called inducible) or inhibitory (or repressible) to the entire process. This scheme is often called the 'operon concept'.

2.1. Kinetic aspects of regulation of transcription and translation

We first outline the relatively simple molecular dynamics of both inducible and repressible operons and how effector concentrations can modify transcription rates. If transcription rates are constant and unaffected by any effector, then this is called a 'no control' situation.

2.1.1. Inducible regulation

The *lac* operon considered below in Section 3.1 is the paradigmatic example of inducible regulation. In an inducible operon when the effector (E) is present then the repressor (R) is inactive and unable to bind to the operator (O) region so DNA transcription can proceed unhindered. E binds to the active form R of the repressor and we assume that this binding reaction is

$$R + nE \underset{k_1^-}{\overset{k_1^+}{\rightleftharpoons}} RE_n,$$
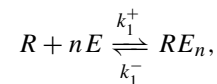
in which k_1^+ and k_1^- are the forward and backward reaction rate constant, respectively. The equilibrium equation for the reaction above is

$$K_1 = \frac{RE_n}{R \cdot E^n}, \qquad (2.1)$$

where $K_1 = k_1^- / k_1^+$ is the reaction dissociation constant and n is the number of effector molecules required

to inactivate repressor R. The operator O and repressor R are also assumed to interact according to
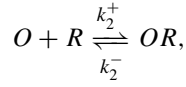
$$O + R \underset{k_2^-}{\overset{k_2^+}{\rightleftharpoons}} OR,$$

which has the following equilibrium equation:

$$K_2 = \frac{OR}{O \cdot R}, \quad K_2 = \frac{k_2^-}{k_2^+}.$$

The total operator O_{tot} is given by

$$O_{tot} = O + OR = O + K_2 O \cdot R = O(1 + K_2 R),$$

while the total repressor is R_{tot}

$$R_{tot} = R + K_1 R \cdot E^n + K_2 O \cdot R.$$

Furthermore, by definition the fraction of operators free to synthesize mRNA (*i.e.*, not bound by repressor) is

$$f(E) = \frac{O}{O_{tot}} = \frac{1}{1 + K_2 R}.$$

If the amount of repressor R bound to the operator O is small

$$R_{tot} \simeq R + K_1 R \cdot E^n = R(1 + K_1 E^n)$$

so

$$R = \frac{R_{tot}}{1 + K_1 E^n},$$

and consequently

$$f(E) = \frac{1 + K_1 E^n}{1 + K_2 R_{tot} + K_1 E^n} = \frac{1 + K_1 E^n}{K + K_1 E^n}, \quad (2.2)$$

where $K = 1 + K_2 R_{tot}$. Maximal repression occurs when $E = 0$ and even at that point mRNA is produced (so-called leakage) at a basal level proportional to K^{-1}.

Assume that the maximal transcription rate of DNA (in units of time^{-1}) is $\bar{\varphi}_m$. Assume further that transcription rate φ in the entire population is proportional to the fraction of unbound operators f. Thus we expect that φ as a function of the effector level will be given by $\varphi = \bar{\varphi}_m f$, or

$$\varphi(E) = \bar{\varphi}_m \frac{1 + K_1 E^n}{K + K_1 E^n}. \quad (2.3)$$

2.1.2. Repressible regulation

The tryptophan operon considered below in Section 3.2 is the classic example of a repressible system. This is because the repressor is active (capable of binding to the operator) when the effector molecules are present which means that DNA transcription is blocked. Using the same notation as before, but realizing that the effector binds the inactive form R of the repressor so it becomes active and take this reaction to be the same as in Equation 2.1. However, we now assume that the operator O and repressor R interaction is governed by
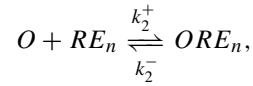
$$O + RE_n \underset{k_2^-}{\overset{k_2^+}{\rightleftharpoons}} ORE_n,$$

with the following equilibrium equation

$$K_2 = \frac{ORE_n}{O \cdot RE_n}, \quad K_2 = \frac{k_2^-}{k_2^+}. \quad (2.4)$$

The total operator is

$$O_{tot} = O + ORE_n = O + K_1 K_2 O \cdot R \cdot E^n$$
$$= O(1 + K_1 K_2 R \cdot E^n),$$

so the fraction of operators not bound by repressor is

$$f(E) = \frac{O}{O_{tot}} = \frac{1}{1 + K_1 K_2 R \cdot E^n}.$$

Assuming, as before, that the amount of R bound to O is small compared to the amount or repressor gives

$$f(E) = \frac{1 + K_1 E^n}{1 + (K_1 + K_1 K_2 R_{tot}) E^n} = \frac{1 + K_1 E^n}{1 + K E^n},$$

where $K = K_1(1 + K_2 R_{tot})$. In this case we have maximal repression when E is large, and even when repression is maximal there is still be a basal level of mRNA production (again known as leakage) which is proportional to $K_1 K^{-1} < 1$. Variation of the DNA transcription rate with effector level is given by $\varphi = \bar{\varphi}_m f$ or

$$\varphi(E) = \bar{\varphi}_m \frac{1 + K_1 E^n}{1 + K E^n}. \quad (2.5)$$

Both (2.3) and (2.5) are special cases of

$$\varphi(E) = \bar{\varphi}_m \frac{1 + K_1 E^n}{A + B E^n} = \bar{\varphi}_m f(E). \quad (2.6)$$

The constants A, $B \geq 0$ are defined in Table 2.1.

Table 2.1
The parameters A, B, Λ, Δ and θ for the inducible and repressible cases. See the text and Section 2.1 for more detail

Parameter	Inducible	Repressible
A	$K = 1 + K_2 R_{tot}$	1
B	K_1	$K = K_1(1 + K_2 R_{tot})$
$\frac{B}{A}$	$\frac{K_1}{K}$	K
$\Lambda = A$	K	1
$\Delta = BK_1^{-1}$	1	KK_1^{-1}
$\theta = \frac{\kappa_d}{n\Delta}\left(1 - \frac{A}{\Lambda}\right)$	$\frac{\kappa_d}{n} \cdot \frac{K-1}{K} > 0$	$\frac{\kappa_d}{n} \cdot \frac{K_1-K}{K} < 0$

2.2. General dynamic considerations

The Goodwin model for operon dynamics [2] considers a large population of cells, each of which contains one copy of a particular operon, and we use that as a basis for discussion. We let (M, I, E) respectively denote the mRNA, intermediate protein, and effector concentrations. For a generic operon with a maximal level of transcription $\bar{b}_d$ (in concentration units), the dynamics are given by [2, 4, 5, 7, 8]

$$\frac{dM}{dt} = \bar{b}_d \bar{\varphi}_m f(E) - \gamma_M M, \qquad (2.7)$$

$$\frac{dI}{dt} = \beta_I M - \gamma_I I, \qquad (2.8)$$

$$\frac{dE}{dt} = \beta_E I - \gamma_E E. \qquad (2.9)$$

It is assumed here that the rate of mRNA production is proportional to the fraction of time the operator region is active, and that the rates of protein and metabolite production are proportional to the amount of mRNA and intermediate protein respectively. All three of the components (M, I, E) are subject to random degradation, and the function f is as determined in Section 2.1 above.

To simplify things we formulate Equations (2.7)–(2.9) using dimensionless concentrations. To start we rewrite Equation (2.6) in the form

$$\varphi(e) = \varphi_m f(e),$$

where φ_m (which is dimensionless) is defined by

$$\varphi_m = \frac{\bar{\varphi}_m \beta_E \beta_I}{\gamma_M \gamma_E \gamma_I} \quad \text{and} \quad f(e) = \frac{1 + e^n}{\Lambda + \Delta e^n}, \quad (2.10)$$

Λ and Δ are defined in Table 2.1, and a (dimensionless) effector concentration (e) is defined by

$$E = \eta e \quad \text{with} \quad \eta = \frac{1}{\sqrt[n]{K_1}}.$$

We continue and define dimensionless intermediate protein (i) and mRNA concentrations (m):

$$I = i\eta \frac{\gamma_E}{\beta_E} \quad \text{and} \quad M = m\eta \frac{\gamma_E \gamma_I}{\beta_E \beta_I},$$

so Equations (2.7)–(2.9) take the form

$$\frac{dm}{dt} = \gamma_M[\kappa_d f(e) - m],$$

$$\frac{di}{dt} = \gamma_I(m - i),$$

$$\frac{de}{dt} = \gamma_E(i - e),$$

with the dimensionless constants

$$\kappa_d = b_d \varphi_m \quad \text{and} \quad b_d = \frac{\bar{b}_d}{\eta}.$$

To continue our simplifications we rename the dimensionless concentrations through $(m, i, e) = (x_1, x_2, x_3)$, and subscripts $(M, I, E) = (1, 2, 3)$ to finally obtain

$$\frac{dx_1}{dt} = \gamma_1[\kappa_d f(x_3) - x_1], \qquad (2.11)$$

$$\frac{dx_2}{dt} = \gamma_2(x_1 - x_2), \qquad (2.12)$$

$$\frac{dx_3}{dt} = \gamma_3(x_2 - x_3). \qquad (2.13)$$

In all of these equations, γ_i for $i = 1, 2, 3$ denotes a degradation rate (units of inverse time), and thus Equations (2.11)–(2.13) are *not* in dimensionless form. The dynamics of this classic operon model have been fully analyzed [9], the results of which we simply summarize here. We set $X = (x_1, x_2, x_3)$ and let $S_t(X)$ be the flow generated by the system (2.11)–(2.13), i.e., the function $t \mapsto S_t(X)$ is a solution of (2.11)–(2.13) such that $S_0(X) = X$. For both inducible and repressible operons, for all initial conditions $X^0 = (x_1^0, x_2^0, x_3^0) \in \mathbb{R}_3^+$ the flow $S_t(X^0) \in \mathbb{R}_3^+$ for $t > 0$.

The steady state solutions of (2.11)–(2.13) are given by the solutions of

$$\frac{x}{\kappa_d} = f(x) \qquad (2.14)$$

and for each solution x^* of Equation (2.14) there is a steady state $X^* = (x_1^*, x_2^*, x_3^*)$ of (2.11)–(2.13) which is given by

$$x_1^* = x_2^* = x_3^* = x^*.$$

Whether there is a single steady state X^* or there are multiple steady states will depend on whether we are considering a repressible or inducible operon.

2.2.1. No control

No control simply means $f(x) \equiv 1$, and in this case there is a single steady state $x^* = \kappa_d$ that is globally asymptotically stable.

2.2.2. Inducible regulation

Single versus multiple steady states. For an inducible operon (with f given by Equation (2.2)) there may be one (X_1^* or X_3^*), two (X_1^*, $X_2^* = X_3^*$ or $X_1^* = X_2^*$, X_3^*), or three (X_1^*, X_2^*, X_3^*) steady states, with the ordering $0 < X_1^* \leq X_2^* \leq X_3^*$, corresponding to the possible solutions of Equation (2.14) (cf. Figure 2.1). The smallest steady state (X_1^*) is typically called the un-induced state, while the largest steady state (X_3^*) corresponds to the induced state. The steady state values of x are easily obtained from (2.14) for given parameter values, and the dependence on κ_d for $n = 4$ and a variety of values of K is shown in Fig. 2.1. Figure 2.2 shows a graph of the steady states x^* versus κ_d for various values of the leakage parameter K.
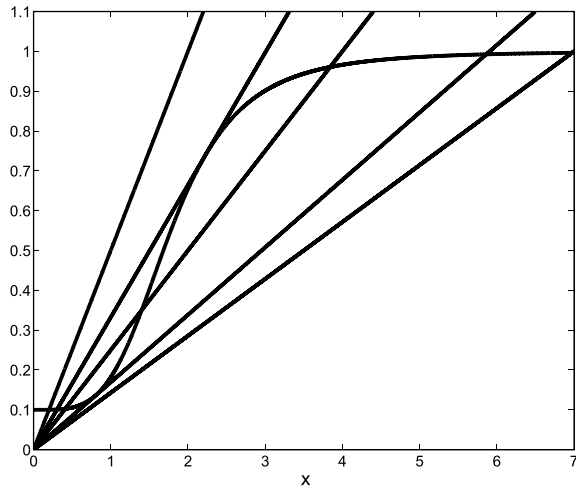

Fig. 2.1. A schematic illustration of the possibility of one, two or three solutions of Equation (2.14) for varying values of κ_d in the presence of inducible regulation. The monotone increasing graph is f of Equation (2.10), and the straight lines correspond to x/κ_d for (in a clockwise direction) $\kappa_d \in [0, \kappa_{d-})$, $\kappa_d = \kappa_{d-}$, $\kappa_d \in (\kappa_{d-}, \kappa_{d+})$, $\kappa_d = \kappa_{d+}$, and $\kappa_{d+} < \kappa_d$. This figure was constructed with $n = 4$ and $K = 10$ for which $\kappa_{d-} = 3.01$ and $\kappa_{d+} = 5.91$ as computed from (2.17). See the text for details. Taken from [9] with permission.
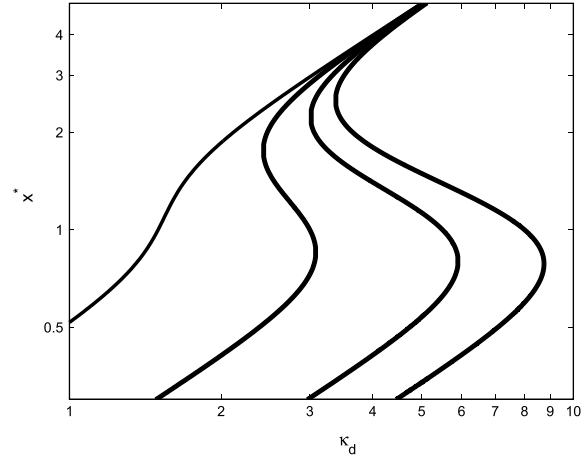

Fig. 2.2. Logarithmic plot of the steady state values of x^* versus κ_d for an inducible operon obtained from Equation (2.14), for $n = 4$ and $K = 2, 5, 10$, and 15 (left to right) illustrating the dependence of the occurrence of bistability on K. See the text for details. Taken from [9] with permission.

Analytic conditions for the existence of one or more steady states come from Equation (2.14) in conjunction with the observation that the delineation points are marked by the values of κ_d at which x/κ_d is tangent to $f(x)$ (see Figure 2.1). Differentiation of (2.14) yields a second condition

$$\frac{1}{\kappa_d n(K-1)} = \frac{x^{n-1}}{(K+x^n)^2}. \qquad (2.15)$$

From Equations (2.14) and (2.15) the values of x at which tangency will occur are given by:

$$x_\pm$$

$$= \sqrt[n]{\frac{K-1}{2}\left\{\left[n - \frac{K+1}{K-1}\right] \pm \sqrt{n^2 - 2n\frac{K+1}{K-1} + 1}\right\}}. \qquad (2.16)$$

The corresponding values of κ_d at which a tangency occurs are given by

$$\kappa_{d\pm} = x_\mp \frac{K + x_\mp^n}{1 + x_\mp^n}. \qquad (2.17)$$

A necessary condition for the existence of two or more steady states is obtained by requiring that the radical in (2.16) is non-negative:
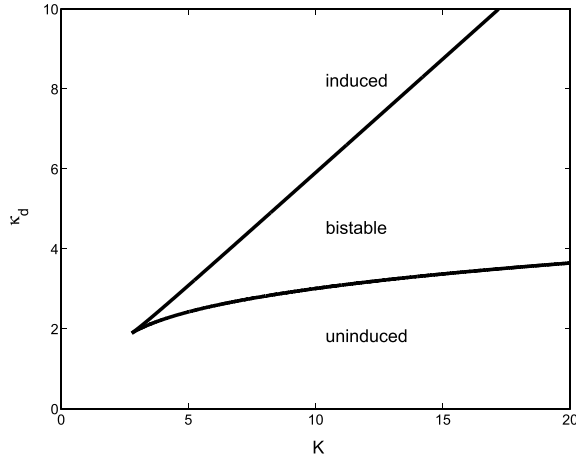
$$K \geq \left(\frac{n+1}{n-1}\right)^2. \qquad (2.18)$$

Fig. 2.3. This figure presents a parametric plot (for $n = 4$) of the bifurcation diagram in (K, κ_d) parameter space separating one from three steady states in an inducible operon as determined from Equations (2.14) and (2.15). The upper (lower) branch corresponds to κ_{d-} (κ_{d+}), and for all values of (K, κ_d) in the interior of the cone there are two locally stable steady states X_1^*, X_3^*, while outside there is only one. The tip of the cone occurs at $(K, \kappa_d) = ((5/3)^2, (5/3)\sqrt[4]{5/3})$ as given by Equations (2.18) and (2.19). For $K \in [0, (5/3)^2)$ there is a single steady state. Taken from [9] with permission.

Thus a second necessary condition follows:

$$\kappa_d \geq \frac{n+1}{n-1} \sqrt[n]{\frac{n+1}{n-1}}. \tag{2.19}$$

Further, from Equations (2.14) and (2.15) we can find the boundaries in (K, κ_d) space in which there are one or three locally stable steady states as shown in Fig. 2.3. There, we have given a parametric plot (x is the parameter) of κ_d versus K, using

$$K(x) = \frac{x^n[x^n + (n+1)]}{(n-1)x^n - 1} \quad \text{and}$$

$$\kappa_d(x) = \frac{[K(x) + x^n]^2}{nx^{n-1}[K(x) - 1]},$$

for $n = 4$ obtained from Equations (2.14) and (2.15). As is clear from the figure, when leakage is appreciable (small K, e.g for $n = 4$, $K < (5/3)^2$) then the possibility of bistable behaviour is lost.

We can make some general comments on the influence of n, K, and κ_d on the appearance of bistability from this analysis. First, the degree of cooperativity (n) in the binding of effector to the repressor plays a significant role and $n > 1$ is a necessary condition for bistability. If $n > 1$ then a second necessary condition for bistability is that K satisfies Equation (2.18)

so the fractional leakage (K^{-1}) is sufficiently small. Furthermore, κ_d must satisfy Equation (2.19) which is quite interesting. For $n \to \infty$ the limiting lower limit is $\kappa_d > 1$ while for $n \to 1$ the minimal value of κ_d becomes quite large. This simply tells us that the ratio of the product of the production rates to the product of the degradation rates must always be greater than 1 for bistability to occur, and the lower the degree of cooperativity (n) the larger the ratio must be.

If n, K and κ_d satisfy these necessary conditions then bistability is only possible if $\kappa_d \in [\kappa_{d-}, \kappa_{d+}]$ (c.f. Figure 2.3). The locations of the minimal (x_-) and maximal (x_+) values of x bounding the bistable region are independent of κ_d. And, finally, ($x_+ - x_-$) is a decreasing function of increasing n for constant κ_d, K while ($x_+ - x_-$) is an increasing function of increasing K for constant n, κ_d.

Local and global stability. Although the local stability analysis of the inducible operon is possible [9], the thing that is interesting is that the global stability is possible to determine.

Theorem 2.1. *[7, 10, Proposition 2.1, Chapter 4] For an inducible operon with φ given by Equation (2.3), define $I_I = [1/K, 1]$. There is an attracting box $B_I \subset \mathbb{R}_3^+$ defined by*

$$B_I = \{(x_1, x_2, x_3) : x_i \in I_I, \ i = 1, 2, 3\}$$

such that the flow S_t is directed inward everywhere on the surface of B_I. Furthermore, all $X^* \in B_I$ and

1. *If there is a single steady state, i.e. X_1^* for $\kappa_d \in [0, \kappa_{d-})$, or X_3^* for $\kappa_{d+} < \kappa_d$, then it is globally stable.*
2. *If there are two locally stable nodes, i.e. X_1^* and X_3^* for $\kappa_d \in (\kappa_{d-}, \kappa_{d+})$, then all flows $S_t(X^0)$ are attracted to one of them. (See [8] for a delineation of the basin of attraction of X_1^* and X_3^*.)*

2.2.3. Repressible regulation

As is clear from a simple consideration of our dynamical equations the repressible operon has a single steady state corresponding to the unique solution x^* of Equation (2.14). Again, rather remarkably, we can characterize the global stability of this single steady state through the following.

Theorem 2.2. *[10, Theorems 4.1 & 4.2, Chapter 3] For a repressible operon with φ given by Equation (2.5),*

define $I_R = [K_1/K, 1]$. *There is a globally attracting box $B_R \subset \mathbb{R}_3^+$ defined by*

$$B_R = \{(x_1, x_2, x_3) : x_i \in I_R, \ i = 1, 2, 3\}$$

such that the flow S_t is directed inward everywhere on the surface of B_R. Furthermore there is a single steady state $X^* \in B_R$. If X^* is locally stable it is globally stable, but if X^* is unstable then a generalization of the Poincare-Bendixson theorem [10, Chapter 3] implies the existence of a globally stable limit cycle in B_R.

2.3. The appearance of cell growth effects and delays due to transcription and translation

The considerations of the previous sections must, however, be tempered by the realization that sometimes cell growth has to be taken into account as well as the fact that significant delays may enter into the dynamical equations [11]. The effects of growth are obvious in that if a cell increases its volume then there is an effect on concentrations. But where do these delays come from? Their origin is simple to understand and arises from the fact that the transcription and translation processes take place at a finite velocity and therefore require a non-zero time for completion. The existence of these delays has been known for some time by modelers [12] and whether the incorporation of the delays will potentially change the qualitative nature of the model dynamics will depend on the type of regulation. Generally when the regulation is that of an inducible operon there will be no change, but if the system is a repressible one then the inclusion of the transcriptional and translational delays may lead to the prediction of limit cycle behaviour.

Once we take growth and these transcriptional and translational delays into account, our basic dynamical equations are modified to the form

$$\frac{dM}{dt} = \bar{b}_d \bar{\varphi}_m f(e^{-\mu \tau_M} E_{\tau_M}) - \bar{\gamma}_M M, \qquad (2.20)$$

$$\frac{dI}{dt} = \beta_I e^{-\mu \tau_I} M_{\tau_I} - \bar{\gamma}_I I, \qquad (2.21)$$

$$\frac{dE}{dt} = \beta_E I - \bar{\gamma}_E E. \qquad (2.22)$$

In Equations (2.20)–(2.22) there are several changes to be noted. The first is the appearance of the terms $e^{-\mu \tau_M}$ and $e^{-\mu \tau_I}$ which account for an effective dilution of the mRNA (M) and intermediate protein (I) because the cell is growing at a rate μ (time^{-1}). The second

is the alteration of the decay rates γ_i to $\bar{\gamma}_i \equiv \gamma_i + \mu$ because the cell growth leads to an effective increase in the rate of destruction. The last is the altered notation $E_{\tau_M}(t) = E(t - \tau_M)$ and $M_{\tau_I}(t) = M(t - \tau_I)$ indicating that both E and M are now to be evaluated at a time in the past due to the non-zero times required for transcription and translation. From a dynamic point of view, the presence of these delays can have a dramatic effect.

Equations (2.20)–(2.22) can be put in a simpler form, just as we did for (2.7)–(2.9), but by now setting

$$E = \eta e \quad \text{with} \quad \eta = \frac{1}{e^{-\mu \tau_M} \sqrt[n]{K_1}},$$

$$I = i\eta \frac{\gamma_E}{\beta_E} \quad \text{and} \quad M = m\eta \frac{\gamma_E \gamma_I}{\beta_E \beta_I} e^{-\mu \tau_I}$$

so Equations (2.20)–(2.22) take the form

$$\frac{dm}{dt} = \bar{\gamma}_M [\kappa_d f(e_{\tau_M}) - m],$$

$$\frac{di}{dt} = \bar{\gamma}_I (m_{\tau_I} - i),$$

$$\frac{de}{dt} = \bar{\gamma}_E (i - e),$$

with

$$\kappa_d = \frac{\bar{b}_d \varphi_m \beta_I \beta_E}{\bar{\gamma}_M \bar{\gamma}_I \bar{\gamma}_E \eta} e^{-\mu \tau_I}. \qquad (2.23)$$

To finish our simplifications, as before rename the dimensionless concentrations $(m, i, e) = (x_1, x_2, x_3)$, and subscripts $(M, I, E) = (1, 2, 3)$ to obtain

$$\frac{dx_1}{dt} = \bar{\gamma}_1 [\kappa_d f(x_{3, \tau_1}) - x_1], \qquad (2.24)$$

$$\frac{dx_2}{dt} = \bar{\gamma}_2 (x_{1, \tau_2} - x_2), \qquad (2.25)$$

$$\frac{dx_3}{dt} = \bar{\gamma}_3 (x_2 - x_3). \qquad (2.26)$$

Again Equations (2.24)–(2.296) are *not* in dimensionless form.

It is important to realize that the appearance of the delays τ_M and τ_I (or τ_1 and τ_2) plays absolutely no role in the determination of the steady state(s) of inducible and repressible systems as discussed above.

For an inducible operon in which $f'(X^*) > 0$ a simple extension of the proof in [10, Proposition 6.1, Chapter 6] shows that the global stability properties are not altered by the presence of the delays (τ_1, τ_2).

However, for a repressible operon there are, at this point in time, no extensions of the global stability results of [10, Theorem 4.1 & Theorem 4.2, Chapter 3] for inducible systems. The best that we can do is to linearize Equations (2.24)–(2.26) in the neighborhood of the unique steady state X^* to obtain the eigenvalue equation $g(\lambda) = P(\lambda) + \vartheta e^{-\lambda\tau} = 0$ wherein

$$P(\lambda) = (\bar{\gamma}_1 + \lambda)(\bar{\gamma}_2 + \lambda)(\bar{\gamma}_3 + \lambda) \quad \text{and}$$
$$\vartheta = -\kappa_d f'(X^*)\bar{\gamma}_1\bar{\gamma}_2\bar{\gamma}_3 > 0 \qquad (2.27)$$

and $\tau = \tau_1 + \tau_2$. Writing out $g(\lambda)$ we have

$$g(\lambda) = \lambda^3 + a_1\lambda^2 + a_2\lambda + a_3 + \vartheta e^{-\lambda\tau}, \quad (2.28)$$

where

$$a_1 = \sum_{i=1}^{3} \gamma_i, \quad a_2 = \sum_{i \neq j=1}^{3} \gamma_i\gamma_j, \quad a_3 = \prod_{i=1}^{3} \gamma_i.$$

Let $\lambda(\tau) = \alpha(\tau) + i\omega(\tau)$ be the root of Equation (2.28) satisfying $\alpha(\tau_0) = 0$ and $\omega(\tau_0) = \omega_0$, and set $p = a_1^2 - 2a_2$, $q = a_2^2 - 2a_1a_3$, $r = a_3^2 - \vartheta^2$, and let $h(z) = z^3 + pz^2 + qz + r$. [13, Theorem 2.4] gives the conditions for X^* to be locally stable and for the existence of a Hopf bifurcation.

Theorem 2.3. *[13, Theorem 2.4]*

1. *If $r \geq 0$ and $\Delta = p^2 - 3q < 0$, then all roots of Equation (2.28) have negative real parts for all $\tau \geq 0$.*
2. *If $r < 0$ or $r \geq 0$, $z_1 > 0$ and $h(z_1) < 0$, then all roots of Equation (2.28) have negative real parts when $\tau \in [0, \tau_0)$.*
3. *If the conditions of (2) are satisfied, $\tau = \tau_0$ and $h'(\omega_0^2) \neq 0$, then $\pm i\omega_0$ is a pair of simple purely imaginary roots of Equation (2.28) and all other roots have negative real parts. Moreover,*

$$\frac{dRe\lambda(\tau_0)}{d\tau} > 0.$$

2.4. Fast and slow variables

Identifying fast and slow variables can give considerable simplification and insight into the long term behaviour of the system. A fast variable in a given dynamical system relaxes much more rapidly to an equilibrium than a slow one [14]. Differences in degradation rates in chemical and biochemical systems lead to the distinction that the slowest variable is the one that has the smallest degradation rate.

Typically the degradation rate of mRNA is much greater than the corresponding degradation rates for both the intermediate protein and the effector ($\gamma_1 \gg \gamma_2, \gamma_3$) so in this case the mRNA dynamics are fast and we have the approximate relationship

$$x_1 \simeq \kappa_d f(x_3).$$

If $\gamma_1 \gg \gamma_2 \gg \gamma_3$ so that the effector is the slowest variable, then we have

$$x_2 \simeq x_3$$

and the three variable system (2.11)–(2.13) describing the generic operon reduces to a one dimensional system

$$\frac{dx_3}{dt} = \gamma_3[\kappa_d f(x_3) - x_3] \qquad (2.29)$$

for the relatively slow effector dynamics. If instead the effector qualifies as a fast variable (as for the *lac* operon) so that $\gamma_1 \gg \gamma_3 \gg \gamma_2$ and

$$x_3 \simeq x_2$$

then the intermediate protein is the slowest variable described by the one-dimensional equation

$$\frac{dx_2}{dt} = \gamma_2[\kappa_d f(x_2) - x_2]. \qquad (2.30)$$

Consequently both Equations (2.30) and (2.29) are of the form

$$\frac{dx}{dt} = \gamma[\kappa_d f(x) - x], \qquad (2.31)$$

where γ is either γ_2 for protein (x_2) dominated dynamics or γ_3 for effector (x_3) dominated dynamics.

Eliminating fast variables, also known as the adiabatic elimination technique [14], has been extended to stochastically perturbed systems when the perturbation is a Gaussian distributed white noise, c.f. [15, Section 11.1], [16, 17], and [18, Section 6.4]. For the case of perturbation being a jump Markov process we refer to [19].

3. Specific examples in various systems

3.1. The lactose operon

Glucose is the favourite carbon and energy source for *E. coli*, as well as for many other organisms. Although this bacterium can also feed on other sugars, it only does so when glucose is absent. A typical
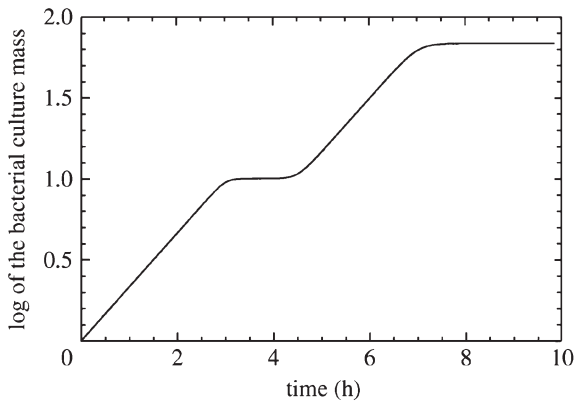
Fig. 3.1. Typical diauxic growth curve. Note the existence of two different exponential growth phases, separated by a short interval in which the culture does not grow. The first (second) phase corresponds to the bacterial culture feeding on glucose (lactose), while the interval with no growth corresponds to the time required for the bacteria to turn on the genes needed to metabolize lactose after glucose exhaustion.

population of *E. coli* doubles its size approximately every hour in presence of a pure sugar, like glucose or lactose. The existence of the lactose operon was conjectured by Jacob and Monod after observing that a population of *E. coli* is initially unable to digest lactose, when it is fed with a mixture of the glucose and lactose sugars.

Monod [20] observed in his PhD work that in the presence of a mixture of glucose and lactose the exponential growth begins as usual, then it pauses for about one hour before resuming at a similar pace. The bacterial growth curve shows two distinctive phases, as can be seen in Fig. 3.1. The key observation was that the timing of the pauses was controlled by the ratio of the initial amounts of glucose and lactose: the larger initial amount of glucose the later the pause would begin. Monod realized that *E. coli* is initially unable to digest lactose, so that the bacteria initially feeds exclusively on glucose, until it is totally consumed and the bacteria then needs to change its internal metabolism to consume lactose. It is worth mentioning at this point that diauxic growth only occurs in batch cultures, and simultaneous usage of sugars is often observed in continuous cultures [21].

Jacob and co-workers [1] proposed the lactose operon model as a mechanism for explaining these features. Thus, the *lac* genes that encode the enzymes necessary for lactose absorption and hydrolysis are all controlled by a single mechanism, and they are all turned off in the presence of glucose or the absence

of lactose. Properly speaking, the lactose operon is a DNA segment composed of a promoter/operator region, followed by the structural genes *lacZ*, *lacY*, and *lacA*, and finally by the corresponding terminator. The promoter/operator is the DNA region where the *transcription factors* (RNA polymerase, lactose repressor, cyclic-AMP receptor protein, *et cetera*) bind in order to initiate the transcription of a corresponding mRNA strand or to regulate the corresponding transcription process. The gene *lacZ* codes for the enzyme β-galactosidase (LacZ) that in *E. coli* cleaves the disaccharide lactose into glucose and galactose. The gene *lacY* codes for the enzyme β-galactoside permease (LacY), an inner membrane-bound symporter that pumps lactose into the cell using a proton gradient. Finally, *lacA* encodes the enzyme β-galactoside transacetylase (LacA) that transfers an acetyl group from acetyl-sides. Nevertheless, it is not completely understood what its precise function is.

The β-galactosidase enzyme. Few genes have a history of study as long and distinguished as *lacZ*. The *lacZ* gene encodes an open reading frame of 1024 amino acids and was one of the first large genes to be completely sequenced. In *E. coli*, the biologically active β-galactosidase protein exists as a tetramer of four identical subunits and has a molecular weight of approximately 480–500 kDa. The primary enzymatic function of β-galactosidase relevant to its role as a biotechnological tool is to cleave the chemical bond between the anomeric carbon and glycosyl oxygen of appropriate substrates; see for example [22].

lacZ was chosen as the target of a very extensive early analysis, in part owing to specific experimental advantages accompanying work with β-galactosidase. These advantages continue to provide a rationale for using this protein in biotechnological applications today.

The β-galactoside permease protein. Active transporters (pumps) require a cellular energy source (*i.e.* ATP hydrolysis) to catalyze the transport of charged components against an electrochemical gradient. Depending on their energy source, active transporters are classified as primary or secondary. In particular, secondary transporters use the free energy stored in a given electrochemical ion gradient, as shown in [23]. β-galactoside permease is a secondary transporter that couples free energy released from downhill translocation of protons to drive the uphill transloca-

tion of galactosides against a concentration gradient. This protein is composed of 417 amino acid residues and has 12 helices that transverse the membrane in a zigzag fashion, connected by relatively hydrophilic loops with both N and C termini on the cytoplasm side. β-galactoside permease is encoded by the *lacY* gene, the second structural gene in the lactose operon. *lacY* was the first gene encoding a membrane transport protein to be cloned into a recombinant plasmid, overexpressed and sequenced; see for example [24] and the references therein. This success in the early days of molecular biology opened the study of secondary active transport at the molecular level. Thus, β-galactoside permease was the first protein of its class to be solubilized and purified in a completely functional state, thereby demonstrating that this single gene product is solely responsible for all the translocation reactions catalyzed by the galactoside transport system in *E. coli*. [24] has also shown that this protein is both structurally and functionally a monomer in the membrane.

The lactose operon regulatory pathway. The lactose operon plays two main important roles in the *E. coli* metabolism: It controls the production of the enzymes necessary for lactose absorption and hydrolysis, but it also closes a positive feedback loop, the so called *lactose operon regulatory pathway*. Once the disaccharide lactose is pumped inside the bacteria by the β-galactoside (*lac*) permease, the second enzyme β-galactosidase has the dual role of transforming the lactose into allolactose and hydrolyzing both (lactose and allolactose) into the monosaccharides galactose and glucose. The positive feedback loop is closed when the intermediary sugar allolactose interacts with the control mechanisms of the lactose operon. Thus the allolactose binds to the lactose repressor *lacI* reducing its ability to repress the transcription and expression of the structural genes *lacZ*, *lacY*, and *lacA*. We refer the reader to the cartoon in Fig. 3.5. for a better understanding. Consequently an increment in the concentration of lactose or allolactose inside the bacteria enhances the production of the enzymes β-galactosidase and β-galactoside permease, via the expression of the structural genes *lacZ* and *lacY*. This incremental enzyme production enhances the absorption of more external lactose and its transformation into allolactose, closing the feedback loop.

In summary, the lactose operon is an excellent example of the inducible operon reviewed in Section 2.2. However, it took a while to interpret the lactose operon subtle behaviour in terms of what we now call bistability. This interpretation was first introduced by Novick and Wiener [25] and Cohn and Horibata [26], who suggested that a single cell may have two alternative states: induced, in which it can metabolize lactose, or uninduced, in which the corresponding genes are switched off and lactose metabolism does not occur. From their results, Novick and Wiener, as well as Cohn and Horibata, interpreted the so called *maintenance effect* as the consequence of a high permease concentration in induced cells, which would enable these cells to maintain the induced state and to transmit it to their progeny, even if placed in a medium with a low concentration of inducer. Although this interpretation accounts for the existence of two distinct phenotypes and provides an explanation of why induced cells placed in media with low inducer concentrations remain indefinitely induced, whereas cells that have never been induced stayed uninduced, it does not explain what makes the cells switch between alternative states. This switching remained a mystery that had to wait for the introduction of the concept of multistability to be fully explained.

We have seen in Section 2.2 that Griffith [5] introduced a mathematical model for a single gene controlled by a positive feedback loop, and found that, under certain conditions, two stable states may be accessible for the system simultaneously. However, Griffith did not use his model to explain the maintenance effect of the *lac* operon. The first models explicitly aimed at unraveling this phenomenon were due to Babloyantz and Sanglier [27] and to Nicolis and Prigogine [28], who were able to interpret the maintenance effect as the biological facet of the physical process of multistability. These models were quite complex, and took into account all the information regarding the lactose operon regulatory pathway available at the time. However, the level of detailed knowledge about the underlying molecular mechanisms has expanded greatly in the intervening decades. Thus, more detailed and sophisticated models are possible. Below, we review some of the most recent modeling studies of the lactose operon, many of which are by our group.

Transcription of the structural genes. Let $P(O_P)$ be the probability that a polymerase is bound to the promoter/operator region of the lactose operon and it is ready to initiate transcription. The dynamical equations for the *lacZ* and *lacY* ribosome binding sites (RBSs) in the mRNA molecule are given [30–32] by
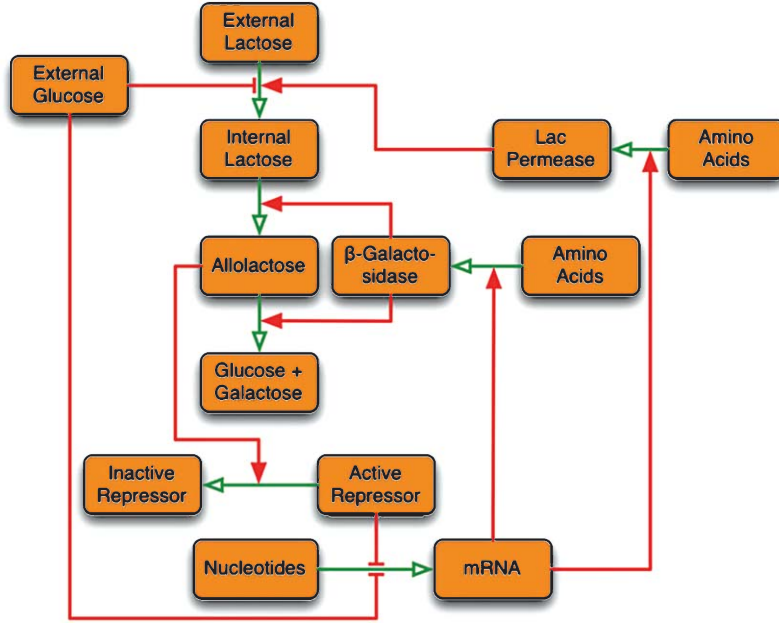
Fig. 3.2. Cartoon representation of the lactose operon regulatory pathway. Labeled rectangles represent chemical species, arrows with empty heads denote processes through which one chemical species is transformed into another, arrows with solid heads indicate interactions that enhance the process they point to, and finally, lines ending in perpendicular bars denote interactions that diminish (inhibit) the process they point to. Taken from [29] with permission.

$$\frac{dM_Z}{dt} = Dk_M P_{\tau_Z}(O_P) - (\gamma_M + \mu)M_Z, \quad (3.1)$$

$$\frac{dM_Y}{dt} = Dk_M P_{\tau_Y}(O_P) - (\gamma_M + \mu)M_Y. \quad (3.2)$$

Variable M_Z and M_Y respectively denote the concentrations of *lacZ* and *lacY* RBSs. D stands for the concentration (number of molecules per average bacteria) of lactose promoters, k_M is the maximum transcription initiation rate of the promoter, γ_M denotes the mRNA degradation rate, and μ is the average bacterias grown rate. μ is included along with the degradation rate γ_M to account for the effective loss due to dilution. Both (3.1) and (3.2) share the same parameters because the structural genes *lacZ* and *lacY* are located in tandem after the promoter, and thus they are transcribed by the same polymerase one after the other. Finally, the notation $P_{\tau_Z}(*)(t)$ stands for $P(*)(t-\tau_Z)$, and we use it to take into account the time delay τ_Z existing between transcription initiation and translation initiation. Hence, τ_Z is the time interval between transcription initiation and the moment when the corresponding RBS is transcribed, so that a ribosome can bind to it and initiate the translation. Obviously, the time delay τ_Y is larger than τ_Z, because the structural

genes *lacZ* are located close to the promoter and so are transcribed first. Note that the symmetry between Equations (3.1) and (3.2) implies that $M_Y(t)$ is equal to $M_Z(t-\tau)$ for the difference $\tau = \tau_Y - \tau_Z$, so that we need to use only one of these equations.

Translation of mRNA. The dynamical equations for the concentration of the proteins encoded by the genes *lacZ* and *lacY* are given [30–32] by

$$\frac{dE_Z}{dt} = k_Z e^{-\mu\tau_Z^*} M_{Z,\tau_Z^*} - (\gamma_Z + \mu)E_Z, \quad (3.3)$$

$$\frac{dE_Y}{dt} = k_Y e^{-\mu\tau_Y^*} M_{Y,\tau_Y^*} - (\gamma_Y + \mu)E_Y. \quad (3.4)$$

The variable E_Z (E_Y) denotes the concentration of LacZ (LacY) polypeptides. The parameter k_Z stands for the maximum translation initiation rate at the *lacZ* RBS, τ_Z^* is the time necessary to fully translate a LacZ polypeptide, γ_Z denotes the protein E_Z degradation rate, and μ is as before. The exponential factor $e^{-\mu\tau_Z^*}$ accounts for dilution of mRNA concentration due to cell growth in the time interval $[t-\tau_Z^*, t]$. Finally, the notation $M_{Z,\tau_Z^*}(t)$ stands for the delayed function $M_Z(t-\tau_Z^*)$. The parameters k_Y, τ_Y^*, and γ_Y in Equation (3.4) have the same meaning as above for the

dynamics of protein E_Y. Since the *lacY* mRNA segment has its own ribosome binding site, it is translated independently from *lacZ* mRNA segment.

Observe that if the set of parameters $(k_Z, \tau_Z^*, \gamma_Z)$ is identically equal to $(k_Y, \tau_Y^*, \gamma_Y)$, then the symmetry between Equations (3.3) and (3.4) implies that $E_Y(t)$ is equal to $E_Z(t-\tau)$ for $\tau = \tau_Y - \tau_Z$, because we already know that M_Y is equal to $M_{Z,\tau}$.

Lactose absorption and hydrolysis into lactose and allolactose. Once the *lacZ* and *lacY* polypeptides are produced, they pass through several biochemical processes like folding and tetramerization in order to produce the corresponding enzymes β-galactosidase and β-galactoside permease. The internal dynamics of these biochemical processes are not modeled in general (the corresponding reversible reactions are assumed to always be in equilibrium), and so one may take

$$B = E_Z/4 \quad \text{and} \quad Q = E_Y, \qquad (3.5)$$

where B is the internal concentration of β-galactosidase and Q denotes the concentration of β-galactoside permease. The factor $1/4$ comes from the fact that β-galactosidase is a homo-tetramer made up of four identical *lacZ* polypeptides. We thus assume that all the β-galactosidase monomers are incorporated into tetramers.

Dynamical equations for the concentration of intracellular lactose L in bacteria were developed in [30] and [31], and then later improved [32] to include explicitly the effects of the external glucose G_e in the absorption of lactose. This latter formulation took the form

$$\frac{dL}{dt} = k_L \beta_L(L_e) \beta_G(G_e) Q - k_\ell \beta_\ell(L) Q$$
$$- \phi_M \mathcal{M}(L) B - (\gamma_L + \mu) L. \qquad (3.6)$$

L, as before, is the concentration of intracellular lactose, while G_E (L_E) denotes the concentration of extracellular glucose (lactose). The first term $k_L \beta_L \beta_G Q$ in (3.6) stands for the gain of intracellular lactose L obtained due to the action of the β-galactoside permease Q in the transport of extracellular lactose L; the second term $k_\ell \beta_\ell(L) Q$ expresses the loss of intracellular lactose to the extracellular fluid due to the reversible nature of the permease mediated transport; the third term $\phi_M \mathcal{M}(L) B$ accounts for the β-galactosidase mediated conversion of lactose into allolactose as well as the hydrolysis of lactose to glu-

cose and galactose. The last term in (3.6) stands for the decrease in internal lactose due to degradation and dilution. $\beta_L(L_e)$ is an increasing function of the extracellular lactose L_e, and $\beta_G(G_e)$ is decreasing with respect to the external glucose G_e to take into account the negative influence of the glucose on the absorption of lactose:

$$\beta_L(L_e) = \frac{L_e}{\kappa_L + L_e} \quad \text{and} \quad \beta_G(G_e) = 1 - \frac{\phi_G \, G_e}{\kappa_G + G_e}. \tag{3.7}$$

Furthermore, the terms $\beta_\ell(L)$ and $\mathcal{M}(L)$ are both functions of the internal lactose

$$\beta_\ell(L) = \frac{L}{\kappa_\ell + L} \quad \text{and} \quad \mathcal{M}(L) = \frac{L}{\kappa_M + L}. \quad (3.8)$$

The dynamical equation for the concentration of allolactose A is much simpler:

$$\frac{dA}{dt} = \alpha \phi_M \mathcal{M}(L) B - \phi_A \mathcal{M}(A) B - (\gamma_A + \mu) A, \tag{3.9}$$

where α is the fraction of internal lactose L transformed by β-galactosidase B into allolactose instead of being hydrolyzed into glucose and galactose. The term $\phi_A \mathcal{M}(A) B$ represents the hydrolysis of allolactose into glucose and galactose mediated by β-galactosidase, while the last term in (3.9) stands for the decrease in internal allolactose due to degradation and dilution. We implicitly assume that the dynamics of lactose and allolactose hydrolysis are so similar that the same functions $\mathcal{M}(L)$ and $\mathcal{M}(A)$ can be used to represent both.

In particular, if $\alpha \phi_M \simeq \phi_A$ holds, $\gamma_A + \mu$ is close to zero, and the allolactose dynamics are fast (so that Equation (3.9) is always close to equilibrium), then we conclude that $A \approx L$ and is independent of B.

The lactose operon control system. The system of Equations (3.1) to (3.9) gives a mathematical model of the biochemical reactions involved in the transcription and translation of the *lac* structural genes, the absorption of the extracellular lactose, its later transformation into allolactose, and the hydrolysis of lactose and allolactose into glucose and galactose. The one thing left to specify is an exact expression for the probability $P(O_P)$ that a polymerase is bound to the promoter/operator region of the lactose operon and it is ready to initiate transcription. We need an explicit formula for $P(O_P)$ in order to substitute it into Equations (3.1)–(3.2) and to model how allolactose and glucose control the production of the enzymes nec-
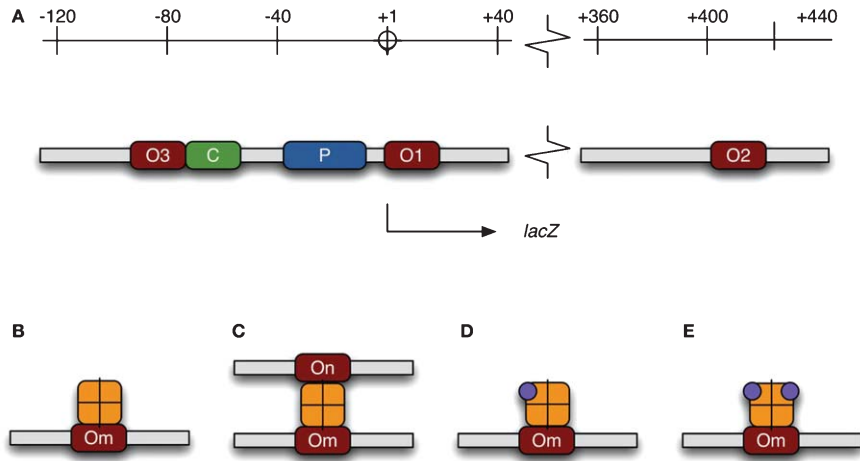
Fig. 3.3. (A) Schematic of the regulatory elements located in lactose operon DNA. P denotes the promoter, O1, O2, and O3 correspond to the three operators (repressor-binding sites), and C is the binding site for the cAMP–CRP complex. The different ways in which a repressor molecule can interact with the operator sites are represented in panels B to E. Namely, a free repressor molecule (B), one with a single subunit bound by allolactose (D) or one with the two subunits in the same side bound by allolactose (E) can bind a single operator. Moreover, a free repressor molecule can bind two different operators simultaneously (C). Taken from [29] with permission.

essary for the lactose absorption, transformation, and hydrolysis, closing in this way the positive feedback loop described previously.

The system (3.1) to (3.9) was presented in [30–32] and has not been significantly modified since the time it was originally developed. However, the probability $P(O_P)$ has changed significantly from the original form

$$P(O_P) = \frac{a + A^n}{b + A^n}$$

proposed by [30].

Other investigators [29, 33–35] have proposed different formulas for $P(O_P)$ adding more and more new details on the lactose operon control system, which is quite complex as the most recent discoveries show. Thus [36] and [37] have established that the lactose operon regulatory elements (pictured in Fig. 3.3a) are distributed along the DNA chain as follows: the lactose promoter is located between bp -36 (bp stands for base pair, and positions are referred relative to the starting point of gene *lacZ*, bp $+1$) and bp -7. Operator O1 is 21 bp long and centred around bp $+11$. There are two additional operators, denoted O2 and O3, which are, respectively, located at 401 bp downstream and 92 bp upstream from O1. Finally, the activator (CAP)-binding site spans from bp -72 to bp -50.

The lactose repressor is a homo-tetramer (consisting of two functional homo-dimers) of *lacI* polypeptides,

according to [38] and [39]. Each functional dimer can bind operators O1, O2 and O3. Furthermore, DNA can also fold in such a way that a single repressor binds two operators simultaneously, one per dimer. Each monomer in the lactose repressor can be bound by an allolactose molecule, inhibiting the capability of the corresponding dimer to bind an operator. This means that free repressors can bind one operator (Fig. 3.3b) or two of them simultaneously (Fig. 3.3c), repressors with three free monomers can bind one but not two operators (Fig. 3.3d), repressors with two free monomers can bind one operator, if the bound monomers belong to the same dimer (Fig. 3.3e), or none at all, and that repressors with only one free monomer are unable to bind any operator, as are repressors with all four monomers bound by allolactose; see for example [40].

Deletion experiments [41] have shown that a repressor bound to O1 inhibits transcription initiation, while a repressor bound to either O2 or O3 has almost no effect on the expression of the lactose operon structural genes. Nevertheless, O2 and O3 do have an indirect effect because the complex formed by a single repressor simultaneously bound to O1 and either O2 or O3 is far more stable than that of a repressor bound only to O1. The consequence of this is that interacting with the lactose repressor operator O1 is only capable of decreasing the expression of the operon genes 18 times; when it cooperates with O2, the repression level can be as high as 700-fold; when O1 and O3 act together, they

can reduce the operon activity up to 440 times; when all three operators are present, the repression intensity can be as high as 1300-fold.

Also, in [36] it has been established that the intracellular production of cyclic AMP (cAMP) decreases as the concentration of extracellular glucose increases. cAMP further binds a specific receptor molecule (CRP) to form the so-called CAP complex. Finally, CAP binds a specific DNA site (denoted here as C) upstream from the lac promoter, and by doing so it increases the affinity of the mRNA polymerase for this promoter. This regulatory mechanism is known as catabolite repression.

A novel source of cooperativity has been recently discovered [42] in the lactose operon: when a CAP complex is bound to site C, it bends DNA locally and increases the probability of the complex in which a repressor simultaneously binds operators O1 and O3.

The last regulatory mechanism in the lac operon is a so-called inducer exclusion. In it, external glucose decreases the efficiency of lac permease to transport lactose, and by doing so negatively affects the induction of the operon genes; see for example [36].

These regulatory mechanisms which we have briefly reviewed above are summarized in Fig. 3.3. As we have seen, the activity of the lactose operon is regulated by extracellular glucose and lactose. While extracellular glucose decreases the operon activity via catabolite repression and inducer exclusion, extracellular lactose increases the operon expression level by deactivating the repressor. Another important point is the existence of a positive feedback loop: as more molecules of lactose permease and β-galactosidase are produced, there is an elevated lactose uptake flux and an increased lactose metabolism rate; this further increases the production of allolactose and, as a consequence, diminishes the amount of active repressor. This, in turn, increases the operon activity, and thus more lactose permease and β-galactosidase molecules are produced.

The reader interested in the details of the lac operon regulatory mechanisms is referred to the excellent review [43] and the references therein. A good description of the operon regulatory elements and their location on the DNA chain can be found in [36]. The most recent discoveries regarding the cooperativity between CAP-binding site and operator O3 are [42].

Probability that a polymerase is bound to the promoter and a transcription initiates. Santillan and co-workers

([34] and [29]) have taken into account all the details of the lactose operon control system described above and deduced an explicit formula for the probability $P(O_P)$ as a function of the allolactose A and external glucose G_e concentrations. This rather complicated expression is given by

$$P(O_P) = p_{pc}(G_e)\mathcal{P}_R(A), \tag{3.10}$$

$$p_{pc}(G_e) = p_p \frac{1 + (k_{pc}-1)p_c(G_e)}{1 + (k_{pc}-1)p_p p_c(G_e)}, \tag{3.11}$$

$$p_{cp}(G_e) = p_c(G_e) \frac{1 + (k_{pc}-1)p_p}{1 + (k_{pc}-1)p_p p_c(G_e)}, \tag{3.12}$$

$$p_c(G_e) = \frac{K_G^m}{K_G^m + G_e^m}, \tag{3.13}$$

$$\mathcal{P}_R(A) = \frac{(1 + \xi_2\rho(A))(1 + \xi_3\rho(A)) + \xi_1^*\rho(A)^2}{\mathcal{Z}(A) + \prod_{j=1,2,3}(1 + \xi_j\rho(A))}, \tag{3.14}$$

$$\mathcal{Z}(A) = \sum_{j=1,2,3} p_{cp}(G_e)^{\delta_{2j}}\left(1 + \xi_j\rho(A)\right)\xi_j^*\rho(A)^2, \tag{3.15}$$

$$\rho(A) = \left(\frac{K_A}{K_A + A}\right)^2. \tag{3.16}$$

In the following few paragraphs we explain, step by step, the elements of this expression.

The function $\mathcal{P}_R(A)$ in (3.14) accounts for the regulation of transcription initiation by active repressors, giving the probability that the lactose promoter is not repressed by an active repressor bound to Operator O1. It accounts for the interactions of the repressor and allolactose molecules, of the repressor molecules and the three different lactose operators (including DNA looping), of the CAP activator and the mRNA polymerase, and of CAP and the DNA loop involving operators O1 and O3.

Repressor molecules are tetramers formed by the union of two active dimers. Every one of the four repressor subunits can be bound by an allolactose molecule. According to [40], free repressors, repressors bound by one allolactose, and repressors bound by two allolactoses in the same dimer can bind a single operator. The fraction of repressors able to do so is denoted by $\rho(A)$ in (3.16). Conversely, only free repressors, whose fraction is given by $\rho(A)^2$, can bind two different operators simultaneously.

The function $p_{pc}(G_e)$ in (3.11) denotes the modulation of transcription initiation by the cooperative interaction between a CAP activator and a polymerase, each bound to its respective site. Production of cyclic AMP (cAMP) is inhibited by extracellular glucose G_e. cAMP further binds the so-called cAMP receptor protein to form the CAP complex. Finally, CAP binds a specific site near the lactose promoter and enhances transcription initiation. The probability of finding a CAP molecule bound to its corresponding site is given by $p_c(G_e)$ in (3.13).

The probability that a CAP activator is bound to its corresponding site is given by the function $p_{cp}(Ge)$ in (3.12). Its presence in the definition of $\mathcal{P}_R(A)$ in (3.14) accounts for the fact that it affects the formation of the DNA loop in which a single repressor binds operators O1 and O3 at the same time. Note that $(p_{cp})^{\delta_{2j}}$ is equal to p_{cp} only when $j = 2$ and it is equal to one in any other case.

Reduced model. The system of equations developed above can be reduced after assuming that the set of parameters $(\tau_Z, k_Z, \tau_Z^*, \gamma_Z)$ is equal to $(\tau_Y, k_Y, \tau_Y^*, \gamma_Y)$, because in this case Equations (3.1) and (3.2) are identical, and in the same way (3.3) is identical to (3.4). Thus, recalling Equations (3.6) and (3.10), we obtain the reduced system

$$\frac{dM_Z}{dt} = Dk_M p_{pc}(G_e)\mathcal{P}_R(A) - (\gamma_M + \mu)M_Z, \quad (3.17)$$

$$\frac{dE_Z}{dt} = k_Z e^{-\mu\tau_Z^*} M_{Z,\tau_Z^*} - (\gamma_Z + \mu)E_Z, \quad (3.18)$$

$$\frac{dL}{dt} = k_L \beta_L(L_e)\beta_G(G_e)Q - k_\ell \beta_\ell(L)Q$$
$$\qquad -\phi_M \mathcal{M}(L)B - (\gamma_L + \mu)L. \quad (3.19)$$

The functions $p_{pc}(G_e)$ and $\mathcal{P}_R(A)$ are given in Equations (3.10) to (3.16). Finally, if we assume in (3.9) that the equality $\alpha\phi_M = \phi_A$ holds, the sum $\gamma_A + \mu$ is very small (close to zero), and the allolactose dynamics is very fast, then we can assume that $A = L$. Thus, we complete the model for the lactose operon by adding the Equations (3.5) to (3.8),

$$B = E_Z/4, \quad (3.20)$$

$$Q = E_Z, \quad (3.21)$$

$$A = L, \quad (3.22)$$

$$\beta_L(L_e) = \frac{L_e}{\kappa_L + L_e}, \quad (3.23)$$

Table 3.1
Values of the parameters in the lactose operon model equations. The parameter K_A is the only one that we were unable to estimate. The unit *mpb* stands for molecules per average-size bacterium

$D \approx 2\text{mpb}$	$k_M \approx 0.18/\text{min}$
$\mu \approx 0.02/\text{min}$	$\gamma_M \approx 0.46/\text{min}$
$p_p \approx 0.127$	$k_{pc} \approx 30$
$K_G \approx 2.6\mu M$	$m \approx 1.3$
$\xi_1 \approx 17$	$\xi_2 \approx 0.85$
$\xi_3 \approx 0.17$	$\xi_1^* \approx 0$
$\xi_2^* \approx 430.6$	$\xi_3^* \approx 1261.7$
$K_A \geq 0$ mpb	$k_Z \approx 18.8/\text{min}$
$\gamma_Z \approx 0.01/\text{min}$	$\tau_Z \approx 0.1\text{min}$
$\tau_Z^* \approx 0.42\text{min}$	$k_L \approx 6.0 \times 10^4/\text{min}$
$k_\ell \approx 0/\text{min}$	$\kappa_\ell \approx 680\mu M$
$\phi_M \in [0, 3.8 \times 10^4]/\text{min}$	$\kappa_L \approx 680\mu M$
$\phi_G \approx 0.35$	$\kappa_G \approx 1.0\mu M$
$\kappa_M \approx 7.0 \times 10^5\text{mpb}$	

$$\beta_G(G_e) = 1 - \phi_G \frac{G_e}{\kappa_G + G_e}, \quad (3.24)$$

$$\beta_\ell(L) = \frac{L}{\kappa_\ell + L}, \quad (3.25)$$

$$\mathcal{M}(L) = \frac{L}{\kappa_M + L}. \quad (3.26)$$

The parameters of the model (3.10) to (3.26) are given in Table 3.1 as estimated from the experimental literature, see [32, 34] and [29].

Comparison with experimental results. In [44], experiments were carried out in which *E. coli* cultures were grown in M9 minimal medium, with succinate as the main carbon source, supplemented with varying amounts of glucose and trimethylglycine (TMG). They engineered a DNA segment in which the *gfp* gene was under the control of the wild-type lactose promoter, and inserted this segment into the chromosome of the cultured *E. coli* bacteria, at the λ-insertion site. In these mutant bacteria, Ozbudak et al. estimated the lactose operon expression level in each bacterium by simply measuring the intensity of green fluorescence.

Experimentally [44] it has been observed that the histograms of fluorescence intensities were unimodal, and that the mean value corresponded to low induction levels of the lactose operon, when the bacterial growth medium had low TMG levels. After the TMG concentration surpassed a given threshold, the histograms became bimodal, which can be viewed as evidence for bistability: the original (new) mode corresponds to the uninduced (induced) steady state. With further increments of the TMG concentration,
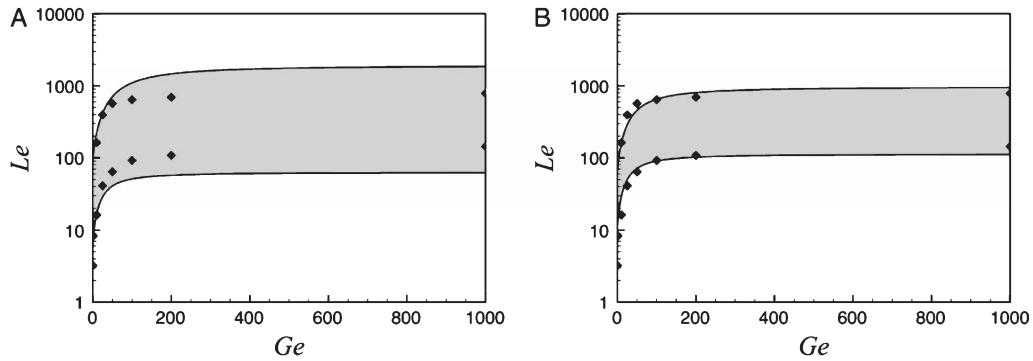
Fig. 3.4. Bifurcation diagrams in the L_e versus G_e parameter space, calculated with the present model under various conditions. In both cases, the bistability region is shaded, while the monostable induced (top) and uninduced (bottom) regions are uncoloured. In the left graph we used the parameter values tabulated in Table 3.1, and set $K_A = 8.2 \times 10^5 mpb$. In the right graph all parameters ξ_j and ξ_j^* were reduced to 0.15 times the value reported in Table 3.1, and K_A was set to $2.8 \times 10^6 mpb$. To calculate the graphs, we set $\phi_M = 0/min$ to simulate induction of the lac operon with the non-metabolizable TMG. The K_A values referred to above were chosen to fit the experimental results of [44], which are shown with solid diamonds. Taken from [34] with permision.

the mode corresponding to the uninduced state disappeared, and the histogram became unimodal again. When the experiment was repeated by decreasing the concentrations of TMG, the opposite behaviour was observed. Ozbudak et al. measured the range of TMG concentrations for which bistability was obtained, for several concentrations of external glucose. When they repeated the same experiments with the natural inducer (lactose), they were unable to find analogous evidence for bistability, even when lactose was given at saturation levels. In these last experiments, they employed glucose concentrations in the same range as in the experiments with TMG.

Noting that TMG inactivates the lactose repressor, but it is not metabolizable, we simulate the Ozbudak et al. experiments. For this, we set $\phi_M = 0/min$ to account for the presence of a reliable carbon source (succinate) and induction with TMG, which is not metabolized by β-galactosidase. Then, we calculated the bifurcation points and plotted them in the L_e versus G_e parameter space. We took K_A as a free parameter, and found that $K_A = 8.2 \times 10^5 mpb$ (here and thereafter mpb means molecules per average-size bacterium) gives a reasonable fit to the experimental points of Ozbudak et al. Both the model bifurcation diagram and the experimental points are presented in Fig. 3.4A. Note that the bistability region predicted by the model is wider than the experimental one. There are three possible explanations for this discrepancy: 1) the lactose promoter-*gfp* fusion employed by Ozbudak et al. as a reporter lacks operators O2 or O3; 2) the diffi-

culty in measuring exactly the L_e values at which the bimodal histograms appear and disappear; and 3) the phase diagram of Fig. 3.4A is based upon a mean-field analysis, and so biochemical noise can change the phase boundaries; see Section 4 below. A fourth possible explanation for the disagreement between the model and the experimental results is that our estimated parameter values differ from those corresponding to the *E. coli* strain used by Ozbudak et al. To account for this possibility, we explored the parameter space looking for a better fit. We found that it can be obtained by decreasing the parameters ξ_j and ξ_j^* to 15% of the values reported in Table 3.1, and by setting $K_A = 2.8 \times 10^6 mpb$. The results are shown in Fig. 3.4B.

3.2. The tryptophan operon

Tryptophan is one of the 20 amino acids out of which all proteins are made. Arguably, tryptophan is the most expensive amino acid to synthesize, biochemically speaking. Perhaps, for this reason, humans and many other mammals do not have the enzymes necessary to catalyze tryptophan synthesis and instead they find this amino acid in their diet.

However, microorganisms like *E. coli* generally posses the machinery to produce tryptophan, but the production process is tightly regulated in all cases. In the particular case of *E. coli*, the tryptophan operon is a DNA segment containing a promoter (*trpR*) where transcription starts and regulation by repression takes place, a leader region (*trpL*) where regulation by
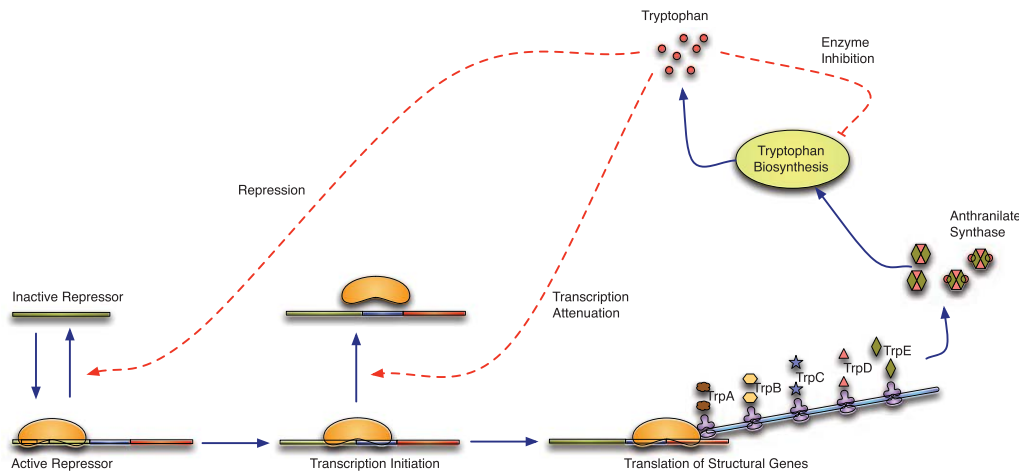
Fig. 3.5. Schematic representation of the tryptophan operon regulatory pathway. Solid lines represent the processes involved in gene expression and tryptophan synthesis, while dashed lines correspond to the operon regulatory mechanisms.

transcriptional attenuation occurs, and five structural genes (*trpE* to *trpA*) that code for the polypeptides comprising the enzymes responsible for the catalysis of tryptophan biosynthesis. There are three different regulatory mechanisms involved in the control of the tryptophan operon dynamics: repression, transcriptional attenuation, and enzyme inhibition. The tryptophan regulatory pathway is illustrated in Fig. 3.5.

Repression occurs when an active repressor binds to one of the three available binding sites within the promoter, inhibiting the binding of a RNA polymerase, and so of transcription initiation. The repressor molecule is a homo-dimer made up of two TrpR polypeptides. Each subunit has a binding site for tryptophan, and the repressor molecules activate when both tryptophan binding sites are occupied. Of the three repressor binding sites within the promoter, the two closest to the transcription initiation site interact cooperatively. That is, when two are bound to such sites, the resulting complex is much more stable than it would be expected from the addition of the individual binding energies.
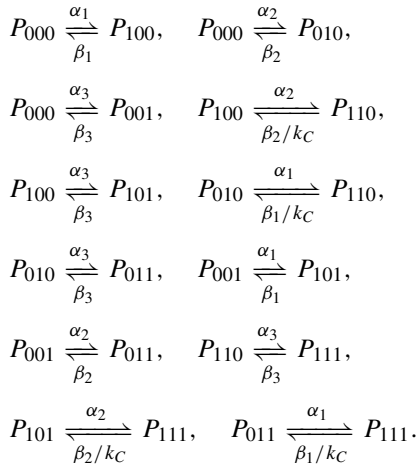
Transcriptional attenuation is regulated by the DNA leading region. The RNA strand resulting from transcription of *trpL* can fold into three alternative hairpin-like structures, as a result of Watson-Crick base pairing. Soon after transcription initiation, the first hairpin structure is formed, and this causes the polymerase to pause transcription. When a ribosome binds to the nascent RNA strand to start translation, it eventually disrupts the hairpin and both transcription and translation proceed together. Not long after that, the

ribosome encounters two tryptophan codons in tandem. Under conditions of abundant tryptophan, there is a large amount of charged *trp* transfer RNA, and so the two consecutive tryptophan codons are rapidly translated. When this occurs, a second hairpin structure, that serves as a transcription termination signal, forms and transcription is prematurely aborted. Conversely, if tryptophan is scarce, the ribosome stops at the *trp* codons while the RNA polymerase continues transcribing the rest of the leading region. This prevents the formation of the transcription-terminating hairpin and instead promotes the formation of a third structure that allows the polymerase to go into the structural genes to transcribe them.

Tryptophan biosynthesis takes place through a series of reactions, each one catalyzed by enzymes formed from the polypeptides coded by genes *trpE-A*. The first of those reactions, and the slowest one, is catalyzed by the enzyme anthranilate synthase. In this reaction, anthranilate is synthesized out of chorismic acid. Being the slowest reaction of the tryptophan synthesis path, anthranilate synthesis determines the velocity of the whole process. Furthermore, anthranitale synthase is a heterotetramer made up of two *TrpE* and two *TrpD* subunits. Each *TrpE* subunit has a binding site for tryptophan, and when they are bound by this amino acid, the whole enzyme suffers an allosteric transformation that makes it unable catalyze the corresponding reaction. This regulatory mechanism is known as enzyme inhibition.

A deterministic model for this regulatory pathway can be constructed as follows. Consider first the

dynamics of promoter switching. Denote the state of repression of the promoter as (i, j, k)—with $i, j, k = 0, 1$; a value of 1 means that the corresponding repressor binding site is occupied, while a value of 0 means that it is empty. If P_{ijk} represents the average number of promoters whose state is (i, j, k), the chemical reactions through which the promoter switches between its different available states are:

$$P_{000} \underset{\beta_1}{\overset{\alpha_1}{\rightleftharpoons}} P_{100}, \quad P_{000} \underset{\beta_2}{\overset{\alpha_2}{\rightleftharpoons}} P_{010},$$

$$P_{000} \underset{\beta_3}{\overset{\alpha_3}{\rightleftharpoons}} P_{001}, \quad P_{100} \underset{\beta_2/k_C}{\overset{\alpha_2}{\rightleftharpoons}} P_{110},$$

$$P_{100} \underset{\beta_3}{\overset{\alpha_3}{\rightleftharpoons}} P_{101}, \quad P_{010} \underset{\beta_1/k_C}{\overset{\alpha_1}{\rightleftharpoons}} P_{110},$$

$$P_{010} \underset{\beta_3}{\overset{\alpha_3}{\rightleftharpoons}} P_{011}, \quad P_{001} \underset{\beta_1}{\overset{\alpha_1}{\rightleftharpoons}} P_{101},$$

$$P_{001} \underset{\beta_2}{\overset{\alpha_2}{\rightleftharpoons}} P_{011}, \quad P_{110} \underset{\beta_3}{\overset{\alpha_3}{\rightleftharpoons}} P_{111},$$

$$P_{101} \underset{\beta_2/k_C}{\overset{\alpha_2}{\rightleftharpoons}} P_{111}, \quad P_{011} \underset{\beta_1/k_C}{\overset{\alpha_1}{\rightleftharpoons}} P_{111}.$$

In these reactions α_i represents the effective reaction rate constant for the binding of an active repressor to the i-th binding site in the promoter, β_i is the corresponding unbinding reaction rate constant, and k_C accounts for the cooperativity between the first two repressor binding sites.

By making use of the theory of chemical kinetics we can write the following set of differential equations governing the dynamics of variables P_{ijk}:

$$\frac{dP_{000}}{dt} = -(\alpha_1 + \alpha_2 + \alpha_3)P_{000} + \beta_1 P_{100}$$
$$+\beta_2 P_{010} + \beta_3 P_{001}, \tag{3.27}$$

$$\frac{dP_{100}}{dt} = -(\beta_1 + \alpha_2 + \alpha_3)P_{100} + \alpha_1 P_{000}$$
$$+\frac{\beta_2}{k_C} P_{110} + \beta_3 P_{101}, \tag{3.28}$$

$$\frac{dP_{010}}{dt} = -(\alpha_1 + \beta_2 + \alpha_3)P_{010} + \frac{\beta_1}{k_C} P_{110}$$
$$+\alpha_2 P_{000} + \beta_3 P_{011}, \tag{3.29}$$

$$\frac{dP_{001}}{dt} = -(\alpha_1 + \alpha_2 + \beta_3)P_{001} + \beta_1 P_{101}$$
$$+\beta_2 P_{011} + \alpha_3 P_{000}, \tag{3.30}$$

$$\frac{dP_{110}}{dt} = -\left(\frac{\beta_1}{k_C} + \frac{\beta_2}{k_C} + \alpha_3\right) P_{110} + \alpha_1 P_{010}$$
$$+\alpha_2 P_{100} + \beta_3 P_{111}, \tag{3.31}$$

$$\frac{dP_{101}}{dt} = -(\beta_1 + \alpha_2 + \beta_3) P_{101} + \alpha_1 P_{001}$$
$$+\frac{\beta_2}{k_C} P_{111} + \alpha_3 P_{100}, \tag{3.32}$$

$$\frac{dP_{011}}{dt} = -(\alpha_1 + \beta_2 + \beta_3) P_{011} + \frac{\beta_1}{k_C} P_{111}$$
$$+\alpha_2 P_{001} + \alpha_3 P_{010}, \tag{3.33}$$

$$\frac{dP_{111}}{dt} = -\left(\frac{\beta_1}{k_C} + \frac{\beta_2}{k_C} + \beta_3\right) P_{111} + \alpha_1 P_{011}$$
$$+\alpha_2 P_{101} + \alpha_3 P_{110}. \tag{3.34}$$

These equations do not constitute a complete set because the effective binding reaction rate constants α_i are directly proportional to the amount of active repressors, R_A, which is, in turn, a function of the intracellular tryptophan concentration.

To complete the differential equation system let M represent the concentration of mRNA molecules resulting from transcription of the tryptophan operon, E be the concentration of anthranilate synthase enzymes, and T denote the intracellular tryptophan level. Following the development in previous sections, the differential equations accounting for the dynamics of these variables are:

$$\frac{dM}{dt} = k_M P_{000} \mathcal{A}(T) - \gamma_M M, \tag{3.35}$$

$$\frac{dE}{dt} = k_E M - \gamma_E E, \tag{3.36}$$

$$\frac{dT}{dt} = k_T E \mathcal{I}(T) - \gamma_T \frac{T}{T + K_T}, \tag{3.37}$$

in which k_M is the transcription initiation rate, $\mathcal{A}(T)$ represents the probability that a newly initiated transcriptional event is not prematurely aborted due to attenuation, γ_M accounts for the mRNA degradation rate, k_E is the enzyme synthesis rate per mRNA molecule, γ_E is the enzyme degradation rate, k_T represents the tryptophan synthesis rate per active enzyme, $\mathcal{I}(T)$ is the probability that an enzyme is not inhibited by tryptophan, γ_T is the maximal tryptophan consumption rate due to the cellular metabolism, and K_T is the corresponding half saturation constant.

The reaction rate constants for repressor binding are proportional to the concentration of active repressors $R_A(T)$. That is,

$$\alpha_i = a_i R_A(T). \tag{3.38}$$

Thus, expressions for $R_A(T)$, $\mathcal{A}(T)$, and $\mathcal{I}(T)$ are required to complete the model. These functions correspond to the three known regulatory mechanisms in this system: repression, transcriptional attenuation, and enzyme inhibition, respectively. Functions $R_A(T)$, $\mathcal{A}(T)$, and $\mathcal{I}(T)$ were derived in [45] from chemical kinetics considerations by taking into account all the chemical reactions behind the corresponding regulatory mechanisms. The resulting expressions are:

$$R_A(T) = R_{Tot}\left(\frac{T}{T + K_A}\right)^2, \tag{3.39}$$

$$\mathcal{A}(T) = \frac{1 + 2\alpha\frac{T}{K_G+T}}{\left(1 + \alpha\frac{T}{K_G+T}\right)^2}, \tag{3.40}$$

$$\mathcal{I}(T) = \frac{K_I^n}{K_I^n + T^n}. \tag{3.41}$$

Here, R_{Tot} is the total number of repressor molecules, K_T the dissociation constant between tryptophan and one binding site of a repressor, K_G is the dissociation constant between tryptophan and the corresponding transfer RNA, α the probability per unit time that a charged tRNA$^{\text{Trp}}$ arrives at a tryptophan codon so that it is translated, K_I is the dissociation constant between tryptophan and one of the TrpE subunits in anthranilate synthase, and n is a Hill coefficient.

Equations (3.27)–(3.41) constitute a complete system of differential equations that model the dynamics of the tryptophan operon. However, due to its high dimensionality, this system is quite difficult to analyze. For that reason, some simplifying assumptions are useful. One which has been widely employed consists in supposing that the dynamics of repressor binding and unbinding are much faster than those of mRNA and protein synthesis and degradation, as well as those of tryptophan production and consumption. If this is the case, the subsystem given by Equations (3.27)–(3.34) is much faster than that given by Equations (3.35)–(3.37), and so one can make a quasi steady state approximation (also known as adiabatic elimination) for Equations (3.27)–(3.34), with which the model transforms into

$$\frac{dM}{dt} = k_M P_{000}(T)\mathcal{A}(T) - \gamma_M M, \tag{3.42}$$

$$\frac{dE}{dt} = k_E M - \gamma_E E, \tag{3.43}$$

$$\frac{dT}{dt} = k_T E \mathcal{I}(T) - \gamma_T \frac{T}{T + K_T}. \tag{3.44}$$

The concentration of non-repressed promoters is given in this case by

$$\begin{aligned}
P_{000}(T) = \Big(&1 + \frac{a_1}{\beta_1}R_A(T) + \frac{a_2}{\beta_2}R_A(T) + \frac{a_3}{\beta_3}R_A(T) \\
&+ k_C\frac{a_1}{\beta_1}\frac{a_2}{\beta_2}R_A^2(T) + \frac{a_1}{\beta_1}\frac{a_3}{\beta_3}R_A^2(T) \\
&+ \frac{a_2}{\beta_2}\frac{a_3}{\beta_3}R_A^2(T) + k_C\frac{a_1}{\beta_1}\frac{a_2}{\beta_2}\frac{a_3}{\beta_3}R_A^3(T)\Big)^{-1}.
\end{aligned} \tag{3.45}$$

$P_{000}(T)$, $\mathcal{A}(T)$, and $\mathcal{I}(T)$ are monotonic sigmoidally decreasing functions of T, and so is the product $P_{000}(T)\mathcal{A}(T)$. This product is sometimes replaced by a decreasing Hill function [46].

As explored extensively in Section 2, an elementary classification of systems subject to feedback regulation includes those with negative feedback or, alternately, those with positive feedback. This is important because the type of feedback determines the kind of expected dynamic behaviour. Thus, positive feedback is necessary for bistability, while negative feedback is the mechanism underlying cyclic behaviour. Given that the tryptophan operon has been experimentally studied for several decades (and, thus, is one of the best known molecular systems), and that it is regulated by three different negative feedback loops, this system has become a paradigm for studying the effects of negative feedback regulation on gene expression. Below we review some of the most prominent past studies of the tryptophan operon.

As discussed in Section 2.2, the first mathematical model for a repressible operon was due to Goodwin [3], who developed a model with a structure equivalent to that in Equations (3.42)–(3.44), except that the regulatory functions accounting for transcriptional attenuation and enzyme inhibition were not taken into account. The repression regulatory function was modeled in the Goodwin model by a monotone decreasing Hill function. In a later paper, Goodwin [2] presented analog computer simulations of limit cycles (sustained oscillations) obtained from this model with a Hill expo-

nent of one. However, Griffith [4] later demonstrated that the steady state is locally stable up to a Hill exponent equal to 8, making limit cycle oscillations highly unlikely for low exponent values. In a large number of simulations Griffith found limit cycles only if the steady state was unstable. Apparently, there was an error in Goodwin's analog simulation. The controversy was finally resolved by Tyson [47], who analytically proved the existence of at least one periodic solution whenever the steady state is unstable.

As we saw in the previous paragraph, the first modeling studies on a repressible operon focused on the possibility of sustained oscillations, and ended with a negative conclusion. This question was revisited in [48], who modified the Goodwin model to include the transcriptional and translational time delays as well as the regulatory function accounting for enzyme inhibition. Bliss et al. demonstrated that time delays can induce sustained oscillations, but only when enzyme inhibition is weakened. They also presented experimental results with a mutant strain of *E. coli* in which the enzyme anthranilate synthase cannot be inhibited by tryptophan. This strain was first grown in a tryptophan-rich medium and then suddenly changed to a tryptophan-less medium to induce expression of the tryptophan operon genes. Both the simulations and the experiments showed periodic oscillation in both the enzyme and the tryptophan intracellular concentrations.

In later work [49] the Goodwin model was further refined by deriving a repression regulatory function from first principles, taking into consideration the underlying chemical reactions. Nonetheless, they dismissed the regulatory functions corresponding to transcription attenuation and enzyme inhibition. In [49] the possible complex behaviours the tryptophan operon can show, given the architecture of the regulatory network, was investigated. They found that the steady state, although normally stable, becomes unstable for super-repressing strains, even at low values of the cooperativity of repression. However, in order for this to happen it is necessary that the demand for end-product saturates at large end-product concentrations. Finally, in [49] it was proved that the system can also show bistability, in which a stable steady state and a stable limit cycle coexist.

In 1990, other investigators [50] introduced one more model for the tryptophan operon regulatory pathway, and used it to investigate the possibility of engineering an *E. coli* strain to overproduce trypto-

phan. The model [50] has a similar structure to that in Equations (3.42)–(3.44) but, as some of the models reviewed in the former paragraphs, it ignores the transcriptional attenuation and the enzyme inhibition regulatory mechanisms. Through analytical studies and numerical simulations the authors were able to demonstrate that stable overproduction is feasible. Nevertheless, under some specific circumstances the operon may become unstable and lead to periodic synthesis. In [51] the models of [49] and [50] were further refined and employed to study the influence of periodic fluctuations in the intracellular demand for tryptophan.

In our group we have studied the dynamic behaviour of the tryptophan operon regulatory pathway for some time. In [52] we developed a mathematical model that accounts for all known regulatory mechanisms, as well as for the time delays due to transcription and translation. Moreover, we put special attention to estimating all of the model parameters from reported experimental data. Although involving one extra differential equation, a more careful analysis reveals that this model is equivalent to that in Equations (3.42)–(3.44). To test the model feasibility, we compared its predictions with available dynamic experiments from wild type and two mutant strains. Later [45], we simplified our original model, but still considered all three existing regulatory mechanisms, and analyzed their influence on the system dynamic behaviour. We numerically showed that enzyme inhibition is the fastest responding mechanism. However, although it could suffice to efficiently control tryptophan biosynthesis, it would be very expensive because it would imply continuous production of enzymes. Although repression and transcription attenuation respond considerably more slowly, they allow bacteria to diminish the energy expended in enzyme synthesis when tryptophan demand is low for longer periods of time. In other words, the redundancy of feedback regulatory mechanisms allows *E. coli* to efficiently respond to both slow (via repression and transcription attenuation) and fast (via enzyme inhibition) fluctuations of tryptophan demand. These numerical results were analytically corroborated in [53], where we studied the global stability of the tryptophan operon model using the second Lyapunov method.

As we have seen, the first modeling studies on the tryptophan operon focused on the possibility that this system shows sustained oscillations under given circumstances. Interestingly, there is only one experimental report of such oscillatory behaviour in
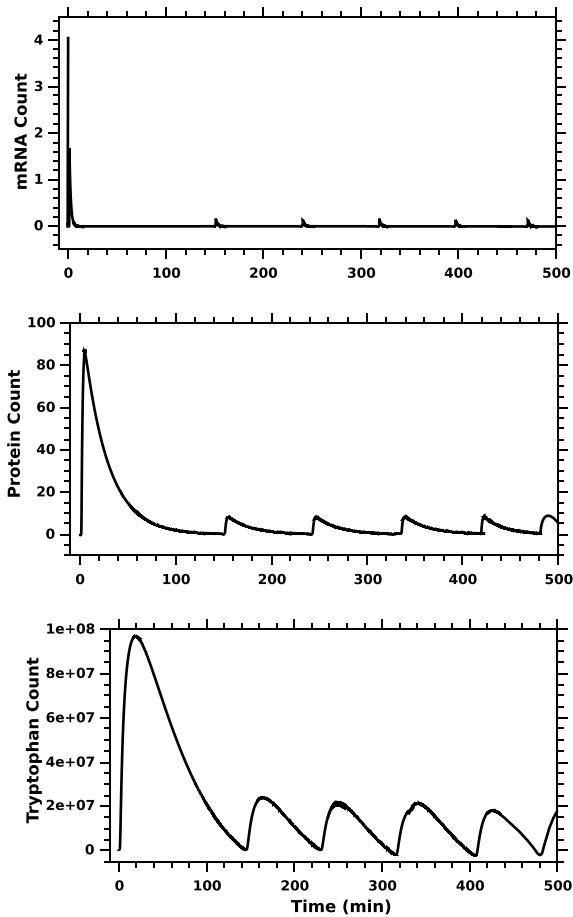
Fig. 3.6. Oscillatory behaviour during a de-repression experiment with an enzyme-inhibition-less *E. coli* mutant strain, as predicted by the deterministic model in [54].
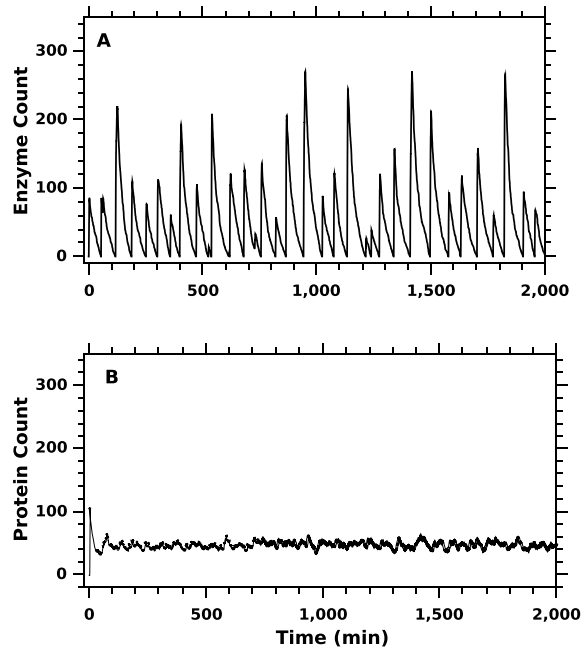

Fig. 3.7. (A) Stochastic quasi-oscillatory behaviour observed during a de-repression experiment with an enzyme-inhibition-less *E. coli* mutant strain, as predicted by the stochastic model in [54]. (B) Average behaviour of 100 independent cells.

the tryptophan operon [48]. Taking into consideration the lack of experimental evidence, as well as recent discoveries regarding the existence of multiple repressor binding sites within the *trp* promoter, and of cooperativity between two of them, we have further investigated the possibility of observing sustained oscillations in this system [54]. To that end, we improved the model in [45] by incorporating the discoveries discussed above and analyzed it numerically. We found that indeed, a mutant bacterial strain lacking enzyme inhibition can behave cyclically, and that the time delays due to transcription and translation are essential for this behaviour. In Fig. 3.6 we show the model results, which show a very good agreement with the experimental results in [48].

On the other hand, regular periodic oscillations are observed in the model of [54], but only when the system

intrinsic stochasticity is ignored. When the so-called intrinsic biochemical noise is taken into account, the system shows oscillations with variable periods, and this causes the global system behaviour in a cell population to be non cyclic overall. These results stress the necessity of further studying the appearance of oscillations in the tryptophan operon, both analytically and experimentally; not only to satisfy some people's scientific curiosity, but also because answering this question may shed some light into the dynamics of gene regulation. In Fig. 3.7 we show the stochastic quasi-periodic dynamic behaviour predicted by the mathematical model, as well as the average of 100 independent cells.

All the models reviewed so far have the structure of the model represented by Equations (3.42)–(3.44). This means that, either explicitly or implicitly, they assume that promoter gating between the various repressed and the non-repressed states is much faster than the transcription and translation processes. Nevertheless, recent detailed measurements of the repressor-promoter kinetics revealed that this assumption is not valid—see [55] and references therein. This further implies that the assumed separation of
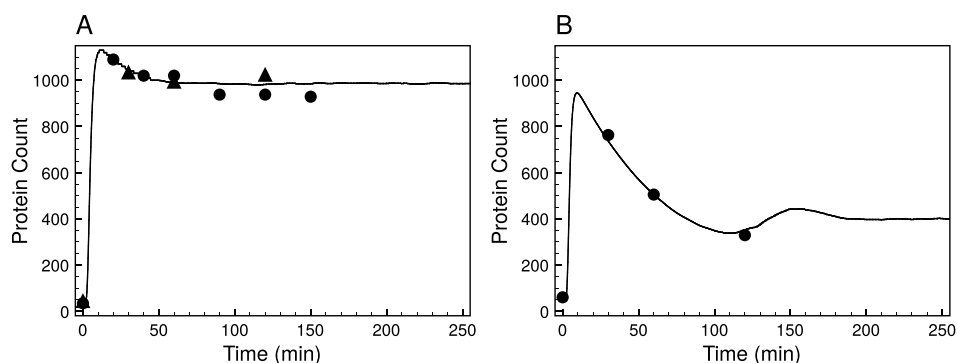
Fig. 3.8. Comparison of experimental results (circles and triangles) and model predictions for de-repression experiments carried out with a wild-type and an enzyme-inhibition-less mutant strain of *E. coli.*

time scales employed to obtain the simplified model in Equations (3.42)–(3.44) does not exist, and one is obliged to work with the full model: Equations (3.27)–(3.37). In a recent paper [55] we studied the stochastic behaviour of such a model, but analytical and numerical studies of the deterministic counterpart are still missing.

We wish to emphasize that in our modeling studies we have followed the strategy of producing models as detailed as possible, given the available experimental evidence. This meant that not only we included all known mechanisms into the model equations, but also that we estimated all of the model parameters from reported experimental data. Understandably, this is not always possible when developing models for biological systems. However, in this particular case, the tryptophan operon of *E. coli* is so well studied that developing this kind of model is completely feasible. A natural consequence of having quite detailed models is the possibility of accurately reproducing dynamic experiments. In particular, we have employed the experimental results of [56] to compare with our models' predictions. In Fig. 3.8 we show comparisons of model predictions and the Yanofsky and Horn experimental measurements for a wild type and for a enzyme-inhibition-less mutant *E. coli* strain. The theoretical simulations in Fig. 3.8 were carried out with our most detailed model [55], but qualitatively similar results are obtained with all the model versions previously reviewed. In our opinion, it is essential for a model to be able to reproduce existing dynamical experimental data, before it can be employed to answer dynamical questions not easily addressed experimentally.

E. coli is not the only bacterium with a tryptophan operon. Other bacteria also have an equivalent

system, in particular *B. subtilis*. Interestingly, the structure of the regulatory pathway in both systems is very similar, although the specific mechanisms are very different. For instance, instead of repression, the tryptophan operon in *B. subtilis* involves a so-called TRAP molecule that promotes premature transcription termination when it is bound by 11 tryptophan molecules. Instead of transcriptional attenuation, *B. subtilis* has a secondary *at* operon that is regulated by tryptophan and produces a protein that modulates the effect of TRAP proteins. The only mechanism that *E. coli* and *B. subtilis* share in common is enzyme inhibition. A model for the tryptophan operon of *B. subtilis* was developed in [57] and shown that not only its regulatory pathway has a similar structure to that of *E. coli*, but the analogous mechanisms in both systems play similar roles from a dynamic perspective. Given that the lineages of both organisms evolved separately several millions of years ago, these similarities may be the result of evolutionary convergence.

4. Noise effects in gene regulation: Intrinsic versus extrinsic

In all areas of science, when making experimental measurements it is noted that the quantity being measured does not have a smooth temporal trajectory but, rather, displays apparently erratic fluctuations about some mean value when the experimental precision is sufficiently high. These fluctuations are commonly referred to as 'noise' and usually assumed to have an origin outside the dynamics of the systems on which measurements are being made–although there have been many authors who have investigated the possi-

bility that the 'noise' is actually a manifestation of the dynamics of the system under study. Indeed, a desire to find ways to quantitatively characterize this 'noise' is what led, in large part, to the development of the entire mathematical field loosely known as stochastic processes, and the interaction of stochastic processes with deterministic dynamics is of great interest since it is important to understand to what extent fluctuations or noise can actually affect the operation of the system being studied.

Precisely the same issues have arisen in molecular biology as experimental techniques have allowed investigators to probe temporal behaviour at ever finer levels, even to the level of individual molecules. Experimentalists and theoreticians alike who are interested in the regulation of gene networks increasingly focus on trying to assess the role of various types of fluctuations on the operation and fidelity of both simple and complex gene regulatory systems. Recent reviews [58–60] give an interesting perspective on some of the issues confronting both experimentalists and modelers.

As in other areas of science, in gene regulation the debate often swirls around whether the fluctuations are extrinsic to the system under consideration [61–64], or whether they are an intrinsic part of the fundamental processes they are affecting (*e.g.* bursting, see below). The dichotomy is rarely so sharp however, but in [65] an elegant experimental technique has been presented to operationally distinguish between the two, see also [66], while [67] and [68] have partially set the stage for a theoretical consideration of this question. One issue that is raised persistently in considerations of the role of fluctuations or noise in the operation of gene regulatory networks is whether or not they are 'beneficial' [69] or 'detrimental' [70] to the operation of the system under consideration. This is, of course, a question of definition and not one that we will be further concerned with here since it is a question without scientific meaning.

In this section we study the density of the molecular distributions in generic bacterial operons in the presence of 'bursting' (commonly known as intrinsic noise in the biological literature) as well as inherent (extrinsic) noise using an analytical approach. In a very real sense, the whole field of intrinsic noise behaviour owes its basis to the pioneering work of Berg [71] who first studied the statistical fluctuations of protein numbers in bacterial population (with division) through the master equation approach, and introduced the concept of what is now called bursting. Our work is further motivated by the well documented production of mRNA

and/or protein in stochastic bursts in both prokaryotes and eukaryotes [72–79], and follows other mathematical contributions [80–101]. We stress, however, that we have not referenced studies in which stochasticity was studied solely using Gillespie simulations since these have become *de rigeur* for almost all supposed modeling efforts in spite of the fact that in and of themselves they yield little if any real insight.

Because of its relevance to the analysis of experimental data, we emphasize the behaviour of densities of gene regulatory constituents. To our knowledge, the analytical solution of the steady state density of the molecular distributions in the presence of bursting was first derived in [81]. Our approach emphasized here extends these results to show the global stability of the limiting densities and examines their bifurcation structure to give a complete understanding of the effect of bursting on molecular distributions.

4.1. Dynamics with bursting

4.1.1. Generalities

In this section we model the amount of the dominant protein as a Markov process $\{x(t)\}_{t \geq 0}$ with values in $(0, \infty)$. Let $x(t)$ denote the amount of the protein in a cell at time t, $t \geq 0$. Following [73, 81] we assume that the *amplitude* of protein production through bursting translation of mRNA is exponentially distributed, that the frequency of bursting φ is dependent on the level of the protein, and that protein molecules undergo degradation with rate γ. We take here $\varphi(x) = \gamma \kappa_b f(x)$ and $\kappa_b \equiv \varphi_m$ in contrast to the deterministic case where $\kappa_d = b_d \varphi_m$. If only degradation were present, then $x(t)$ would satisfy the equation

$$x'(t) = -\gamma x(t), \quad t \geq 0.$$

However, we interrupt the degradation at random times

$$t_1 < t_2 < \ldots$$

occurring with intensity φ, *i.e.*,

$$\Pr(t_k - t_{k-1} > t | x(t_{k-1}) = x) = e^{-\int_0^t \varphi(e^{-\gamma s} x) ds},$$

$$t, x > 0.$$

At each t_k a random amount e_k of protein molecules is produced according to an exponential distribution with density

$$h(x) = \frac{1}{b}e^{-x/b}. \qquad (4.1)$$

Consequently the process is given by

$$x(t) = \begin{cases} e^{-\gamma(t-t_{k-1})}x(t_{k-1}), & t_{k-1} \le t < t_k, \\ e^{-\gamma(t_k-t_{k-1})}x(t_{k-1}) + e_k, & t = t_k, k = 1, 2, \ldots \end{cases}$$

The corresponding master equation for the evolution of the density $u(t, x)$ of $x(t)$ is given by

$$\frac{\partial u(t, x)}{\partial t} - \gamma \frac{\partial(xu(t, x))}{\partial x}$$
$$= -\gamma \kappa_b f(x)u(t, x)$$
$$+ \gamma \kappa_b \int_0^x f(y)u(t, y)h(x - y)dy. \qquad (4.2)$$

A stationary solution of Equation (4.2), which now becomes

$$-\frac{d(xu_*(x))}{dx} = -\kappa_b f(x)u_*(x)$$
$$+ \kappa_b \int_0^x f(y)u_*(y)h(x - y)dy,$$

with h given by (4.1) and nonnegative f, is of the form

$$u_*(x) = \frac{\mathcal{C}}{x}e^{-x/b}\exp\left[\kappa_b \int^x \frac{f(y)}{y}dy\right], \qquad (4.3)$$

where $\mathcal{C}$ is a normalizing constant, if u_* is integrable. The next result follows from [102].

Theorem 4.1. *Suppose that h is exponential as in (4.1) with $b > 0$ and that*

$$\mathcal{C} := \int_0^\infty \frac{1}{x}e^{-x/b}\exp\left[\kappa_b \int^x \frac{f(y)}{y}dy\right]dx < \infty.$$

Then u_* defined in (4.3) is the unique stationary density of (4.2) and the solution $u(t, x)$ of (4.2) is asymptotically stable in the sense that

$$\lim_{t\to\infty} \int_0^\infty |u(t, x) - u_*(x)|dx = 0$$

for all initial densities $u(0, x)$.

4.1.2. Distributions in the presence of bursting

We consider the situation in which the function f in the burst frequency $\varphi = \gamma \kappa_b f$ is given [103] by

$$f(x) = \frac{1 + \Theta x^n}{\Lambda + \Delta x^n},$$

where Λ, Δ, n are positive constants and $\Theta \ge 0$. We take $\Theta = 1$ to get f as defined in (2.10) for both the generic inducible and repressible operons treated in Section 2.1 with the constants Λ, Δ enumerated in Table 2.1. We have

$$\kappa_b \int^x \frac{f(y)}{y}dy = \int^x \frac{\kappa_b}{y}\left[\frac{1 + \Theta y^n}{\Lambda + \Delta y^n}\right]dy$$
$$= \ln\left\{x^{\kappa_b\Lambda^{-1}}(\Lambda + \Delta x^n)^\theta\right\},$$

where

$$\theta = \frac{\kappa_b}{n\Delta}\left(\Theta - \frac{\Delta}{\Lambda}\right).$$

Thus, the stationary density (4.3) explicitly becomes

$$u_*(x) = \mathcal{C}e^{-x/b}x^{\kappa_b\Lambda^{-1}-1}(\Lambda + \Delta x^n)^\theta. \qquad (4.4)$$

Observe that in the absence of control, *i.e.*, if $f \equiv 1$ or, equivalently, $\Theta = \Lambda = \Delta = 1$, we obtain, as in [81], the density of the gamma distribution:

$$u_*(x) = \frac{1}{b^{\kappa_b}\Gamma(\kappa_b)}e^{-x/b}x^{\kappa_b-1},$$

where $\Gamma(\cdot)$ denotes the gamma function. In particular, the first two terms of Equation 4.4 are proportional to the density of a gamma distribution.

The analysis of the qualitative nature of the stationary density (4.4) leads to different conclusions for the inducible and repressible operon cases, since the parameter θ is either positive or negative. In the rest of this section we assume that $\Theta = 1$. First note that we have $u_*(0) = \infty$ if $0 < \kappa_b\Lambda^{-1} < 1$ while $u_*(0) = 0$ for $\kappa_b\Lambda^{-1} > 1$ in which case there is at least one maximum at a value of $x > 0$. To calculate the number of maxima we use the fact that $u_*(x) > 0$ for all $x > 0$ and that

$$u_*'(x) = u_*(x)\left(\frac{\kappa_b f(x)}{x} - \frac{1}{b} - \frac{1}{x}\right), \quad x > 0.$$

Consequently, we have $u_*'(x) = 0$ for $x > 0$ if and only if

$$\kappa_b\left(\frac{x}{b} + 1\right) = \frac{1 + x^n}{\Lambda + \Delta x^n}. \qquad (4.5)$$

For $\theta \le 0$, as in the case of no control or a repressible operon, we have $\Lambda = 1$, $\Delta \ge 1$, and graphical arguments (see Fig. 4.1) easily show that Equation (4.5) may have none or one solution. Therefore, we have a stationary density which we can classify as
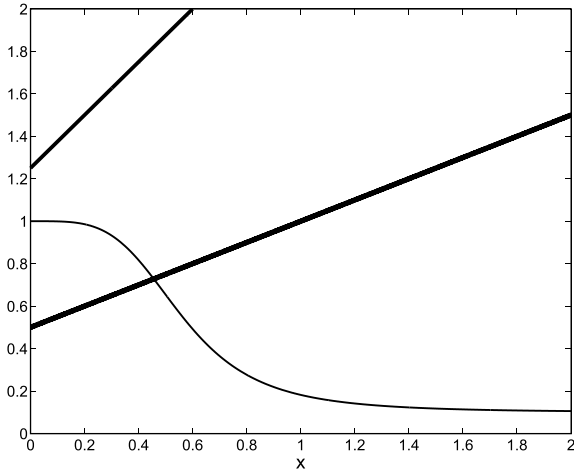
Fig. 4.1. Schematic illustration that there can be one or no solution of Equation (4.5), depending on the value of κ_b, with repressible regulation. The straight lines correspond (in a clockwise direction) to $\kappa_b = 2$ and $\kappa_b = 0.8$. This figure was constructed with $n = 4$, $\Delta = 10$ and $b = 1$. See the text for further details. Taken from [9] with permission.

Unimodal type 1 if $u_*(0) = \infty$ and u_* is decreasing, or

Unimodal type 2 if $u_*(0) = 0$ and u_* has a single maximum at a value of $x > 0$.

Observe that the stationary density u_* in the case of the repressible operon is Unimodal of type 1 if $0 < \kappa_b < 1$ and Unimodal of type 2 if $1 < \kappa_b$.

For $\theta > 0$, as in the case of an inducible operon, the stationary density becomes

$$u_*(x) = \mathcal{C}e^{-x/b}x^{\kappa_b K^{-1}-1}(K + x^n)^\theta,$$

$$\theta = \frac{\kappa_b}{n}(1 - K^{-1})$$

and there is the possibility that u_* may have more than one maximum, indicative of the existence of bistable behaviour. Graphical arguments (see Fig. 4.2) show that there may be up to three roots of

$$\frac{1}{\kappa_b}\left(\frac{x}{b} + 1\right) = \frac{1 + x^n}{K + x^n}. \quad (4.6)$$

There are two cases to distinguish. If $0 < \kappa_b < K$ then $u_*(0) = \infty$ and there can be none, one, or two positive solutions to equation (4.6). If $0 < K < \kappa_b$ then $u_*(0) = 0$ and there may be one, two, or three positive roots of equation (4.6). If there are three we label them as $\tilde{x}_1 < \tilde{x}_2 < \tilde{x}_3$. The values $\tilde{x}_1, \tilde{x}_3$ will correspond to the location of maxima in u_* while $\tilde{x}_2$ will be the loca-
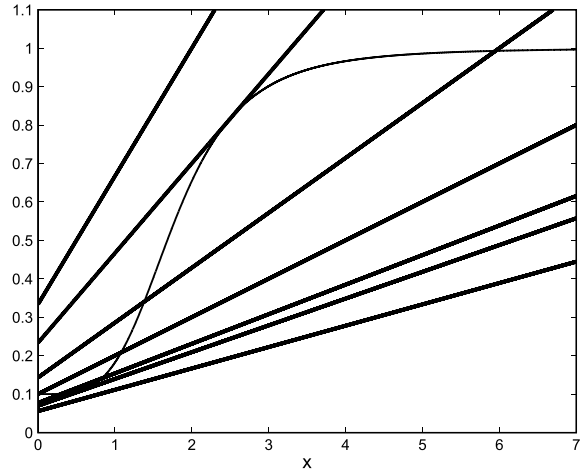

Fig. 4.2. Schematic illustration of the possibility of one, two or three solutions of equation (4.6) for varying values of κ_b with bursting inducible regulation. The straight lines correspond (in a clockwise direction) to $\kappa_b \in (0, \kappa_{b-})$, $\kappa_b = \kappa_{b-}$, $\kappa_b \in (\kappa_{b-}, \kappa_{b+})$ (and respectively $\kappa_b < K$, $\kappa_b = K$, $K < \kappa_b$), $\kappa_b = \kappa_{b+}$, and $\kappa_{b+} < \kappa_b$. This figure was constructed with $n = 4$, $K = 10$ and $b = 1$ for which $\kappa_{b-} = 4.29$ and $\kappa_{b+} = 14.35$ as computed from (4.9). See the text for further details. Taken from [9] with permission.

tion of the minimum between them. Consequently, the stationary density u_* can be classified as Unimodal type 1, type 2, as well as

Bimodal type 1 if $u_*(0) = \infty$ and u_* has a single maximum at $x > 0$, or

Bimodal type 2 if $u_*(0) = 0$ and u_* has two maxima at $\tilde{x}_1, \tilde{x}_3$, $0 < \tilde{x}_1 < \tilde{x}_3$.

There are two different bifurcation patterns that are possible. In what will be referred as **Bifurcation type 1**, the maximum at $x = 0$ disappears when there is a second peak at $x = \tilde{x}_3$. The sequence of densities encountered for increasing values of κ_b is then: Unimodal type 1 to a Bimodal type 1 to a Bimodal type 2 and finally to a Unimodal type 2 density. Figure 4.3 illustrates **Bifurcation type 1**, when $n = 4$, $K = 4$, $b = 1$, and κ_b increases from low to high values. In the **Bifurcation type 2** situation, the sequence of density types for increasing values of κ_b is: Unimodal type 1 to a Unimodal type 2 and then a Bimodal type 2 ending in a Unimodal type 2 density. Figure **??** shows **Bifurcation type 2**, when $n = 4$, $K = 4$, $b = \frac{1}{10}$, and the parameter κ_b increases.

To find the analogy between the bistable behaviour in the deterministic system and the existence of
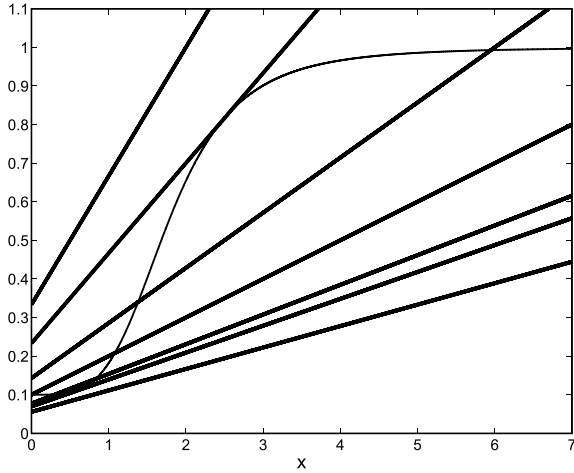
Fig. 4.3. In this figure we illustrate **Bifurcation type 1** when intrinsic bursting is present. The stationary density u_* is plotted versus x between 0 and 6. The values of the parameters used in this figure are $b = 1$, $K = 4$, and $n = 4$. The parameter κ_b was taken to be 3, 3.7, 4, 4.5, 5, 6.5, where for $\kappa_b = 3$ we have unimodal type 1 density, and increasing κ_b we obtain bimodal type 2 density for $\kappa_b = 4.5, 5$.
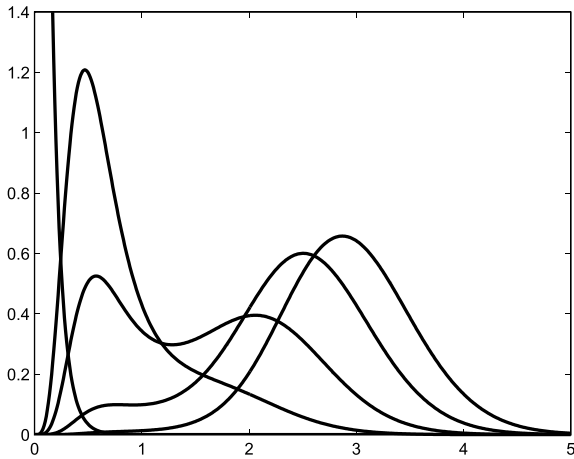

Fig. 4.4. An illustration of **Bifurcation type 2** for intrinsic bursting. The stationary density u_* is plotted versus x between 0 and 5. The parameters used are $b = 1$, $K = 4$, and $n = 4$. The parameter κ_b was taken to be 2.5, 22, 25, 28, 31, where we have Unimodal type 1 density for $\kappa_b = 2.5$, then Unimodal type 2 for $\kappa_b = 22$, Bimodal type 2 for $\kappa_b = 25, 28$, and back Unimodal type 2.

bimodal stationary density u_* we fix the parameters $b > 0$ and $K > 1$ and vary κ_b as in Fig. 4.2. In general we can cannot determine when there are three roots of (4.6). Instead, using the argument of Section 2.2.2 one can determine when there are only two roots. Differentiation of (4.6) yields the condition

$$n \frac{x^{n-1}}{(K + x^n)^2} = \frac{1}{\kappa_b b(K - 1)}. \qquad (4.7)$$

Equations (4.6) and (4.7) can be combined to give an implicit equation for the value of $x_\pm$ at which tangency will occur

$$x^{2n} - (K - 1) \left[n - \frac{K + 1}{K - 1} \right]$$
$$x^n - nb(K - 1)x^{n-1} + K = 0 \qquad (4.8)$$

and the corresponding values of $\kappa_{b\pm}$ are given by

$$\kappa_{b\pm} = \left(\frac{x_\mp + b}{b} \right) \left(\frac{K + x_\mp^n}{1 + x_\mp^n} \right). \qquad (4.9)$$

We see then that the different possibilities depend on the respective values of $K, \kappa_{b-}, \kappa_{b+}$, and κ_b. Note that it is necessary for $0 < K < \kappa_b$ in order to obtain Bimodal type 2 behaviour.

We now choose to see how the average burst size b affects bistability in the density u_* by looking at the parametric plot of $\kappa_b(x)$ versus $K(x)$. Define

$$F(x, b) = \frac{x^n + 1}{nx^{n-1}(x + b)}. \qquad (4.10)$$

Then

$$K(x, b) = \frac{1 + x^n F(x, b)}{1 - F(x, b)} \quad \text{and}$$

$$\kappa_b(x, b) = [K(x, b) + x^n] \frac{x + b}{b(x^n + 1)}. \qquad (4.11)$$

Figure 4.5 presents the regions of bimodality in the presence of bursting in the $(K, b \cdot \kappa_b)$ parameter space, which should be compared to the region of bistability in the deterministic case in the (K, κ_d) parameter space ($b\kappa_b$ is the mean number of proteins produced per unit of time, as is κ_d).

4.1.3. Recovering the deterministic case

The deterministic behaviour can be recovered from the bursting dynamics with a suitable scaling limit of parameters. The frequency κ_b and the amplitude b are two important parameters in the bursting production, while in the deterministic production there is only κ_d. Thus, if we take the limit

$$b \to 0, \quad \kappa_b \to \infty \quad \text{with} \quad b\kappa_b \equiv \kappa_d,$$

in the implicit Equations (4.5) which define the maximum points of the steady state density, then we obtain
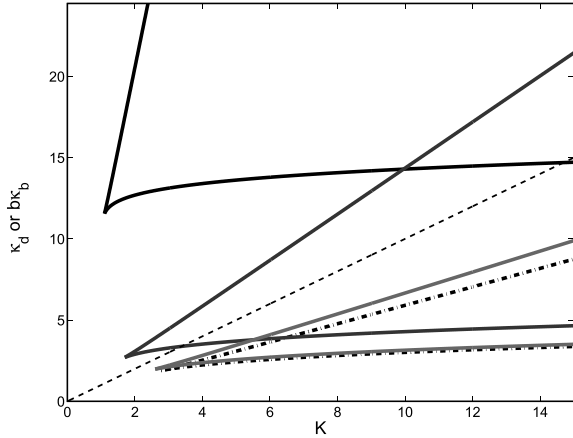
Fig. 4.5. The presence of bursting can drastically alter regions of bimodal behaviour as shown in this parametric plot (for $n = 4$) of the boundary in $(K, b \cdot \kappa_b)$ parameter space delineating unimodal from bimodal stationary densities u_* in an inducible operon with bursting and in (K, κ_d) parameter space delineating one from three steady states in the deterministic inducible operon. From top to bottom, the regions are for $b = 10$, $b = 1$, $b = 0.1$ and $b = 0.01$. The lowest (heavy dashed line) is for the deterministic case. Note that for $b = 0.1$, the two regions of bistability and bimodality coincide and are indistinguishable from one another. Taken from [9] with permission.

Equations (2.14) and (2.15) which define the stable steady states in the deterministic case.

Recovering Equation (2.17) in the limit implies that the bifurcations will also take place at the same points. Since we have $\kappa_b > K$ when $\kappa_b \to \infty$, Bimodality type 1 as well as the Unimodal type 1 behaviours will no longer be present. Moreover, the steady-state density u_* will became more sharply peaked as $b \to 0$ and the mass will be more concentrated around the larger maximum of u_*.

4.1.4. A discrete space bursting model

The number of protein molecules in a single cell can also be described as a Markov process with values in the discrete state space $\{0, 1, 2, \ldots\}$. Here we follow the approach of [103]. Let $X(t)$ be the number of gene product molecules at time t. If we have $X(t) = m$ then in a small time interval the change in the number of molecules is

$$m \xrightarrow{\lambda_m} m + k, \quad m \xrightarrow{\gamma_m} m - 1,$$

where $\gamma_m, \lambda_m, m \geq 0$, are constants satisfying

$$\lambda_0 > 0, \quad \gamma_0 = 0, \quad \gamma_m > 0, \quad \lambda_m \geq 0,$$
$$m = 1, 2, \ldots, \qquad (4.12)$$

while k is randomly chosen, independently of the actual number of molecules, according to a probability density function $h = (h_k)_{k \geq 1}$, so that $\sum_{k=1}^{+\infty} h_k = 1$, $h_k \geq 0$, $k \geq 1$. Of particular interest is the case when h is geometric

$$h_k = (1 - b)b^{k-1}, \quad k = 1, 2, \ldots, \qquad (4.13)$$

with $b \in (0, 1)$, which is the discrete space analog of the exponential distribution given by (4.1). Let $P_m(t)$ be the probability that the cell at time t has m protein molecules of the gene product. Our general master equation is an infinite set of differential equations

$$\frac{dP_m}{dt} = \gamma_{m+1}P_{m+1} - \gamma_m P_m$$
$$+ \sum_{k=1}^{m} h_k \lambda_{m-k} P_{m-k} - \lambda_m P_m,$$
$$m = 0, 1, \ldots, \qquad (4.14)$$

where we use the convention that $\sum_{k=1}^{0} = 0$. We supplement (4.14) with the initial condition $P_m(0) = v_m$, $m = 0, 1, \ldots,$ where $v = (v_m)_{m \geq 0}$ is a probability density function of the initial amount of the gene product.

The equation for the steady state $p^* = (p_m^*)_{m \geq 0}$ of (4.14) is of the form

$$\gamma_{m+1}p_{m+1}^* - \gamma_m p_m^*$$
$$+ \sum_{k=1}^{m} h_k \lambda_{m-k} p_{m-k}^* - \lambda_m p_m^* = 0,$$
$$m = 0, 1, \ldots,$$

which is uniquely solvable (up to a multiplicative constant) by

$$p_{m+1}^* = \frac{1}{\gamma_{m+1}} \sum_{k=0}^{m} \overline{h}_{m-k} \lambda_k p_k^*, \quad m = 0, 1, \ldots, \qquad (4.15)$$

where

$$\overline{h}_l = \sum_{j=l+1}^{\infty} h_j, \quad l \geq 0.$$

We have the following general result.

Theorem 4.2. *[103, Theorem 3.1] Assume condition (4.12) and suppose that a strictly positive $p^* = (p_m^*)_{m \geq 0}$ given by (4.15) satisfies*

$$\sum_{m=0}^{\infty} p_m^* = 1 \quad \text{and} \quad \sum_{m=0}^{\infty} (\lambda_m + \gamma_m) p_m^* < \infty.$$

Then for each initial probability density function Equation (4.14) has a unique solution and

$$\lim_{t \to \infty} \sum_{m=0}^{\infty} |P_m(t) - p_m^*| = 0.$$

In particular, if condition (4.12) holds and h is geometric as in (4.13) then $p^* = (p_m^*)_{m \geq 0}$ as in (4.15) is given by

$$p_m^* = \frac{p_0^* \lambda_0}{\gamma_m} \prod_{k=1}^{m-1} \frac{\lambda_k + b\gamma_k}{\gamma_k}, \quad m = 1, 2, \ldots. \quad (4.16)$$

If additionally $\gamma_m = \gamma m, m \geq 1$, with $\gamma > 0$ and λ_m is a Hill function of the form

$$\lambda_m = \lambda \frac{1 + \Theta m^n}{\Lambda + \Delta m^n}, \quad (4.17)$$

where $\Lambda, \Delta, n > 0$ and $\Theta \geq 0$ are constants, then all assumptions of Theorem 4.2 are satisfied, implying that the steady-state density $p^* = (p_m^*)_{m \geq 0}$ given by (4.16) is the discrete state space analog of (4.4).

4.2. Distributions with fluctuations in the degradation rate

We now examine the situation in which fluctuations appear in the degradation rate γ of the generic Equation (2.31). If the fluctuations are Gaussian distributed then it follows from standard chemical kinetic arguments [104] that the mean numbers of molecules decaying in a time dt is simply $\gamma x dt$ and the standard deviation of these numbers is proportional to $\sqrt{x}$. Consequently, we replace Equation (2.31) with a stochastic differential equation in the form

$$dx = \gamma[\kappa_d f(x) - x]dt + \sigma \sqrt{x} dw,$$

where w is a standard Brownian motion and we use the Ito interpretation of the stochastic integral. The corresponding Fokker Planck equation for the evolution of the ensemble density $u(t, x)$ is given by [105]

$$\frac{\partial u(t, x)}{\partial t} = -\frac{\partial \left[(\gamma \kappa_d f(x) - \gamma x) u(t, x) \right]}{\partial x}$$
$$+ \frac{\sigma^2}{2} \frac{\partial^2 (x u(t, x))}{\partial x^2}. \quad (4.18)$$

Since concentrations of molecules cannot become negative the boundary at $x = 0$ is reflecting and the stationary solution of Equation (4.18) is given by

$$u_*(x) = \frac{\mathcal{C}}{x} e^{-2\gamma x/\sigma^2} \exp \left[\frac{2\gamma \kappa_d}{\sigma^2} \int^x \frac{f(y)}{y} dy \right].$$

Set $\kappa_e = 2\gamma \kappa_d / \sigma^2$. Then the stationary density is given explicitly by

$$u_*(x) = \mathcal{C} e^{-2\gamma x/\sigma^2} x^{\kappa_e \Lambda^{-1} - 1} [\Lambda + \Delta x^n]^\theta, \quad (4.19)$$

where $\Lambda, \Delta \geq 0$ and θ are given in Table 2.1. It follows from [106, Theorem 2] that the unique stationary density of Equation (4.18) is given by Equation (4.19) and that $u(t, x)$ is asymptotically stable.

The form of the stationary density for the situation with bursting (intrinsic noise) and extrinsic noise are identical, provided that one replaces the average burst amplitude b with $b \to \sigma^2/2\gamma \equiv b_w$ and $\kappa_b \to \kappa_e = 2\gamma \kappa_d / \sigma^2 \equiv \kappa_d / b_w$. Consequently, all of the results of the previous section can be carried over here. In particular, the regions of bimodality in the (K, κ_d)-plane can be identified for a fixed value of b_w. We have the implicit equation for $x_\pm$

$$x^{2n} - (K - 1) \left[n - \frac{K + 1}{K - 1} \right] x^n$$
$$- n b_w (K - 1) x^{n-1} + K = 0$$

and the corresponding values of κ_d are given by

$$\kappa_{d\pm} = (x_\mp + b_w) \left(\frac{K + x_\mp^n}{1 + x_\mp^n} \right).$$

Then the bimodality region in the (K, κ_d)-plane with noise in the degradation rate is the same as the bimodality region for bursting in the $(K, b\kappa_b)$-plane.

Finally, we can recover the deterministic behaviour from a limit in the extrinsic fluctuations dynamics. In this case, however, the frequency and the amplitude of the perturbation are already scaled. Then the limit $\sigma \to 0$ gives the same result as in the deterministic case.

5. Discussion and conclusions

Here we have attempted to give an overview of the mathematical techniques that have been used to gain understanding about the operation of bacterial operons. We have looked at generic deterministic models

in a very general sense followed by more realistic considerations of both the lactose and tryptophan operons. These two examples are ones for which we have, arguably, the most extensive knowledge of the underlying biology as well as good data and if we cannot successfully understand their operation from a modeling perspective then there is little hope for more complicated situations. Finally we have discussed very recent results related to the role that noise (either extrinsic or intrinsic) may play in the steady state characteristics of a bacterial population. We have *not* dealt with the use of simulation techniques *per se* in the study of these systems as they fall far from our purpose and are a subject of study in their own right.

Acknowledgments

References

[1] F. Jacob, D. Perrin, C. Sanchez and J. Monod, Operon: A group of genes with the expression coordinated by an operator, *C R Hebd Seances Acad Sci* **250** (1960), 1727–1729.

[2] B.C. Goodwin, Oscillatory behaviour in enzymatic control process, *Adv Enzyme Regul* **3** (1965), 425–438.

[3] B.C. Goodwin, *Temporal organization in cells*, Academic Press, London-New York, 1963.

[4] J. Griffith, Mathematics of cellular control processes. I. Negative feedback to one gene, *J Theor Biol* **20** (1968), 202–208.

[5] J. Griffith, Mathematics of cellular control processes. II. Positive feedback to one gene, *J Theor Biol* **20** (1968), 209–216.

[6] J. Tyson and H. Othmer, The dynamics of feedback control circuits in biochemical pathways, *in: Progress in Biophysics*, R. Rosen, ed., Academic Press, New York, vol. 5, 1978, pp. 1–62.

[7] H. Othmer, The qualitative dynamics of a class of biochemical control circuits, *J Math Biol* **3** (1976), 53–78.

[8] J. Selgrade, Mathematical analysis of a cellular control process with positive feedback, *SIAM J Appl Math* **36** (1979), 219–229.

[9] M.C. Mackey, M. Tyran-Kamińska and R. Yvinec, Molecular distributions in gene regulatory dynamics, *J Theor Biol* **274** (2011), 84–96.

[10] H. Smith, *Monotone Dynamical Systems, Mathematical Surveys and Monographs*, vol. 41, American Mathematical Society, Providence, RI, 1995.

[11] L. Mier-y-Teran-Romero, M. Silber and V. Hatzimanikatis, The origins of time-delay in template biopolymerization processes, *PLOS Comp Biol* **6** (2010), e1000726–1–15.

[12] R. Heinrich and T.A. Rapoport, Mathematical modeling of translation of mRNA in eucaryotes: Steady states, time-dependent processes and application to reticulocytes, *J Theor Biol* **86** (1980), 279–313.

[13] S. Ruan and J. Wei, On the zeros of a third degree exponential polynomial with applications to a delayed model for the control of testosterone secretion, *IMA J Math Appl Med Biol* **18** (2001), 41–52.

[14] H. Haken, *Synergetics: An introduction, Springer Series in Synergetics*, vol. 1, Springer-Verlag, Berlin, 1983.

[15] R.L. Stratonovich, *Topics in the theory of random noise. Vol. I: General theory of random processes. Nonlinear transformations of signals and noise*, Revised English edition. Translated from the Russian by Richard A. Silverman, Gordon and Breach Science Publishers, New York, 1963.

[16] G. Wilemski, On the derivation of Smoluchowski equations with corrections in the classical theory of Brownian motion, *J Stat Phys* **14** (1976), 153–169.

[17] U. Titular, A systematic solution procedure for the Fokker-Planck equation of a Brownian particle in the high-friction case, *Physica A* **91A** (1978), 321–344.

[18] C. Gardiner, *Handbook of Stochastic Methods*, Springer Verlag, Berlin, Heidelberg, 1983.

[19] R. Yvinec, C. Zhuge, J. Lei and M.C. Mackey, Adiabatic reduction of a model of stochastic gene expression with jump Markov process, *J Math Biol* **68** (2014), 1051–1070.

[20] J. Monod, *Recherches sur la croissance des cultures bactériennes*, Ph.D. thesis, Université de Paris, Paris, 1941.

[21] U. Lendenmann, M. Snozzi and T. Egli, Kinetics of the simultaneous utilization of sugar mixtures by *Escherichia coli* in continuous culture, *Appl Environ Microbiol* **62** (1996), 1493–1499.

[22] I.G. Serebriiskii and E.A. Golemis, Uses of lacz to study gene function: Evaluation of β-galactosidase assays employed in the yeast twohybrid system, *Anal Biochem* **285** (2000), 1–15.

[23] J. Abramson, S. Iwata and H.R. Kaback, Lactose permease as a paradigm for membrane transport proteins (review), *Mol Membr Biol* **21** (2004), 227–236.

[24] H.R. Kaback, Structure and mechanism of the lactose permease, *C R Biol* **328** (2005), 557–567.

[25] A. Novick and M. Weiner, Enzyme induction as an all-or-none phenomenon, *Proc Natl Acad Sci U S A* **43** (1957), 553–566.

[26] M. Cohn and K. Horibata, Analysis of the differentiation and of the heterogeneity within a population of *Escherichia coli* undergoing induced beta-galactosidase synthesis, *J Bacteriol* **78** (1959), 613–623.

[27] A. Babloyantz and M. Sanglier, Chemical instabilities of "all-or-none" type in beta - galactosidase induction and active transport, *FEBS Lett* **23** (1972), 364–366.

[28] G. Nicolis and I. Prigogine, *Self-organization in nonequilibrium systems. From dissipative structures to order through fluctuations*, John Wiley and Sons, New York, 1977.

[29] O. Díaz-Hernández and M. Santillán, Bistable behavior of the lacoperon in *E. coli* when induced with a mixture of lactose and TMG, *Front Physiol* **1**.

[30] N. Yildirim and M.C. Mackey, Feedback regulation in the lactose operon: A mathematical modeling study and comparison with experimental data, *Biophys J* **84** (2003), 2841–2851.

[31] N. Yildirim, M. Santillán, D. Horike and M.C. Mackey, Dynamics and bistability in a reduced model of the lac operon, *Chaos* **14** (2004), 279–292.

[32] M. Santillán and M.C. Mackey, Inuence of catabolite repression and inducer exclusion on the bistable behavior of the lac operon, *Biophys J* **86** (2004), 1282–1292.

[33] M. Santillán, M.C. Mackey and E.S. Zeron, Origin of bistability in the lac operon, *Biophys J* **92** (2007), 3830–3842.

[34] M. Santillán, Bistable behavior in a model of the lac operon in *Escherichia coli* with variable growth rate, *Biophys J* **9** (2008), 2065–2081.

[35] M. Santillán and M.C. Mackey, Quantitative approaches to the study of bistability in the lac operon of *Escherichia coli*, *J R Soc Interface* **5** (2008), S29S39.

[36] W.S. Reznikoff, The lactose operon-controlling elements: A complex paradigm, *Mol Microbiol* **6** (1992), 2419–2422.

[37] B. Müller-Hill, The function of auxilaty operators, *Mol Microbiol* **29** (1998), 13–18.

[38] M. Lewis, The lac repressor, *C R Biol* **328** (2005), 521–548.

[39] C.J. Wilson, H. Zhan, L. Swint-Kruse and K.S. Matthews, The lactose repressor system, paradigms for regulation, allosteric behavior and protein folding, *Cell Mol Life Sci* **64** (2007), 3–16.

[40] A. Narang, Effect of DNA looping on the induction kinetics of the lacoperon, *J Theor Biol* **247** (2007), 695–712.

[41] S. Oehler, E.R. Eismann, H. Krämer and B. Müller-Hill, The three operators of lac operon cooperate in repression, *EMBO J* **9** (1990), 973–979.

[42] T. Kuhlman, Z. Zhang, M.H. Saier and T. Hwa, Combinatorial transcriptional control of the lactose operon of *Escherichia coli*, *Proc Natl Acad Sci U S A* **104** (2007), 6043–6048.

[43] J. Beckwith, The lactose operon, in: *Escherichia coli and Salmonella thyphymurium: Cellular and molecular biology*, F. Neidhart, J. Ingraham, K. Low, B. Magasanik and H. Umbarger, eds., American Society for Microbiology, Washington, DC, vol. 2, 1987, pp. 1439–1443.

[44] E. Ozbudak, M. Thattai, H. Lim, B. Shraiman and A. van Oudenaarden, Multistability in the lactose utilization network of *Escherichia coli*, *Nature* **427** (2004), 737–740.

[45] M. Santillán and E.S. Zeron, Dynamic inuence of feedback enzyme inhibition and transcription attenuation on the tryptophan operon response to nutritional shifts, *J Theor Biol* **231** (2004), 287–298.

[46] M. Santillán, On the use of the Hill functions in mathematical models of gene regulatory networks, *Math Model Nat Phenom* **3** (2008), 85–97.

[47] J.J. Tyson, On the existence of oscillatory solutions in negative feedback cellular control processes, *J Theor Biol* **1** (1975), 311–315.

[48] R.D. Bliss, P.R. Painter and A.G. Marr, Role of feedback inhibition in stabilizing the classical operon, *J Theor Biol* **97** (1982), 177–193.

[49] S. Sinha, Complex behaviour of the repressible operon, *J Theor Biol* **132** (1988), 307–318.

[50] A.K. Sen and W.-M. Liu, Dynamic analysis of genetic control and regulation of amino acid synthesis: The tryptophan operon in *Escherichia coli*, *Biotechnol Bioeng* **35** (1990), 185–194.

[51] M. Giona and A. Adrover, Modified model for the regulation of the tryptophan operon in *Escherichia coli*, *Biotechnol Bioeng* **80** (2002), 297–304.

[52] M. Santillán and M.C. Mackey, Dynamic regulation of the tryptophan operon: Modeling study and comparison with experimental data, *Proc Natl Acad Sci U S A* **98** (2001), 1364–1369.

[53] M. Santillán and E.S. Zeron, Analytical study of the multiplicity of regulatory mechanisms in the tryptophan operon, *Bull Math Biol* **68** (2006), 343–359.

[54] A. Hernández-Valdez and M. Santillán, Cyclic expression and cooperative operator interaction in the trp operon of *Escherichia coli*, *J Theor Biol* **263** (2010), 340–352.

[55] E. Salazar-Cavazos and M. Santillán, Optimal performance of the tryptophan operon of *E. coli*: A stochastic, dynamical, mathematical modeling approach, *Bull Math Biol* **76** (2013), 314–334.

[56] C. Yanofsky and V. Horn, Role of regulatory features of the trp operon of *Escherichia coli* in mediating a response to a nutritional shift, *J Bacteriol* **176** (1994), 6245–6254.

[57] C. Zamora-Chimal, M. Santillán and J. Rodríguez-González, Inuence of the feedback loops in the trp operon of B. subtilis on the system dynamic response and noise amplitude, *J Theor Biol* **310** (2012), 119–131.

[58] M. Kaern, T. Elston, W. Blake and J. Collins, Stochasticity in gene expression: From theories to phenotypes, *Nat Rev Genet* **6** (2005), 451–464.

[59] A. Raj and A. van Oudenaarden, Nature, nurture, or chance: Stochastic gene expression and its consequences, *Cell* **135** (2008), 216–226.

[60] V. Shahrezaei and P. Swain, The stochastic nature of biochemical networks, *Curr Opin Biotechnol* **19** (2008), 369–374.

[61] V. Shahrezaei, J. Ollivier and P. Swain, Colored extrinsic uctuations and stochastic gene expression, *Mol Syst Biol* **4** (2008), 196–205.

[62] A. Ochab-Marcinek, Predicting the asymmetric response of a genetic switch to noise, *J Theor Biol* **254** (2008), 37–44.

[63] A. Ochab-Marcinek, Extrinsic noise passing through a Michaelis-Menten reaction: A universal response of a genetic switch, *J Theor Biol* **263** (2010), 510–520.

[64] A. Ochab-Marcinek and M. Tabaka, Bimodal gene expression in noncooperative regulatory systems, *Proc Natl Acad Sci U S A* **107** (2010), 22096–22101.

[65] M. Elowitz, A. Levine, E. Siggia and P. Swain, Stochastic gene expression in a single cell, *Science* **297** (2002), 1183–1186.

[66] J. Raser and E. O'Shea, Control of stochasticity in eukaryotic gene expression, *Science* **304** (2004), 1811–1814.

[67] P. Swain, M. Elowitz and E. Siggia, Intrinsic and extrinsic contributions to stochasticity in gene expression, *Proc Natl Acad Sci U S A* **99** (2002), 12795–12800.

[68] M. Scott, B. Ingallls and M. Kærn, Estimations of intrinsic and extrinsic noise in models of nonlinear genetic networks, *Chaos* **16** (2006), 026107–1–15.

[69] W. Blake, G. Balázsi, M. Kohanski, F. Issacs, K. Murphy, Y. Kuang, C. Cantor, D. Walt and J. Collins, Phenotypic consequences of promoter-mediated transcriptional noise, *Mol Cell* **24** (2006), 853–865.

[70] H. Fraser, A. Hirsh, G. Glaever, J. Kumm and M. Eisen, Noise minimization in eukaryotic gene expression, *PLoS Biol* **2** (2004), 8343–838.

[71] O.G. Berg, A model for the statistical uctuations of protein numbers in a microbial population, *J Theor Biol* **71** (1978), 587–603.

[72] W. Blake, M. Kaern, C. Cantor and J. Collins, Noise in eukaryotic gene expression, *Nature* **422** (2003), 633–637.

[73] L. Cai, N. Friedman and X. Xie, Stochastic protein expression in individual cells at the single molecule level, *Nature* **440** (2006), 358–362.

[74] J. Chubb, T. Trcek, S. Shenoy and R. Singer, Transcriptional pulsing of a developmental gene, *Curr Biol* **16** (2006), 1018–1025.

[75] I. Golding, J. Paulsson, S. Zawilski and E. Cox, Real-time kinetics of gene activity in individual bacteria, *Cell* **123** (2005), 1025–1036.

[76] A. Raj, C. Peskin, D. Tranchina, D. Vargas and S. Tyagi, Stochastic mRNA synthesis in mammalian cells, *PLoS Biol* **4** (2006), 1707–1719.

[77] A. Sigal, R. Milo, A. Cohen, N. Geva-Zatorsky, Y. Klein, Y. Liron, N. Rosenfeld, T. Danon, N. Perzov and U. Alon, Variability and memory of protein levels in human cells, *Nature* **444** (2006), 643–646.

[78] D. Suter, N. Molina, D. Gatfield, K. Schneider, U. Schibler and F. Naef, Mammalian genes are transcribed with widely different bursting kinetics, *Science* **332** (2011), 472–474.

[79] J. Yu, J. Xiao, X. Ren, K. Lao and X. Xie, Probing gene expression in live cells, one protein molecule at a time, *Science* **311** (2006), 1600–1603.

[80] T. Kepler and T. Elston, Stochasticity in transcriptional regulation: Origins, consequences, and mathematical representations, *Biophy J* **81** (2001), 3116–3136.

[81] N. Friedman, L. Cai and X. Xie, Linking stochastic dynamics to population distribution: An analytical framework of gene expression, *Phys Rev Lett* **97** (2006), 168302–1/4.

[82] L. Morelli and F. Julicher, Precision of genetic oscillators and clocks, *Phys Rev Lett* **9** (2007), 228101.

[83] A. Bobrowski, T. Lipniacki, K. Pichór and R. Rudnicki, Asymptotic behavior of distributions of mRNA and protein levels in a model of stochastic gene expression, *J Math Anal Appl* **333** (2007), 753–769.

[84] V. Shahrezaei and P. Swain, Analytic distributions for stochastic gene expression, *Proc Natl Acad Sci U S A* **105** (2008), 17256–17261.

[85] S. Iyer-Biswas, F. Hayot and C. Jayaprakash, Stochasticity of gene products from transcriptional pulsing, *Phys Rev E* **7** (2009), 031911.

[86] A. Mugler, A. Walczak and C. Wiggins, Spectral solutions to stochastic models of gene expression with bursts and regulation, *Phys Rev E* **8** (2009), 041921.

[87] A. Ribeiro, O. Smolander, T. Rajala, A. Hakkinen and O. Yli-Harja, Delayed stochastic model of transcription at the single nucleotide level, *J Comput Biol* **16** (2009), 539–53.

[88] V. Elgart, T. Jia and R. Kulkarni, Applications of Little's law to stochastic models of gene expression, *Phys Rev E* **8** (2010), 021901.

[89] J. Lei, Stochasticity in single gene expression with both intrinsic noise and uctuation in kinetic parameters, *J Theor Biol* **256** (2010), 485–492. *Erratum appears in J Theor Biol* **262**(2) (2010), 381.

[90] T. Rajala, A. Hakkinen, S. Healy, O. Yli-Harja and A. Ribeiro, Effects of transcriptional pausing on gene expression dynamics, *PLoS Comput Biol* **6** (2010), e1000704.

[91] M. Tang, The mean frequency of transcriptional bursting and its variation in single cells, *J Math Biol* **60** (2010), 27–58.

[92] G. Bett, Q. Zhou and R. Rasmusson, Models of HERG gating, *Biophys J* **101** (2011), 631–642.

[93] T. Jia and R. Kulkarni, Intrinsic noise in stochastic models of gene expression with molecular memory and bursting, *Phys Rev Lett* **10** (2011), 058102.

[94] D. Cottrell, P. Swain and P. Tupper, Stochastic branching-diffusion models for gene expression, *Proc Natl Acad Sci U S A* **109** (2012), 9699–9704.

[95] H. Feng, Z. Hensel, J. Xiao and J. Wang, Analytical calculation of protein production distributions in models of clustered protein expression, *Phys Rev E* **85** (2012), 155–160.

[96] M. Ferguson, D. Le Coq, M. Jules, S. Aymerich, O. Radulescu, N. Declerck and C. Royer, Reconciling molecular regulatory mechanisms with noise patterns of bacterial metabolic promoters in induced and repressed states, *Proc Natl Acad Sci U S A* **109** (2012), 155–160.

[97] H. Kuwahara and R. Schwartz, Stochastic steady state gain in a gene expression process with mRNA degradation control, *J R Soc Interface* **9** (2012), 1589–1598.

[98] A. Singh and P. Bokes, Consequences of mRNA transport on stochastic variability in protein levels, *Biophys J* **103** (2012), 1087–1096.

[99] T. Earnest, E. Roberts, M. Assaf, K. Dahmen and Z. Luthey-Schulten, DNA looping increases the range of bistability in a stochastic model of the lac genetic switch, *Phys Biol* **1** (2013), 026002.

[100] R. Robinson, Bursting with randomness: A simple model for stochastic control of gene expression, *PLoS Biol* **11** (2013), e1001622.

[101] T. Tian, Chemical memory reactions induced bursting dynamics in gene expression, *PLoS One* **8** (2013), e52029.

[102] M.C. Mackey and M. Tyran-Kamińska, Dynamics and density evolution in piecewise deterministic growth processes, *Ann Polon Math* **94** (2008), 111–129.

[103] M.C. Mackey, M. Tyran-Kamińska and R. Yvinec, Dynamic behavior of stochastic gene expression models in the presence of bursting, *SIAM J Appl Math* **73** (2013), 1830–1852.

[104] I. Oppenheim, K. Schuler and G. Weiss, Stochastic and deterministic formulation of chemical rate equations, *J Chem Phys* **50** (1969), 460–466.

[105] A. Lasota and M.C. Mackey, *Chaos, fractals, and noise, Applied Mathematical Sciences*, vol. 97, Springer-Verlag, New York, 1994.

[106] K. Pichór and R. Rudnicki, Continuous Markov semigroups and stability of transport equations, *J Math Anal Appl* **249** (2000), 668–685.