In Silico Biology 12 (2012-2014) 69–82
DOI 10.3233/ISB-150465
IOS Press

A simple model for Lutz and Bujard's controllable promoters and its application for analyzing a simple genetic oscillator

C.G. Zamora-Chimal[a] and E.S. Zeron[b,*]
[a]*Centro de Investigación y de Estudios Avanzados del IPN, Unidad Monterrey, Parque de Investigación e Innovación Tecnológica, Apodaca NL, México*
[b]*Centro de Investigación y de Estudios Avanzados del IPN, Departamento de Matemáticas, Av. Instituto Politécnico Nacional México DF, México*

Received 7 June 2014
Accepted 5 February 2015

Abstract. We develop an exact and flexible mathematical model for Lutz and Bujard's controllable promoters. It can be used as a building block for modeling genetic systems based on them. Special attention is paid to deduce all the model parameters from reported (*in vitro*) experimental data. We validate our model by comparing the regulatory ranges measured *in vivo* by Lutz and Bujard against the ranges predicted by the model, and which are calculated as the reporter activity obtained under inducing conditions divided by the activity measured under maximal repression. In particular, we verify Bond et al. assertion that the cooperativity between two *lac* operators can be assumed to be negligible when their central base pairs are separated by 22 or 32 bp [Gene repression by minimal *lac* loops *in vivo*, Nucleic Acids Res, **38** (2010) 8072–8082]. Moreover, we also find that the probability that two repressors *LacI* bind to these operators at the same time can be assumed to be negligible as well. We finally use the model for the promoter $P_{LlacO\text{-}1}$ to analyze a synthetic genetic oscillator recently build by Stricker et al. [A fast, robust and tunable synthetic gene oscillator, Nature, **456** (2008) 516–519].

Keywords: Synthetic promoters, *lac* and *tet* operators, biochemical oscillations

1. Introduction

A defining goal of synthetic biology is to design and build genetic systems with specific functions that resemble those easily find nowadays in the electronic industry, such as genetic switches and oscillators. One of the main problems in the area is to design and characterize reliable synthetic promoters and operators that can control the expression of a particular gene via specific repressors, activators, or inducers. Lutz and Bujard [1] published in 1997 a seminal paper announcing the development of the synthetic controllable promoters $P_{lac/ara\text{-}1}$, $P_{A1lacO\text{-}1}$, $P_{LlacO\text{-}1}$, $P_{LtetO\text{-}1}$, and $P_{N25tetO\text{-}1}$ in *E.coli*. For example, the promoters $P_{LlacO\text{-}1}$ and $P_{LtetO\text{-}1}$ have two sequences of the respective *lac* and *tet* operators (binding sites for the repressors *LacI* and *TetR*) inserted into regions adjacent to the -10 and -33 hexamers in the phage lambda promoter P_L; see the topography and sequences presented in Fig. 1. So that, $P_{LlacO\text{-}1}$ and $P_{LtetO\text{-}1}$ are respectively repressed by *LacI* and *TetR*, and so induced by IPTG and anhydrotetracycline.

*Corresponding author: Eduardo S. Zeron, Centro de Investigación y de Estudios Avanzados del IPN, Departamento de Matemáticas, Av. Instituto Politécnico Nacional 2508, 07360 México DF, México. Tel.: +52 55 5747 3800/Ext. 6429/3871; Fax: +52 55 5747 3876; E-mail: ESZeron@math.cinvestav.edu.mx.

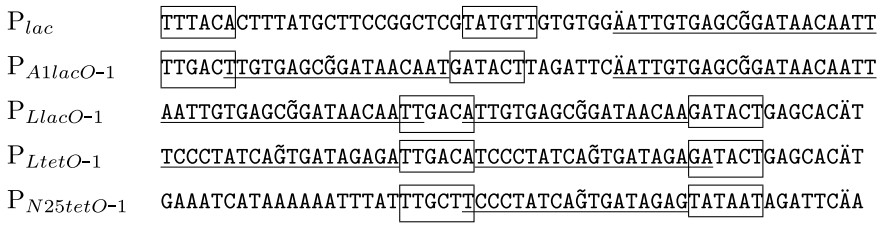

Fig. 1. Topography and sequences of the central parts of the *lac* promoter P_{lac} and four of the Lutz and Bujard promoter/operators. The -10 and -33 hexamers are boxed and the transcriptional starting site is marked with an umlaut diacritic $\ddot{\text{A}}$. The corresponding sequences of the *lac* and *tet* operators (binding sites for the repressors *LacI* and *TetR*) are all underlined; they are between 18 and 21bp long; and their central base pairs are indicated with a tilde diacritic $\tilde{\text{G}}$. The promoters P_{lac}, $P_{A1lacO\text{-}1}$, and $P_{LlacO\text{-}1}$ include the *lac* operator; while $P_{LtetO\text{-}1}$ and $P_{N25tetO\text{-}1}$ include the *tet* operator. The central base pairs of the operators in the promoters $P_{A1lacO\text{-}1}$, $P_{LlacO\text{-}1}$, and $P_{LtetO\text{-}1}$ are respectively separated by 32, 22, and 24bp.

The EXPRESSYS company [2] is nowadays commercializing pZ vector systems that contain a module with one of the five Lutz and Bujard's synthetic promoters followed by a ribosomal binding site and by either a multiple cloning site, the *lacZ* gene, or the *luciferase* gene. The pZ vector systems also have a second module with four possible origins of replication and a third module with one of five antibiotic resistance markers that carry their genuine promoters and ribosomal binding sites. The EXPRESSYS pZ vectors have been used by several authors [3–7] for *in vivo* experiments. In particular, Stricker et al. [4] used the pZ vectors for implementing *in vivo* a pair of genetic oscillators.

Considering the wide applications of Lutz and Bujard's controllable promoters, it is important to explain in detail how they function; i.e., how polymerases and repressors are competing for occupying their respective binding sites in the promoter. In particular, it is important to explain why the *in vivo* regulatory ranges measured by Lutz and Bujard have the values reported in Table 1 of [1], because there are some apparent contradictions with several *in vivo* measurements done by Oehler et al. [8, 9] on the *lac* promoter. Recall that regulatory ranges are calculated as the reporter activity measured under inducing conditions divided by the activity measured under maximal repression. For example, the promoters $P_{LlacO\text{-}1}$ and $P_{A1lacO\text{-}1}$ both contain a pair of *lac* operators and their *in vivo* regulatory ranges are 620 and 350, respectively. These values are at least one order of magnitude larger than the *in vivo* regulatory range ($= 18$) reported by Oehler, Eismann et al. in 1990 for a *lac* promoter with only one functional *lac* operator next to the transcriptional starting site; see configuration $\lambda Ewt100$ in Fig. 2 of [8]. Moreover, the regulatory ranges 620 and 350 are at least one order of magnitude smaller than the regulatory range ($\geq 12 \times 10^3$) reported by Oehler, Amouyal et al. in 1994 for a synthetic *lac* promoter that includes two functional *lac* operators inside its structure; see Fig. 5 of [9].

In this work we develop an exact and flexible mathematical model for Lutz and Bujard's controllable promoters P_{lac}, $P_{A1lacO\text{-}1}$, $P_{LlacO\text{-}1}$, and $P_{LtetO\text{-}1}$. Special attention is paid to deduce all the model parameters from reported (*in vitro*) experimental data. We propose a simple model (written in terms of chemical reactions) that explains how the polymerases and repressors compete for occupying their respective binding sites in the promoter. We also use this model for verifying Bond et al. assertion that the cooperativity between two *lac* operators can be assumed to be negligible when their central base pairs are separated by 22 or 32 bp [10]. Recall that two *lac* operators cooperate to produce a larger repression when a tetrameric repressor *LacI* binds simultaneously to them creating a repression loop [11], so that the cooperativity is proportional to the probability that a single repressor binds to two operators. Bond et al. and Müller et al. [10, 12] have found that, for *lac* operators whose central base pairs are separated by a distance between 49 and 98bp, the maxima and minima in the cooperativity are periodically distributed; i.e., the cooperativity is maximal when the separation between central base pairs is 59, 70, 81, or 92bp; while it is minimal (negligible) when the separation is around 52, 63, 74, 85, or 96bp. Moreover, Bond et al. also suggested that the non-coplanarity of the repressor *LacI* [11] makes repression loops impossible for distances shorter than 50bp, unless there are dramatic changes in the protein–DNA interface.

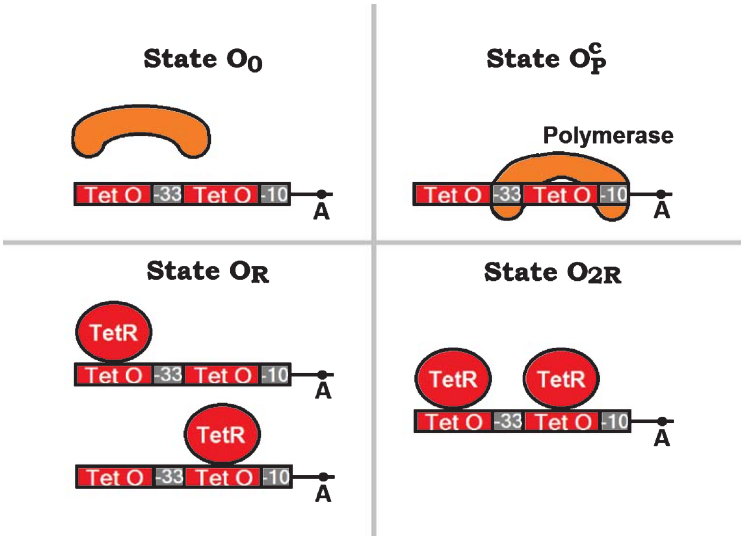

Fig. 2. Graphic representation of the different states O_0, O_P^c, O_R and O_{2R} assumed by the promoter $P_{LtetO\text{-}1}$ in presence of the polymerase and the (dimeric) tetracycline repressor *TetR*. The transcriptional starting site is marked with the letter A, the -10 and -33 hexamers are indicated by small grey rectangles, the polymerase is represented by an orange kidney shaped figure, the *tet* operators are indicated by narrow red rectangles, and the repressors *TetR* are represented by big red ovals.

Finally, once we have a flexible mathematical model for Lutz and Bujard's controllable promoters, we use it for analyzing the simplest synthetic genetic oscillator implemented *in vivo* by Stricker et al. in 2008 [4]. We analyze this oscillator because of its simplicity, it has only one feedback loop, so that the existence and frequency of oscillations are quite sensible to variations in the parameters. We find that the period of the simulated oscillations coincides with the period of the oscillations ($\approx 35min$) measured *in vivo* by Stricker et al. [4], so that we can take this coincidence as a second confirmation that our model for Lutz and Bujard's controllable promoters is correct.

2. Methods and models

Consider the topography and sequences of the central parts of the *lac* promoter P_{lac} and four of the Lutz and Bujard promoter/operators presented in Fig. 1. The *lac* and *tet* operators are located besides (downstream) of the transcriptional starting site, besides (upstream) of the -33 hexamer, or between the -10 and -33 hexamers. Since the -10 and -33 hexamers are a principal landing mark for the polymerases in order to get attached to the promoter and to initiate the transcription, we use Chamberlin proposal that polymerases and repressors physically exclude each other from the promoter, and so we model this competition defining four basic states for the promoter: O_0 represents the state when the promoter is free and ready to be occupied by a polymerase or a repressor; O_P^c represents the state when a polymerase P is bound to the promoter but the DNA is still *closed*; O_R represents the state when one repressor R is bound to the respective operator (or to any one of the operators if there are two of them); O_{2R} represents the state when there are two repressors bound to two operators. Notice that the state O_{2R} only exists for the promoters $P_{A1lacO\text{-}1}$, $P_{LlacO\text{-}1}$, and $P_{LtetO\text{-}1}$, because only them contain two operators. For example, the different states found in promoter $P_{LtetO\text{-}1}$ are graphically represented in Fig. 2. The transitions between the states O_0, O_P^c and O_R are represented as chemical reactions:

$$O_0 + R \underset{k_a}{\overset{nk_A}{\rightleftharpoons}} O_R, \quad O_R + R \overset{K_2}{\rightleftharpoons} O_{2R},$$

$$O_0 + P \underset{k_b}{\overset{k_B}{\rightleftharpoons}} O_P^c, \quad O_P^c \xrightarrow{k_m} O_0 + P + M, \qquad (1)$$

where M represents a nascent mRNA molecule, the constant k_m is the promoter transcription initiation rate, and the parameters k_A, k_a, k_B, and k_b are the respective kinetic association and dissociation rate constants for the complexes repressor/operator O_R and polymerase/promoter O_P^c. The parameter n denotes the number of active operators located in the core (central)

part of the promoter and K_2 is the association constant for the complex (2 repressors)/(2 operators) O_{2R}, so that $n = 1$ and $K_2 = 0$ for the promoters P_{lac} and $P_{N25tetO\text{-}1}$; while $n = 2$ for the other three promoters $P_{A1lacO\text{-}1}$, $P_{LlacO\text{-}1}$, and $P_{LtetO\text{-}1}$ presented in Fig. 1. Obviously, if there is only $n = 1$ operator in the promoter, then the state O_{2R} does not exist and we may fix $K_2 = 0$. Moreover, if there are $n = 2$ operators in the promoter, the resulting kinetic association constant nk_A must be twice as large as the original k_A, because the repressor has now two options (operators) to get bound to. The last reaction in (1) represents the transcription process, minus the initiation part; i.e.: promoter clearance, elongation, and termination. We must include it because there is a transition involved from the *closed* state O_P^c into the free one O_0. The reaction begins with a polymerase P recently bound to the operator (*closed* state O_P^c), the polymerase then opens the DNA strands, frees the promoter site (yielding the state O_0), transcribes a nascent mRNA molecule M in the elongation process, and at the end detaches itself from the DNA.

We could not find experimental data on the value of K_2, and so we must use indirect information. Hillen et al. [13] have made several experiments weighting fragments of DNA that contain one or two *tet* operators, and in presence or absence of the repressor *TetR*. In Fig. 2 of their paper, Hillen et al. indicate that a fragment of DNA containing only one *tet* operator weights the equivalent of 95 base-pairs more when it is saturated with repressors *TetR* than when there are no repressors at all. The same experiment was repeated with a fragment of DNA that contains two *tet* operators. The central base pairs of these operators are separated by 29bp. Hillen et al. indicate that the fragment with two operators weights the equivalent of 174 base-pairs more when it is saturated with repressors *TetR* than when there are no repressors at all. Since 174 is approximately equal to 1.83 times 95, we can interpret the latter results as follows: When the fragment of DNA with two *tet* operators is saturated with repressors, the probability that it is actually bound by two repressors is 83%, while the probability that it is bound by only one repressor is 17%. Finally, since the central base pairs of the *tet* operators in the promoter $P_{LtetO\text{-}1}$ are separated by only 24bp, instead of 29bp, we make the following assumption: When the promoter $P_{LtetO\text{-}1}$ is mixed with repressors *TetR*, the probability that it is in the state O_{2R} is 75%, while the probability that it is in the state O_R is 25%.

We make a different assumption for the promoters $P_{A1lacO\text{-}1}$ and $P_{LlacO\text{-}1}$, which include a pair of *lac* operators inside their structure. The central base pairs of the operators are separated by less than 32.5bp according to Fig. 1. Since the tetrameric repressor *LacI* has a molecular mass (= 154.52 kDa [11]) that is at least three times larger than the mass of the dimeric repressor *TetR* (= 46.71 kDa [14]), we shall assume that the state O_{2R} never happens for the promoters $P_{A1lacO\text{-}1}$ and $P_{LlacO\text{-}1}$; i.e., once a repressor *LacI* binds to one *lac* operator, it physically prevents a second repressor getting bound to the adjacent operator. The consistency of the previous assumptions is validated by comparing the regulatory ranges predicted by the resulting model against those measured *in vivo* by Lutz and Bujard. We rewrite below these assumptions to simplify the presentation.

Assumption 1. *Let $P(O_R)$ and $P(O_{2R})$ be the probability that any one of the promoters $P_{A1lacO\text{-}1}$, $P_{LlacO\text{-}1}$, or $P_{LtetO\text{-}1}$ is at the state O_R and O_{2R}, respectively. We suppose that $P(O_{2R})$ is equal to $\beta P(O_R)$ in the presence of repressors, where $\beta = 3$ for the promoter $P_{LtetO\text{-}1}$ and $\beta = 0$ for all the other promoters.*

This assumption means in particular that the probability $P(O_{2R})$ is identically equal to zero for all the promoters analyzed in this paper, with the only exception of the promoter $P_{LtetO\text{-}1}$. On the other hand, as we have said in the introduction, the tetrameric repressor *LacI* is able to bind simultaneously to two operators creating a repression loop. Hence, we should take into consideration the possibility that the *lac* operators inside the promoters $P_{A1lacO\text{-}1}$ and $P_{LlacO\text{-}1}$ could cooperate to produce a larger repression. Bond et al. created *in vivo* a minimal system in *E. coli* for testing the shortest possible DNA repression loop between two *lac* promoters [10]. The authors suggested in particular that the non-coplanarity of the repressor *LacI* [11] makes repression loops impossible when the central base pairs of the promoters are separated by distances shorter than 50bp, unless there are dramatic changes in the protein–DNA interface. Since the central base pairs of the *lac* operators are separated by only 32bp (or 22bp) in the promoter $P_{A1lacO\text{-}1}$ (or $P_{LlacO\text{-}1}$), we made the following assumption.

Assumption 2. *The cooperativity is negligible between the lac operators in the promoters $P_{A1lacO\text{-}1}$ and $P_{LlacO\text{-}1}$. I.e.; the probability that one repressor LacI*

binds to two different operators is negligible when their central base pairs are separated by 22 or 32bp.

The assumptions made in the previous paragraphs automatically imply that we only need to consider the following states for the promoter: O_0, O_P^c, O_R, and O_{2R}. The last state O_{2R} is only needed for the promoter $P_{LtetO\text{-}1}$. We must introduce some notation at this point. Let $P(O_0)$, $P(O_P^c)$, and $P(O_R)$ be the probability that the promoter is free (state O_0), it forms the complex polymerase/promoter O_P^c, and it forms the complex repressor/operator O_R, respectively. Moreover, we also need the probability $P(O_{2R})$ that the promoter is at the state O_{2R}. The state of equilibrium for the reactions (1) is easily given by the following identities,

$$P(O_R) = \frac{nk_A[R]}{k_a} P(O_0) \quad \text{and}$$

$$P(O_P^c) = \frac{k_B[P]}{k_b + k_m} P(O_0). \tag{2}$$

Since $1 = P(O_P^c)+P(O_R)+P(O_{2R})+P(O_0)$, we can easily deduce the following equation after dividing by $P(O_P^c)$ and applying the above identities and the considerations made in Assumption 1,

$$\frac{1}{P(O_P^c)} = 1 + \frac{nk_A}{k_B}\left(\frac{k_b+k_m}{k_a}\right)\frac{[R]}{[P]}(1+\beta) + \left(\frac{k_b+k_m}{k_B[P]}\right). \tag{3}$$

We can easily verify that the last term $\frac{k_b+k_m}{k_B[P]}$ is really small ($< 10^{-3}$) with respect to the number 1, when we substitute the value of the parameters k_b, k_m, k_B, and $[P]$ enlisted in (13), (14), (15), or (16). Notice that $1M$ is equal to 1.8×10^9 mpb (molecule count per average sized bacterium) according to the conversion calculated at the beginning of Appendix E. Whence, we can eliminate the last term in equation (3) without altering it, and so the probability that the promoter forms the complex polymerase/promoter O_P^c is given by the formula

$$P(O_P^c) = \frac{[P]}{[P] + m[R]} \qquad \text{for}$$

$$m := n(1+\beta)\frac{k_A}{k_B}\left(\frac{k_b+k_m}{k_a}\right). \tag{4}$$

Finally, since the regulatory ranges are calculated as the reporter activity measured under inducing conditions divided by the activity measured under maximal repression, we can also calculate them as the probability $P(O_P^c)$ obtained when $[R] = 0$ divided by the probability obtained when the concentration $[R]$ is maximal; i.e.,

$$\text{Regulatory Range} = 1 + n(1+\beta)\frac{k_A}{k_B}\left(\frac{k_b+k_m}{k_a}\right)\frac{[R]}{[P]}. \tag{5}$$

It is known that the transcription process is a very complex one, which involves many consecutive chemical reactions; see for example the illustrative paper by deHaset et al. [15]. Nevertheless, even when the system of reactions (1) is a very simplistic way to represent the repression/transcription process, we assert that it is an effective and flexible model that can be used as a building block for modeling the genetic systems based on Lutz and Bujard's controllable promoters. We compare in Section 3 the regulatory ranges measured *in vivo* for the promoters P_{lac}, $P_{A1lacO\text{-}1}$, $P_{LlacO\text{-}1}$, and $P_{LtetO\text{-}1}$ against those ranges obtained from equation (5) by substituting the value of the corresponding chemical kinetic constants measured *in vitro* by several authors. Since we obtain a good agreement between the measured and calculated ranges (the error is less than 7.7% in all cases), we can conclude that the system of reactions (1) and Assumptions 1 and 2 are consistent with the experimental data.

2.1. Model for Stricker et al. genetic oscillator

As we have said in the previous paragraph, we obtain a good agreement between the measured and calculated regulatory ranges. However, since the regulatory ranges are calculated with data measured in steady state, we also want to verify that our model can make accurate predictions in a fully dynamical situation. To do so, we build a model for a synthetic genetic oscillator implemented *in vivo* by Stricker et al. [4]. The authors actually implemented two synthetic genetic oscillators using pZ vectors commercialized by the EXPRESSYS company. The topology of one oscillator has two feedback loops with different sign (that compete one against each other), while the topology of the second one is simpler and has only a single negative feedback loop [4]. We analyze the simpler oscillator, because its own simplicity means that the frequency and existence of oscillations are quite sensible to variations in the parameters. Elowitz and Leibler have already implemented *in vivo* a version of the repressilator (genetic oscillator with only one feedback loop) and observed that small variations in the parameters produce large variations in the dynamics [3]. Hence, an agreement between the period of the oscillations measured *in vivo* by Stricker et al. and the period calculated in silico with our model would

be a strong confirmation that the system of reactions (1) can be used as a building block for modeling the genetic systems based on Lutz and Bujard controllable promoters.

Stricker et al. implemented *in vivo* their simplest genetic oscillator by introducing two different EXPRESSYS pZ vectors into a strain of *E.coli* (denoted JS013). The first pZ vector contains a copy of the *lacI* structural genes controlled by the promoter $P_{LlacO\text{-}1}$. Since the promoter is repressed by $LacI$ and the *lacI* genes are transcribed and translated into the same repressor $LacI$, the production of $LacI$ forms a simple negative feedback loop by repressing its own expression. The second pZ vector contains a $yemGFP$ gene controlled by the promoter $P_{LlacO\text{-}1}$, so that it works as a reporter for the activity of $LacI$. We begin our model by considering the system of reactions (1) that describe the dynamics (repression/transcription) of the $P_{LlacO\text{-}1}$ operator,

$$O_0 + R \underset{k_a}{\overset{2k_A}{\rightleftharpoons}} O_R, \quad O_0 + P \underset{k_b}{\overset{k_B}{\rightleftharpoons}} O_P^c,$$
$$O_P^c \xrightarrow{k_m} O_0 + P + M, \tag{6}$$

We do not include the state O_{2R}, because it is only needed for the promoter $P_{LtetO\text{-}1}$. The value of the parameters k_a, k_A, k_b, k_B, and k_m is deduced in Appendix C. We assume that the transcription is an instantaneous reaction, because the translation of the nascent mRNA molecule (M) can begin when the first two segments of the leader region have been transcribed and a ribosome can bind to them [11]. We add to our model the reactions corresponding to the translation of the lactose repressor monomer E; the dimerization of E into a functional molecule D that can bind to the *lac* operator; and the tetramerization into a complete (tetrameric) lactose repressor R that can bind to two different operators [11],

$$\begin{array}{c|c|c} M \overset{k_E}{\underset{\tau_E}{\rightarrow}} E + M & 2E \underset{k_d}{\overset{k_D}{\rightleftharpoons}} D & 2D \underset{k_r}{\overset{k_R}{\rightleftharpoons}} R \\ \text{translation,} & \text{dimerization,} & \text{tetramerization.} \end{array} \tag{7}$$

The parameter τ_E stands for the time delay involved in the translation process. The lactose repressor ($LacI$) is a tetrameric molecule, whose function as a repressor is blocked by the inducer 1-isopropyl-β-D-thiogalactoside (better known as IPTG). Properly speaking, the tetrameric lactose repressor R is a V-shaped molecule, whose branches are two dimers; each dimer D is able to bind independently to a *lac* operator; and molecules of IPTG can bind to a dimer D decreasing by one thousand its affinity for the *lac* operator itself [11, 16]. The different iterations between the IPTG (represented by I) and the molecules D and R can be represented by the following reactions,

$$D + I \underset{k_i}{\overset{k_I}{\rightleftharpoons}} D^I, \quad R + I \underset{k_i}{\overset{2k_I}{\rightleftharpoons}} R^I, \quad R^I + I \underset{k_i}{\overset{k_I}{\rightleftharpoons}} R^{2I}, \tag{8}$$

where D^I is a dimer *contaminated* with IPTG and R^I (or R^{2I}) is a tetramer with one branch (or both branches) also *contaminated* with IPTG. The *contaminated* dimers can participate in the tetramerization process, so that we must add the following two reactions,

$$2D^I \underset{k_r}{\overset{k_R}{\rightleftharpoons}} R^{2I} \quad \text{and} \quad D + D^I \underset{k_r}{\overset{k_R}{\rightleftharpoons}} R^I. \tag{9}$$

Moreover, the tetrameric repressor R^I with only one branch *contaminated* with IPTG and the dimer D can both bind to one of the *lac* operators in $P_{LlacO\text{-}1}$, rendering the transcription impossible, so that we add the following two reactions,

$$O_0 + D \underset{k_a}{\overset{k_A}{\rightleftharpoons}} O_D \quad \text{and} \quad O_0 + R^I \underset{k_a}{\overset{k_A}{\rightleftharpoons}} O_R^I. \tag{10}$$

It is important to point out that the propensity k_A for the formation of the complexes O_D and O_R^I is half of the propensity $2k_A$ for the formation of the original complex O_R given in (6). Finally, we complete our model for Stricker et al. oscillator by including the degradation of the mRNA (M) and the different proteins,

$$\begin{array}{l} M \xrightarrow{\gamma_M} \emptyset, \ E \xrightarrow{\gamma_E} \emptyset, \quad D \xrightarrow{\gamma_D} \emptyset, \quad D^I \xrightarrow{\gamma_D} \emptyset, \\ R \xrightarrow{\gamma_R} \emptyset, \ R^I \xrightarrow{\gamma_R} \emptyset, \ R^{2I} \xrightarrow{\gamma_R} \emptyset, \end{array} \tag{11}$$

The value of the chemical kinetic parameters involved in the system of reactions (6) to (11) is deduced in Appendices C and E. We include below a list of these parameters. The standard units are minutes (min) and "molecule count per average sized bacterium" (mpb),

$$\begin{array}{ll} k_A = 233.3/mpb/min, & k_E = 10/min, \\ k_a = 0.0372/min, & k_D = k_R = 0.018/mpb/min, \\ k_B = 0.4667/mpb/min, & k_I = 1.53 \times 10^{-3}/mpb/min, \end{array}$$

$$\begin{array}{l|l} k_b = 0.0408/min, & k_d = k_r = k_i = 0/min, \\ k_m = 2/min, & \gamma_M = 0.693/min, \\ [P] = 8300 mpb, & \gamma_E = \gamma_D = \gamma_R = 0.1/min, \\ [I] \in [0, 10^3] mpb, & \tau_E = 0.375 min. \end{array} \quad (12)$$

Finally, Stricker et al. have found that the *in vivo* implementation of their simplest genetic oscillator displays strong variations in the period of the oscillations [4]. The average period is approximately equal to 30 minutes and Stricker et al. also explained in the supplementary material of their article that the variations in the oscillations are a consequence of both the biochemical noise and the fact that their genetic oscillator has only one feedback loop. These results indicate that we must simulate the system of reactions (6) to (11) using an adequate stochastic algorithm, which allows us to include in our simulations the effects of both the biochemical noise and the delayed reaction in (7) that represents the translation. We choose the rejection method of Barrio et al., which is a modification of the Doob-Gillespie algorithm [17, 18] and which was proved by Cai to be an exact stochastic simulation algorithm (SSA) able to simulate the correct time evolution of a chemical reacting system that involves delayed reactions [19]. The pseudocode for the rejection method of Barrio et al. is presented in Appendix F. We perform one million simulations with this SSA; and to do so we consider thirteen chemical species: O_0, O_P^c, O_D, O_R, O_R^I, P, M, E, D, D^I, R, R^I, R^{2I}, which interact through the chemical system of reactions (6) to (11).

Now then, we look for evidence of oscillations, and if it is the case, for the average period of such oscillations. Hence, we calculate the average power spectrum for the active repressor molecular counting, because any oscillation appears as a local maximum located at a frequency different from zero in the power spectrum. Moreover, the average period is given by the frequency ($\neq 0$) where the local maximum is located. We proceed as follows in order to analyze the system of reactions (6) to (11): We perform one million trajectories $x_k(t)$ with the rejection method of Barrio et al. described in Appendix F. Each trajectory $x_j(t)$ is performed from zero to 500 minutes and it measures the molecular counting of the active lactose repressor tetramer R at the time t. We have observed that the transitory state always happens between the first 100 minutes. Hence, we account only for the last 300 minutes from all the trajectories in order to eliminate completely the transitory state; and then we sample them recording a measurement every minute (we obviously get 300 samples from every trajectory). Finally, we calculate the fast Fourier transform of every sampled trajectory and work out the corresponding average power spectra. I.e., if $c_k(n)$ are the coefficients obtained from the fast Fourier transform of the sampled trajectory $x_k(t)$, the average power spectrum is given by the formula,

$$C(n) = \frac{1}{N}\sum_{k=1}^{N} |c_k(n)|^2 \quad \text{for} \quad n = 0, 1, 2, ..., 299,$$

where $N = 10^6$ is the number of performed trajectories. The results are presented in Section 4.

3. Results on regulatory ranges

We verify in this section that equation (5) correctly predicts the regulatory ranges measured *in vivo* for the promoters P_{lac}, $P_{A1lacO\text{-}1}$, $P_{LlacO\text{-}1}$, and $P_{LtetO\text{-}1}$. The error is less than 7.7% in the four cases. Recall that the parameter n in (5) denotes the number of active operators located in the core part of the promoter, while the parameter $\beta = 0$ for all the promoters, with exception of $P_{LtetO\text{-}1}$, for which we assume that $\beta = 3$.

3.1. Regulatory range of the lac promoter P_{lac}

Oehler, Amouyal, et al. measured *in vivo* a regulatory range of 200 for the *lac* promoter with only one functional *lac* operator next to the transcriptional starting site (P_{lac}) and a concentration of 50 tetrameric repressors (200 monomeric subunits) per cell [9, p. 3350]. We obtain the value 215.2 from equation (5) after substituting the value of the parameters below; see Appendix A.

$$\begin{array}{l|l} k_A = 7 \times 10^9/M/s, & k_a = 6.2 \times 10^{-4}/s, \\ k_B = 2 \times 10^6/M/s, & k_b = 0.0033/s, \\ [R] = 50\, mpb, & k_m \approx 0.003/s, \\ [P] = 8300\, mpb, & n(1+\beta) = 1. \end{array} \quad (13)$$

As we said in the introduction, Oehler, Eismann et al. also measured *in vivo* a regulatory range of 18 for the promoter P_{lac} in the *E.coli* strain CSH 9 rec A [8, p. 974]. We obtain the value 18.1 from equation (5) after substituting the value of the parameters above, with the exception of the number of active repressors per cell $[R]$, for which we take $4\, mpb$ instead of $50\, mpb$. The concentration $[R] = 4\, mpb$ seems to be

compatible with the assessments done by Gilbert et al. and Müller-Hill et al. [20, 21], who indicate that the repressor *LacI* is synthesized about 5 to 10 molecules each generation.

3.2. Regulatory range of promoter $P_{A1lacO\text{-}1}$

Lutz and Bujard measured *in vivo* a regulatory range of 350 for the promoter $P_{A1lacO\text{-}1}$, when controls the *luciferase* reporter gene in a pZ vector [1, p. 1205]. This promoter is build from the bacteriophage promoter P_{A1} by adding two functional *lac* operators, one of them between the -10 and -33 hexamers. Lutz and Bujard induced a constitutive overproduction of the repressor *LacI* by placing a copy of the *lacI* gene under the control of the mutant constitutive $laci^q$ promoter [21, p. 1260]. We obtain the value 372 from equation (5) after substituting the value of the parameters below; see Appendix B.

$$\begin{array}{l|l} k_A = 7\times 10^9/M/s, & k_a = 6.2\times 10^{-4}/s, \\ k_B = 6\times 10^7/M/s, & k_b = 9.2\times 10^{-4}/s, \\ [R] = 90\,mpb, & k_m \approx 0.09/s, \\ [P] = 8300\,mpb, & n(1+\beta) = 2. \end{array} \tag{14}$$

3.3. Regulatory range of promoter $P_{LlacO\text{-}1}$

Lutz and Bujard measured *in vivo* a regulatory range of 620 for the promoter $P_{LlacO\text{-}1}$, when controls the *luciferase* reporter gene in a Pz vector [1, p. 1205]. This promoter is build from the coliphage-lambda promoter P_L by adding two functional *lac* operators, one of them between the -10 and -33 hexamers. Lutz and Bujard induced a constitutive overproduction of the repressor *LacI* by placing a copy of the *lacI* gene under the control of the mutant constitutive $laci^q$ promoter [21, p. 1260]. We obtain the value 595.9 from equation (5) after substituting the value of the parameters below; see Appendix C.

$$\begin{array}{l|l} k_A = 7\times 10^9/M/s, & k_a = 6.2\times 10^{-4}/s, \\ k_B = 1.4\times 10^7/M/s, & k_b = 6.8\times 10^{-4}/s, \\ [R] = 90\,mpb, & k_m \approx 0.0333/s, \\ [P] = 8300\,mpb, & n(1+\beta) = 2. \end{array} \tag{15}$$

3.4. Regulatory range of the promoter $P_{LtetO\text{-}1}$

Lutz and Bujard measured *in vivo* a regulatory range 2535 for the promoter $P_{LtetO\text{-}1}$, when controls the *luciferase* reporter gene in a Pz vector [1, p. 1205]. This promoter is build from the coliphage-lambda promoter P_L by adding two functional *tet* operators, one of them between the -10 and -33 hexamers. Lutz and Bujard induced a constitutive overproduction of the repressor *TetR* by placing a copy of the *tetR* gene under the control of the constitutive promoter P_{N25}. We obtain the value 2459.8 from equation (5) after substituting the value of the parameters below; see Appendix D.

$$\begin{array}{l|l} k_A = 4.81\times 10^9/M/s, & k_a = 9.62\times 10^{-4}/s, \\ k_B = 1.4\times 10^7/M/s, & k_b = 6.8\times 10^{-4}/s, \\ [R] = 210\,mpb, & k_m \approx 0.0333/s, \\ [P] = 8300\,mpb, & n(1+\beta) = 8. \end{array} \tag{16}$$

Lutz and Bujard also measured *in vivo* the regulatory ranges 3670 (or 5050) for the same promoter $P_{LtetO\text{-}1}$, when controls the *luciferase* reporter gene in a Pz vector that contains the origin of replication P15A (or pSC101*) instead of ColE1 [1]. We think that these repression ranges were obtained in bacteria that have 6×10^3 (or 4×10^3) active polymerases.

4. Results for Stricker et al. simple genetic oscillator

We verify in this section that the system of chemical reactions (6) to (11) is able to produce oscillations with an average period of approximately 30 minutes, when we use the parameters given in (12). A period of 30 minutes would be a good indication that reactions (6) to (11) correctly model the simplest genetic oscillator build by Stricker et al., because they have found *in vivo* oscillations with the same average period [4]. We look for oscillations by calculating the average power spectrum and observing if there are local maxima located at a frequency different from zero.

In Fig. 3 we present several average power spectra that we calculated for the molecular counting of the active lactose repressor tetramer R under different concentrations of IPTG. One can observe local maxima located at frequencies of approximately 0.5 and 0.6 miliHertz for IPTG's concentrations of $100mpb$ and $150mpb$, respectivelly. These local maxima correspond to oscillations with average periods of 33.3 and 27.8 minutes, and so we may conclude that there is a good agreement with the *in vivo* results obtained by Stricker et al. Recall that oscillations with a frequency of one milliHertz have a period of 16.67 minutes. It is also important to observe from Fig. 3 that the inten-

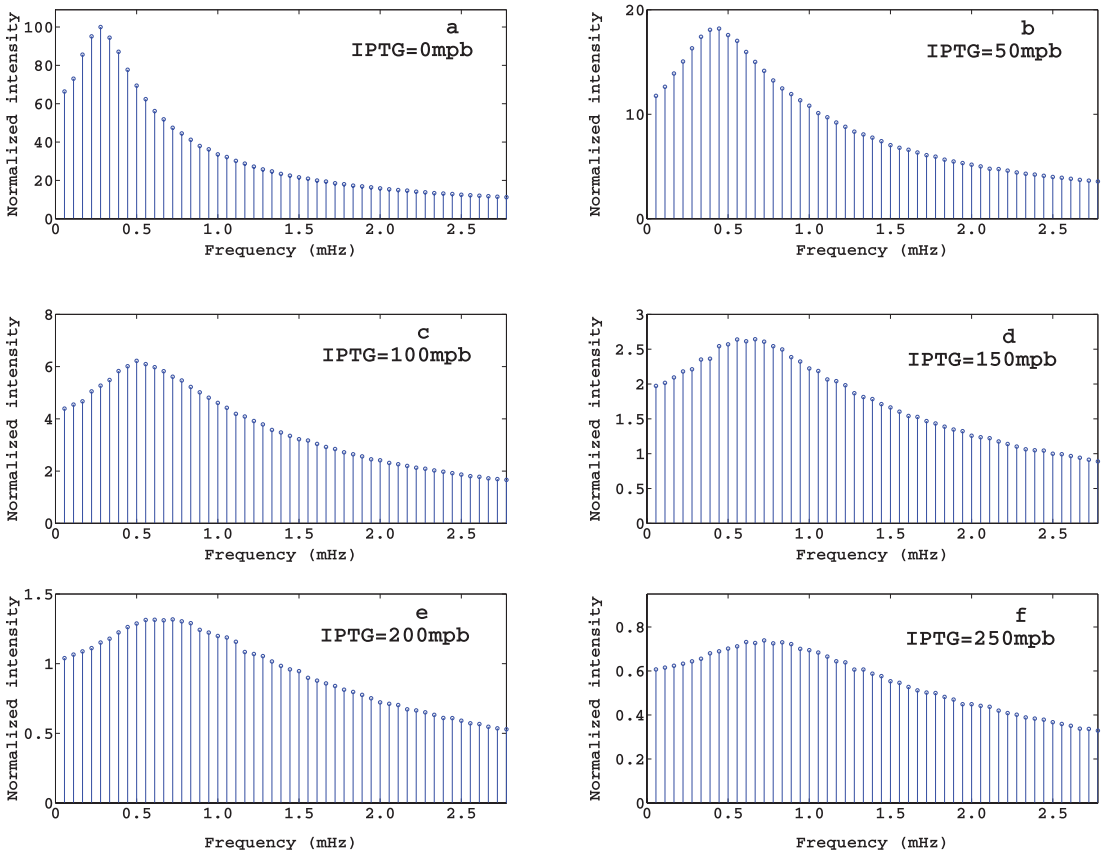

Fig. 3. Average power spectra for the molecular counting of the active lactose repressor tetramer R, when the promoter is $P_{LlacO\text{-}1}$ and the concentration of IPTG is: $0mbp$, **a**; $50mbp$, **b**; $100mbp$, **c**; $150mbp$, **d**; $200mbp$, **e**; $250mbp$, **f**. The vertical axis denotes normalized intensity in arbitrary units and the horizontal axis denotes frequency in miliHertz. Oscillations with an average period of 33.3 and 27.8 minutes can be observed when the concentration of IPTG is $100mpb$ and $150mpb$, respectively.

sity of the oscillations diminish as the concentration of IPTG increases. This effect happens because the IPTG is indeed capturing the active *lac* repressor tetramers R and rendering them inactive.

On the other hand, recall that Stricker et al. implemented *in vivo* their oscillator by introducing two different pZ vector into a strain of *E.coli*. One pZ vector contains a copy of the *lacI* structural genes controlled by the promoter $P_{LlacO\text{-}1}$; while the other contains a *yemGFP* gene controlled by the same kind of promoter. Thus, once we have verified that the system of reactions (6) to (11) correctly models the simplest genetic oscillator constructed by Stricker et al., we may wonder what would be the effect of substituting the promoter $P_{LlacO\text{-}1}$ by any other one.

Therefore, we now calculate the average power spectra for the molecular counting of the active lactose repressor tetramer R, when the promoter $P_{LlacO\text{-}1}$ has been substituted by P_{lac} or $P_{A1lacO\text{-}1}$. To do so, we made again all the calculations indicated at the end of Subsection 1, but we use the value of the parameters k_A, k_a, k_B, k_b, k_m and n indicated in (13) or (14) instead of those given in (12). In particular, one can verify that the average power spectra have the same shape and local maxima when the promoter $P_{LlacO\text{-}1}$ is substituted by $P_{A1lacO\text{-}1}$, so that the *in vivo* versions of the corresponding genetic oscillators should have similar behavior. Nevertheless, one also finds that the oscillations almost disappear behind the biochemical noise when the promoter $P_{LlacO\text{-}1}$ is substituted by P_{lac}. In Fig. 4 we present four average power spectra calculated with the promoter P_{lac} and different concentrations of IPTG. One can observe that the local maxima are not as evident as in Fig. 3-**a**, so that the oscillations are erratic and hidden by the noise.

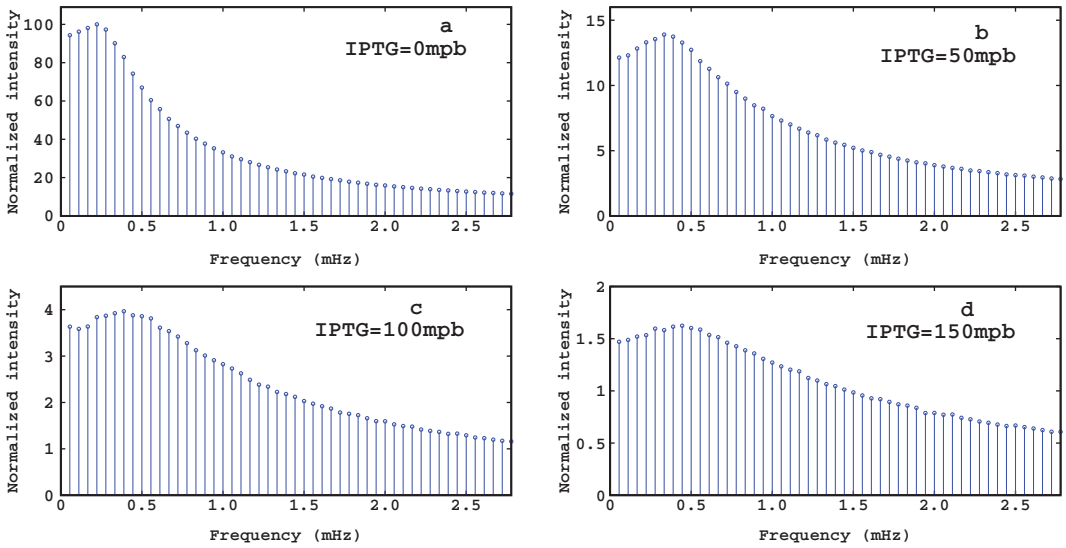

Fig. 4. Average power spectra for the molecular counting of the active lactose repressor tetramer R, when the promoter is P_{lac} and the concentration of IPTG is: $0mbp$, **a**; $50mbp$, **b**; $100mbp$, **c**; $150mbp$, **d**. The vertical axis denotes normalized intensity in arbitrary units and the horizontal axis denotes frequency in miliHertz.

5. Conclusions

The positive results presented in previous Section 3 and 4 imply that the premises made in Assumptions 1 and 2 are compatible with the experimental data, in particular, we can conclude that:

1. If the central base pairs of two *lac* operators are separated by less than 32.5bp, the probability that two repressors bind to them at the same time can be assumed to be negligible. I.e., once a repressor *LacI* binds to one of the *lac* operators, it physically prevents a second repressor getting bound to the adjacent one.
2. We verify Bond et al. assertion [10] that the cooperativity between two *lac* operators can be assume to be negligible when their central base pairs are separated by 22 or 32bp.
3. The repression and transcription initiation processes for Lutz and Bujard's synthetic controllable promoters: $P_{A1lacO\text{-}1}$, $P_{LlacO\text{-}1}$, $P_{LtetO\text{-}1}$, and $P_{N25tetO\text{-}1}$ can be modeled as the system of chemical reactions (1); and these reactions can be used as a building block for modeling more complex genetic systems based on Lutz and Bujard's promoters.
4. The model of chemical reactions (1) correctly predicts, via equation (5), the *in vivo* regulatory ranges of the promoters: P_{lac}, $P_{A1lacO\text{-}1}$, $P_{LlacO\text{-}1}$, and $P_{LtetO\text{-}1}$, when one substitutes the given parameters by the corresponding chemical kinetic constants measured *in vitro* by several authors. In particular, the probability $P(O_P^c)$ that a polymerase forms a closed complex with the respective promoter can be expressed as the Hill function introduced in (4); i.e.,

$$P(O_P^c) = \frac{[P]}{[P] + m[R]},$$

where $[R]$ and $[P]$ are the number of active repressors and polymerases, respectively, and the value of m is obtained by substituting into equation (4) the parameters given in (13) to (16):

	P_{lac}	$P_{A1lacO\text{-}1}$	$P_{LlacO\text{-}1}$	$P_{LtetO\text{-}1}$
m	35.6×10^3	34.2×10^3	54.9×10^3	97.2×10^3.

5. We deduce that the *in vivo* simplest genetic oscillator of Stricker et al. [4] must have a similar oscillatory behavior if the promoter $P_{LlacO\text{-}1}$ is substituted by $P_{A1lacO\text{-}1}$. Recall that Stricker et al. oscillator is based on a pZ vector, which contains a copy of the *lacI* structural genes controlled by the promoter $P_{LlacO\text{-}1}$. Moreover, no oscillations must be observed if the promoter $P_{LlacO\text{-}1}$ is

substituted by P_{lac}. These results must be verified experimentally (*in vivo*).

6. As we said in the introduction, Oehler, Eismann et al. measured *in vivo* a regulatory range of 18 for the promoter P_{lac} [8]. This range is one order of magnitude smaller than the ranges reported by Lutz and Bujard (350 and 620), because the concentration of repressor $LacI$ is $4\,mpb$ in Oehler et al. experiments, while the concentration is $90\,mpb$ in Lutz and Bujard experiments.
7. Oehler, Amouyal et al. also measured *in vivo* a regulatory range greater than 12×10^3 for a synthetic *lac* promoter that includes two functional *lac* operators inside its structure [9]. Oehler et al. explained that the magnitude of this range ($\geq 12 \times 10^3$) is a consequence of the cooperativity existing between both *lac* operators, and so it is one order of magnitude larger than the ranges reported by Lutz and Bujard (350 and 620), because in the promoters $P_{A1lacO\text{-}1}$ and $P_{LlacO\text{-}1}$ the *lac* operators do not cooperate to produce a larger repression.

Acknowledgments

We are quite grateful to the reviewers for their observations and suggestions, they have been quite useful for writing down the final version of our paper.

6. Appendix A: Parameters measured *in vitro* for the promoter P_{lac}

1. Complex (repressor $LacI$)/operator formation rate constant k_A. Riggs et al. measured $7(\pm 0.9)\times 10^9/M/s$ *in vitro* [22, p. 408]. We take $k_A = 7\times 10^9/M/s$.
2. Complex (repressor $LacI$)/operator dissociation rate constant k_a. Riggs et al. measured $6.2(\pm 1.3)\times 10^{-4}/s$ *in vitro* [22, p. 404]. We take $k_a = 6.2\times 10^{-4}/s$.
3. Complex polymerase/(promoter P_{lac}) formation rate constant k_B. Lanzer and Bujard state the rate $2(\pm 1)\times 10^6/M/s$ [23, p. 8975]. We take $k_B = 2\times 10^6/M/s$.
4. Complex polymerase/(promoter P_{lac}) dissociation rate constant k_b. Meiklejohn and Gralla measured a half-life time of 3.5 minutes for the wild-type complex in presence of CRP-cAMP [24, p. 771]. We obtain $k_b = \frac{\ln 2}{210s} = 0.0033/s$.
5. *Transcription initiation rate* k_m. The polymerase opens the DNA in the second transcription step. Malan et al. measured an opening rate of $0.003/s$ for the *lac* wild-type promoter in presence of CRP-cAMP [25, p. 891], so we take $k_m \approx 0.003/s$.
6. *Number of repressors LacI per cell* $[R]$. Oehler et al. induced 200 monomeric subunits/cell [9, p. 3350], it corresponds to $[R] = 50$ tetrameric repressors/cell.
7. *Number of polymerases per cell* $[P]$. Bremer and Dennis state there are between 1,500 and 11,400 polymerases/cell [26]. We take $[P] = 8300$ polymerases/cell.

7. Appendix B: Parameters measured *in vitro* for the promoter $P_{A1lacO\text{-}1}$

1. *Complex repressor/operator formation rate constant* $k_A = 7\times 10^9/M/s$; see Appendix A.
2. *Complex repressor/operator dissociation rate constant* $k_a = 6.2\times 10^{-4}/s$; see Appendix A.
3. *Complex polymerase/(promoter* $P_{A1lacO\text{-}1}$*) formation rate constant* k_B. Lanzer and Bujard measured $6(\pm 0.5)\times 10^7/M/s$ for the promoter $P_{A1/O4}$ with a *lac* operator located between the -10 and -33 hexamers [23, p. 8975]. We take $k_B = 6\times 10^7/M/s$.
4. *Complex polymerase/(promoter* $P_{A1lacO\text{-}1}$*) dissociation rate constant* k_b. Brunner and Bujard measured $k_b = 9.2\times 10^{-4}/s$ for the promoter P_{A1} [27, p. 3141].
5. *Transcription initiation rate* k_m. The polymerase opens the DNA in the second transcription step. Dayton et al. measured a DNA opening rate of $0.09/s$ for the bacteriophage promoter P_{A1} [28, p. 1618], so we take $k_m \approx 0.09/s$.
6. *Number of repressors LacI per cell* $[R]$. Müller et al. state that the mutant promoter $laci^q$ induces between 80 and 100 repressors/cell [21, p. 1260]. We take $[R] = 90$ repressors/cell.
7. *Number of polymerases per cell* $[P] = 8300$ polymerases/cell; see Appendix A.

8. Appendix C: Parameters measured *in vitro* for the promoter $P_{LlacO\text{-}1}$

1. *Complex repressor/operator formation rate constant* $k_A = 7\times10^9/M/s$; see Appendix A.
2. *Complex repressor/operator dissociation rate constant* $k_a = 6.2\times10^{-4}/s$; see Appendix A.
3. *Complex polymerase/(promoter* $P_{LlacO\text{-}1}$*) formation rate constant* k_B. Lanzer and Bujard measured $1.4(\pm0.2)\times10^7/M/s$ for the promoter P_L with a *lac* operator located between the -10 and -33 hexamers [23, p. 8975]. We take $k_B = 1.4\times10^7/M/s$.
4. *Complex polymerase/(promoter* $P_{LlacO\text{-}1}$*) dissociation rate constant* k_b. Brunner and Bujard measured $k_b = 6.8\times10^{-4}/s$ for the promoter P_L [27, p. 3141].
5. *Transcription initiation rate* k_m. McClure observed a delay of 60s in the opening of complex polymerase/(promoter P_L) [29, p. 5636]. We take $k_m = \frac{2}{60s}$.
6. *Number of repressors LacI per cell* $[R] = 90$ repressors/cell; see Appendix B.
7. *Number of polymerases per cell* $[P] = 8300$ polymerases/cell; see Appendix A.

9. Appendix D: Parameters measured *in vitro* for the promoter $P_{LtetO\text{-}1}$

1. *Complex (repressor TetR)/operator equilibrium binding constant* (k_A/k_a). Hillen et al. measured a value between $10^{12}/M$ and $10^{13}/M$ [13, p. 715]. We take $k_A/k_a = 5\times10^{12}/M$.
2. *Complex (repressor TetR)/operator dissociation rate constant* k_a. Hillen et al. measured a half-life of 12 min. [13, p. 715]. We obtain $k_a = \frac{\ln 2}{720s} \approx 9.62\times10^{-4}/s$.
3. *Complex polymerase/promoter formation rate constant* $k_B = 1.4\times10^7/M/s$; see Appendix C.
4. *Complex polymerase/promoter dissociation rate constant* $k_b = 6.8\times10^{-4}/s$; see Appendix C.
5. *Transcription initiation rate* $k_m = \frac{2}{60s}$; see Appendix C.
6. *Number of repressors TetR per cell* $[R]$. Lutz and Bujard measured a ratio of 7/3 as the number of repressors *TetR* per cell divided by the number of repressors *LacI* [1]. Since there are 90 repressors *LacI* per cell according to Appendix B, we take $[R] = 210$.
7. *Number of polymerases per cell* $[P] = 8300$ polymerases/cell; see Appendix A.

10. Appendix E: Parameters involved in Stricker et al. simple oscillator

Some parameters were already given in Appendix C above; the remaining necessary parameters are estimated as follows:

- *E.coli volume*. Volkmer and Heinemann measured the volume of a single *E. coli* under different growth conditions, they reported a variation in the volume from 1.6 to 4.4 *fl* [30]. We take a mean volume of 3 *fl*, so that we get the conversion

$$1M = \frac{6.0221\times10^{23} molecules}{1l} \times \frac{3\times10^{-15} l}{1 bacteria} = 1.8\times10^9 \text{mpb}.$$

- *Number of lacI structural genes per cell*. There is only one *lacI* gene in the corresponding pZ vector. Stricker et al. (2008) state that there are between one and three pZ vectors per cell [4], so we take two structural *lacI* genes per cell.
- *Translation initiation rate for the lactose repressor monomer* k_E. Chen et al. state the value $k_E = 10/min$ [31].
- *Dimerization and tetramerization formation rate constants* k_D *and* k_R. We use the values $k_D = k_R = 0.018/mpb/min$ given by Stricker et al. in [4] (suppl. material).
- *Complex (inducer IPTG)/(repressor LacI) formation rate constant* k_I. Dunaway et al. measured *in vitro* the value $k_I = 4.6\times10^4/M/s$ [32].
- *Dimerization and tetramerization dissociation rate constants* k_d *and* k_r. They seem to be negligible in comparison to $k_D = k_R$, so that we take $k_d = k_r = 0/mpb/min$.
- *Complex (inducer IPTG)/(repressor LacI) dissociation rate constant* k_i. It seems to be negligible in comparison to k_I, so that we also take $k_i = 0/mpb/min$.
- *Degradation rate constant for the lacI mRNA:* γ_M. Varmus et al. measured a half-life time of approximately one minute [33, p.2264], so that $\gamma_M \approx \frac{ln(2)}{1min}$.
- *Degradation rate constants for the repressor LacI:* γ_E, γ_D, *and* γ_R. We assume that the

monomeric E, dimeric D, and tetrameric R forms of the repressor $LacI$ all have the same degradation rate. Chen et al. measured it, $\gamma_E = \gamma_D = \gamma_R = 0.1/min$ [31].

- *Time delay in the transcription of the lac mRNA molecule* τ_M. We assume that $\tau_M = 0$, because the translation of the nascent *lac* mRNA molecule can begin when the first two segments of the leader region have been transcribed and a ribosome can bind to them [11].
- *Time delay in the translation of the lactose repressor monomer* τ_E. The lactose repressor monomer is a protein of 360 amino acids residues [11, p. 523]. Bremer and Dennis states that the peptide chain elongation lies between 12 and 21 residues per second [26], so that $\tau_E \approx \frac{360 res.}{16 res./s} = 22.5s$.

11. Appendix F: Algorithms for simulating well-stirred chemical reacting systems

A well-stirred chemical reacting system is naturally described as a set of $n \geq 1$ different chemical species $\{S_1, \ldots, S_n\}$ that interact through $m \geq 1$ different chemical reactions. The dynamic state of such a system is represented by a (state) vector of natural numbers $X(t) \in \mathbb{N}^n$, so that the molecule count at time $t \in \mathbb{R}$ of the k-th chemical species (S_k) is represented by the k-entry $(X_k(t))$ of the vector $X(t)$. The chemical reactions taking place in the reactor are described in the usual form:

$$X \xrightarrow{a(j,X)} X + v[j], \qquad j = 1, 2, ..., m, \tag{17}$$

where $a(j, X) \geq 0$ denotes the propensity of the j-th chemical reaction, and $v[j] \in \mathbb{Z}^n$ is its corresponding stoichiometric vector. The Doob-Gillespie algorithm [17, 18] is an exact stochastic simulation algorithm (SSA) used for calculating the time evolution of the state vector $X(t)$ under some assumptions: the chemical reactions in (17) are all instantaneous, no two chemical reactions can happen simultaneously, and the probability that the j-th reaction happens at the time interval $[t, t + \delta]$ is proportional to $a(j, X(t))\delta$. The pseudocode for the Doob-Gillespie algorithm is the following one:

1. Initialization. Set the time $t = t_0$ and the state vector $X(t_0) = X_0$ to their given initial values.
2. Calculate $\widehat{a} := \sum_{j=1}^{m} a_j$, where $a_j := a(j, X(t))$ is the propensity function for the j-th reaction in (17). Generate the waiting time $\tau = -\ln(u_1)/\widehat{a}$ from a standard uniform random variable u_1 in the interval $[0, 1]$.
3. Generate a standard uniform random variable u_2 in the interval $[0, 1]$ and choose the index $\mu = 1, 2, ..., m$ such that $\sum_{j=1}^{\mu-1} a_j < u_2\widehat{a} \leq \sum_{j=1}^{\mu} a_j$.
4. Set the state vector $X(t + \tau) = X(t) + v[\mu]$ and update the time t to $t + \tau$. Go to the second step or else stop.

Since one can only analyze instantaneous reactions with the original Doob-Gillespie algorithm, it has been modified to produce several SSAs able to simulate the time evolution of a chemical reacting system that includes delayed reactions; i.e., chemical reactions like gene transcription and translation that take some time to deliver their products. Cai proved in [19] that the rejection method of Barrio et al. described below is an exact SSA able to include delayed reactions, and so we use it for calculating the time evolution of the state vector $X(t)$ of the chemical reacting system (6) to (11). For simplicity, we adapt the algorithm for the case when there is only one non-consuming delayed reaction (indexed as the 0-th reaction)

$$X \xrightarrow{a(0,X)} X + v[0], \qquad \text{with a time delay } \tau_0 > 0. \tag{18}$$

By convention, in non-consuming reactions the reactants of an unfinished reaction can participate in new reactions, and so the corresponding molecular counts only change when the reaction is finished, like in gene transcription. The pseudo-code for the rejection method of Barrio et al. is the following one:

1. Initialization. Set the time $t = t_0$ and the state vector $X(t_0) = X_0$ to their given initial values.
2. Calculate $\widehat{a} := \sum_{j=0}^{m} a_j$, where $a_j := a(j, X(t))$ is the propensity function for the j-th reaction in (17) and (18). Generate the waiting time $\tau = -\ln(u_1)/\widehat{a}$ from a standard uniform random variable u_1 in the interval $[0, 1]$. If there is any delayed reaction (18) to finish in the time interval $[t, t + \tau)$, find the time $t_d \in [t, t + \tau)$ of the first delayed reaction to finish, discard τ, set the state vector $X(t_d) = X(t) + v[0]$, update the time $t = t_d$, and repeat step 2. If there is no delayed reaction to finish in $[t, t + \tau)$, proceed to step 3.
3. Generate a standard uniform random variable u_2 in the interval $[0, 1]$ and choose the index

$\mu = 0, 1, ..., m$ such that $\sum_{j=0}^{\mu-1} a_j < u_2\widehat{a} \leq \sum_{j=0}^{\mu} a_j$. If $\mu = 0$, add the ending time $t + \tau_0$ to the list of delayed reactions that are in process of finishing. If $\mu \neq 0$, set the state vector $X(t + \tau) = X(t) + v[\mu]$.

4. Update the time t to $t + \tau$. Go to the second step or else stop.

Notice that we need to include a first-in, first-out queue in this algorithm to save the ending times of the delayed chemical reactions that are in process of finishing.

References

[1] R. Lutz and H. Bujard, Independent and tight regulation of transcriptional units in *Escherichia coli via* the LacR/O, the TetR/O and AraC/I_1-I_2 regulatory elements, *Nucleic Acids Res* **25**(6) (1997), 1203–1210.

[2] Expressys: Tools for gene expression, http://www.expressys.com.

[3] M.B. Elowitz and S. Leibler, A synthetic oscillatory network of transcriptional regulators, *Nature* **403** (2000), 335–338.

[4] J. Stricker, S. Cookson, M.R. Bennett, W.H. Mather, L.S. Tsimring and J. Hasty, A fast, robust and tunable synthetic gene oscillator, *Nature* **456**(7221) (2008), 516–519.

[5] J.-J. Jo and J.-S. Shin, Construction of intragenic synthetic riboswitches for detection of a small molecule, *Biotechnol Lett* **31**(10) (2009), 1577–1581.

[6] S.J. Park, T.W. Lee, S.-C. Lim, T.W. Kim, H. Lee, M.K. Kim, S.H. Lee, B.K. Song and S.Y. Lee, Biosynthesis of polyhydroxyalkanoates containing 2-hydroxybutyrate from unrelated carbon source by metabolically engineered Escherichia coli, *Appl Microbiol Biotechnol* **93**(1) (2012), 273–283.

[7] Y. Soma, K. Tsuruno, M. Wada, A. Yokota and T. Hanai, Metabolic flux redirection from a central metabolic pathway toward a synthetic pathway using a metabolic toggle switch, *Metab Eng* **23** (2014), 175–184.

[8] S. Oehler, E.R. Eismann, H. Krämer and B. Müller-Hill, The three operators of the lac operon cooperate in repression, *EMBO J* **9**(4) (1990), 973–979.

[9] S. Oehler, M. Amouyal, P. Kolkhof, B. von Wilcken-Bergmann and B. Müller-Hill, Quality and position of the three lac operators of *E.coli* define efficiency of repression, *EMBO J* **13**(14) (1994), 3348–3355.

[10] L.M. Bond, J.P. Peters, N.A. Becker, J.D. Kahn and L.J. Maher III, Gene repression by minimal lac loops *in vivo*, *Nucleic Acids Res* **38**(22) (2010), 8072–8082.

[11] M. Lewis, The lac repressor, *C R Biol* **328**(6) (2005) 521–548.

[12] J. Müller, S. Oehler and B. Müller-Hill, Repression of lac promoter as a function of distance, phase and quality of an auxiliary lac operator, *J Mol Biol* **257**(1) (1996), 21–29.

[13] W. Hillen, C. Gatz, L. Altschmied, K. Schollmeier and I. Meier, Control of expression of the Tn10-encoded tetracycline resistance genes, *J Mol Biol* **169**(3) (1983), 707–721.

[14] Uniprot.org: Tetracycline repressor protein TetR, http://www.uniprot.org/uniprot/Q79VX4.

[15] P.L. deHaseth, M.L. Zupancic and M.T. Record Jr., RNA polymerase-promoter interactions: The comings and goings of RNA polymerase, *J Bacteriol* **180**(12) (1998), 3019–3025.

[16] M.D. Barkley, A.D. Riggs, A. Jobe and S. Bourgeois, Interaction of effecting ligands with lac repressor and repressor-operator complex, *Biochemistry* **14**(8) (1975), 1700–1712.

[17] J.L. Doob, Markoff chains–denumerable case, *Trans Am Math Soc* **58**(3) (1945), 455–473.

[18] D.T. Gillespie, Exact stochastic simulation of coupled chemical reactions, *J Phys Chem* **81**(25) (1977), 2340–2361.

[19] X. Cai, Exact stochastic simulation of coupled chemical reactions with delays, *J Chem Phys* **126**(12) (2007), 124108.

[20] W. Gilbert and B. Müller-Hill, Isolation of the lac repressor, *Proc Natl Acad Sci U S A* **56**(6) (1966), 1891–1898.

[21] B. Müller-Hill, L. Crapo and W. Gilbert, Mutants that make more lac repressor, *Proc Natl Acad Sci U S A* **59**(4) (1968), 1259–1264.

[22] A.D. Riggs, S. Bourgeois and M. Cohn, The lac represser-operator interaction: III. kinetic studies, *J Mol Biol* **53**(3) (1970), 401–417.

[23] M. Lanzer and H. Bujard, Promoters largely determine the efficiency of repressor action, *Proc Natl Acad Sci U S A* **85**(23) (1988), 8973–8977.

[24] A.L. Meiklejohn and J.D. Gralla, Entry of RNA polymerase at the lac promoter, *Cell* **43**(3) (1985), 769–776.

[25] T. Malan, A. Kolb, H. Buc and W.R. McClure, Mechanism of CRP-cAMP activation of lac operon transcription initiation activation of the P1 promoter, *J Mol Biol* **180**(4) (1984), 881–909.

[26] H. Bremer and P.P. Dennis, Modulation of chemical composition and other parameters of the cell by growth rate, in: F.C. Neidhardt, et al., (Eds.), Escherichia coli and Salmonella: Cellular and molecular biology, 2nd Edition, ASM press, Washington, D.C., 1996, pp. 1553–1569.

[27] M. Brunner and H. Bujard, Promoter recognition and promoter strength in Escherichia coli system, *EMBO J* **6**(10) (1987), 3139–3144.

[28] C.J. Dayton, D.E. Prosen, K.L. Parker and C.L. Cech, Kinetic measurements of Escherichia coli RNA polymerase association with bacteriophage T7 early promoters, *J Biol Chem* **259**(3) (1984), 1616–1621.

[29] W.R. McClure, Rate-limiting steps in RNA chain initiation, *Proc Natl Acad Sci U S A* **70**(10) (1980), 5634–5638.

[30] B. Volkmer and M. Heinemann, Condition-dependent cell volume and concentration of escherichia coli to facilitate data conversion for systems biology modeling, *PloS one* **6**(7) (2011), e23126.

[31] W. Chen, J.E. Bailey and S.B. Lee, Molecular design of expression systems: Comparison of different control configurations using molecular mechanism models, *Biotechnol Bioeng* **38**(7) (1991), 679–687.

[32] M. Dunaway, J. Olson, J. Rosenberg, O. Kallai, R. Dickerson and K. Matthews, Kinetic studies of inducer binding to lac repressor operator complex, *J Biol Chem* **255**(21) (1980), 10115–10119.

[33] H.E. Varmus, R.L. Perlman and I. Pastan, Regulation of lac messenger ribonucleic acid synthesis by cyclic adenosine 3', 5'-monophosphate and glucose, *J Biol Chem* **245**(9) (1970), 2259–2267.