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

C.G. Zamora-Chimala and E.S. Zeronb

aCentro de Investigación y de Estudios Avanzados del IPN, Unidad Monterrey, Parque de Investigación e Innovación Tecnológica, Apodaca NL, México
bCentro 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 PLlacO-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 PLlacO-1, PLtetO-1, PLtetO-1, PLtetO-1, and PL23tetO-1 in E.coli. For example, the promoters PLlacO-1 and PLtetO-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, PLlacO-1 and PLtetO-1 are respectively repressed by LacI and TetR, and so induced by IPTG and anhydrotetracycline.
are some apparent contradictions with several of the values reported in Table 1 of [1], because there regulatory ranges measured by Lutz and Bujard have particular, it is important to explain why the their respective binding sites in the promoter. In par-
merases and repressors are competing for occupying explaining in detail how they function; i.e., how poly-
tetras operons: (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. The promoters P_{LtetO}−1, and P_{LtetO}+1 are respectively separated by 32, 22, and 28bp.

The EXPRESSYS company [2] is nowadays commercializing \textit{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 \textit{lac} \textit{Z} gene, or the \textit{luciferase} gene. The \textit{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 \textit{pZ} vectors have been used by several authors [3–7] for \textit{in vivo} experiments. In particular, Stricker et al. [4] used the \textit{pZ} vectors for implementing \textit{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 \textit{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 \textit{in vivo} measurements done by Oehler et al. [8, 9] on the \textit{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_{LtetO}−1 and P_{LtetO}+1 both contain a pair of \textit{lac} operators and their \textit{in vivo} regulatory ranges are 620 and 350, respectively. These values are at least one order of magnitude larger than the \textit{in vivo} regulatory range (= 18) reported by Oehler, Eismann et al. in 1990 for a \textit{lac} promoter with only one functional \textit{lac} operator next to the transcriptional starting site; see configuration \textit{λ} \textit{E} \textit{wt} 70 in Fig. 2 of [8]. Moreover, the regulatory ranges 620 and 350 are at least one order of magnitude smaller than the regulatory range (≥ 12 × 10^3) reported by Oehler, Amouyal et al. in 1994 for a synthetic \textit{lac} promoter that includes two functional \textit{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_{LtetO}, P_{LtetO}+1, P_{LtetO}−1, and P_{LtetO}+1−1. Special attention is paid to deduce all the model parameters from reported (\textit{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 \textit{lac} operators can be assumed to be negligible when their central base pairs are separated by 22 or 32 bp [10]. Recall that two \textit{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 \textit{lac} operators whose central base pairs are separated by a distance between 49 and 98bp, the maxi- 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.

| \textbf{Promoter} | \textbf{Central Sequences} |
|------------------|-------------------------|
| P_{LtetO}−1      | TTACGTATGCTCCGCTCGTT   |
| P_{LtetO}+1      | TTACGTATGCTCCGCTCGTT   |
| P_{LtetO}+1−1    | TTACGTATGCTCCGCTCGTT   |
| P_{LtetO}−1      | AATTCGGATGCTGCCATGTT   |
| P_{LtetO}+1      | AATTCGGATGCTGCCATGTT   |
| P_{LtetO}+1−1    | AATTCGGATGCTGCCATGTT   |

Fig. 1. Topography and sequences of the central parts of the \textit{lac} promoter \textit{P}_{Lacet} 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. The corresponding sequences of the \textit{lac} and \textit{trp} 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.
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 (≈ 35 min) 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.

Fig. 2. Graphic representation of the different states $O_0$, $O_cP$, $O_R$ and $O_2R$ assumed by the promoter $P_{LtetO-1}$ 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.

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; $O_cP$ 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 $P_{A1lacO-1}, P_{LlacO-1}$, and $P_{LtetO-1}$, because only them contain two operators. For example, the different states found in promoter $P_{LtetO-1}$ are graphically represented in Fig. 2. The transitions between the states $O_0$, $O_cP$ and $O_R$ are represented as chemical reactions:

$$O_0 + R \overset{k_A}{\rightleftharpoons} O_{cP}, \quad O_{cP} + R \overset{k_R}{\rightarrow} O_{2R},$$

$$O_0 + P \overset{k_B}{\rightarrow} O_P, \quad O_P \overset{k_P}{\rightarrow} O_0 + P + M, \quad (1)$$

where $M$ represents a nascent mRNA molecule, the constant $k_A$ is the promoter transcription initiation rate, and the parameters $k_A$, $k_R$, $k_B$, and $k_P$ are the respective kinetic association and dissociation rate constants for the complexes repressor/operator $O_R$ and polymerase/promoter $O_P$. 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 \( P_{\text{lac}} \), and \( P_{\text{tet}}^{\text{Smult}-1} \); while \( n = 2 \) for the other three promoters \( P_{\text{lac}}^{\text{dimer}-1} \), \( P_{\text{LacI}-1} \), and \( P_{\text{LtetO}-1} \) 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 \( n_k \) must be twice as large as the original \( k_4 \), 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 \( O_0 \) into the free one \( O_0 \). The reaction begins with a polymerase \( P \) recently bound to the operator \( O_{\text{closed}} \), 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 \( K_1 \), 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 only 24bp, instead of 29bp, we make the following assumption: When the promoter \( P_{\text{LtetO}-1} \) 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 \( O_0 \) is 25\%.

We make a different assumption for the promoters \( P_{\text{Allac}}^{\text{dimer}-1} \) and \( P_{\text{LacI}-1} \), 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 \( P_{\text{Allac}}^{\text{dimer}-1} \) and \( P_{\text{LacI}-1} \); 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 \( P(O_{2R}) \) and \( P(O_{2R}) \) be the probability that any one of the promoters \( P_{\text{Allac}}^{\text{dimer}-1} \) or \( P_{\text{LacI}-1} \) is at the state \( O_{2R} \) and \( O_{2R} \), respectively. We suppose that \( P(O_{2R}) \) is equal to \( \beta \) \( P(O_{2R}) \) in the presence of repressors, where \( \beta = 3 \) for the promoter \( P_{\text{LtetO}-1} \) 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 \( P_{\text{LtetO}-1} \). 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 \( P_{\text{Allac}}^{\text{dimer}-1} \) 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 \( P_{\text{Allac}}^{\text{dimer}-1} \) (or \( P_{\text{LacI}-1} \)), we made the following assumption.

**Assumption 2.** The cooperativity is negligible between the lac operators in the promoters \( P_{\text{Allac}}^{\text{dimer}-1} \) and \( P_{\text{LacI}-1} \), 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 promoter is free (state \( O \)), it forms the complex polymerase/promoter \( O_p \), and it forms the complex repressor/operator \( O_R \), respectively. Moreover, we also need the probability \( P(O_{2B}) \) that the promoter is at the state \( O_{2B} \). The state of equilibrium for the reactions (1) is easily given by the following identities, 
\[ P(O_D) = \frac{\eta k_d[R]}{k_u} \]  
\[ P(O_{2B}) = \frac{k_d[P]}{k_p + k_m} P(O_B). \]  
(2)

Since \( 1 = P(O_D) + P(O_{2B}) + P(O_B) + P(O_{1B}) \), we can easily deduce the following equation after dividing by \( P(O_D) \) and applying the above identities and the considerations made in Assumption 1.
\[ \frac{1}{P(O_{2B})} = 1 + \frac{\eta k_d}{k_p} \left( \frac{k_u + k_m}{k_u} \right) \left( \frac{[R]}{[P]} \right)^{1+\beta} + \left( \frac{k_p + k_b}{k_u} \right)^{m} \]  
(3)

We can easily verify that the last term \( \frac{k_p + k_b}{k_u} \) is really small (\( < 10^{-3} \)) with respect to the number 1, when we substitute the value of the parameters \( k_p, k_b, k_u, k_m \), and \( [P] \) enlisted in (13), (14), (15), or (16). Notice that \( 1M \) is equal to \( 1.8 \times 10^9 \) m(h mole 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 \( O_p \) is given by the formula
\[ P(O_p) = \frac{[P]}{[P] + m[R]} \]  
for
\[ m = 0(l+\beta) \frac{k_p + k_b}{k_u}. \]  
(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 \( P(O_p) \) obtained when \( [R] = 0 \) divided by the probability obtained when the concentration \( [R] \) is maximal; i.e.,

\[ \text{Regulatory Range} = 1 + \frac{k_d}{k_p} \left( \frac{k_u + k_m}{k_u} \right) \frac{[R]}{[P]} \]  
(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 deFai 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 \( P_{lac}, P_{Allan1-1}, P_{lac0-1}, \) and \( P_{lac0-1} \) against those ranges obtained from equation (5) by substituting the values 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. \( \text{gene oscillator} \)

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 \( \text{in silico} \) 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 $P_{lacO-1}$. 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 $P_{lacO+1}$, 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 $P_{lacO+1}$ operator,

$$
O_0 + R^{2} \xrightleftharpoons{k_{A}} O_K, \quad O_0 + P^{2} \xrightleftharpoons{k_{A}} O'_P,
$$

$$
O'_P \xrightleftharpoons{k_{D}} O_0 + P + M. \quad (6)
$$

We do not include the state $O_{2R}$, because it is only needed for the promoter $P_{lacO-1}$. The value of the parameters $k_{A}, k_{D}, k_{B}, k_{gb},$ and $k_{mb}$ 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 lac promoter 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 lac repressor $R$ that can bind to two different operators [11],

$$
M \xrightleftharpoons{k_{E}} E + M \quad 2E \xrightleftharpoons{k_{D}} D \quad 2D \xrightleftharpoons{k_{R}} R \quad \text{translation, dimerization, tetramerization.} \quad (7)
$$

The parameter $k_{E}$ 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 l-isopropyl-$D$-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 $D$ increasing 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,

$$
D + I \xrightleftharpoons{k_{I}} D^I, \quad R + I \xrightleftharpoons{k_{I}} R^I, \quad R^I + I \xrightleftharpoons{k_{I}} R_{II}, \quad (8)
$$

where $D^I$ is a dimer contaminated with IPTG and $R^I$ (or $R_{II}$) 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,

$$
2D^I \xrightleftharpoons{k_{D}} R_{II} \quad \text{and} \quad D + D^I \xrightleftharpoons{k_{D}} R^I. \quad (9)
$$

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 $P_{lacO-1}$, rendering the transcription impossible, so that we add the following two reactions,

$$
O_0 + D \xrightleftharpoons{k_{D}} O_D \quad \text{and} \quad O_0 + R^I \xrightleftharpoons{k_{D}} O'_R. \quad (10)
$$

It is important to point out that the propensity $k_{A}$ for the formation of the complexes $O_{D}$ and $O'_R$ is half of the propensity $2k_{A}$ for the formation of the original complex $O_{K}$, 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,

$$
M \xrightarrow{\gamma_{M}} \emptyset, \quad E \xrightarrow{\gamma_{E}} \emptyset, \quad D \xrightarrow{\gamma_{D}} \emptyset, \quad D^I \xrightarrow{\gamma_{D}} \emptyset, \quad R \xrightarrow{\gamma_{R}} \emptyset, \quad R^I \xrightarrow{\gamma_{R}} \emptyset, \quad R_{II} \xrightarrow{\gamma_{R}} \emptyset. \quad (11)
$$

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).

$k_{A} = 233.3$ (mpb/min), $k_{B} = 10$ (min), $k_{B} = 0.0372$ (min), $k_{D} = 2 = 0.018$ (mpb/min), $k_{D} = 0.4667$ (mpb/min), $k_{E} = 1.53 \times 10^{-3}$ (mpb/min).
\[ k_a = 0.008/\text{min}, \quad k_d = k_e = k_i = 0/\text{min}, \]
\[ k_b = 2/\text{min}, \quad k_p = 0.693/\text{min}, \]
\[ [P] = 830\text{mpb}, \quad [I] = 0.0038/\text{min}, \]
\[ [J] = 10^{10} \text{mpb}, \quad [E] = 0.375/\text{min}. \]

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 reactions [19]. The pseudocode for the rejection algorithm (SSA) able to simulate the correct time evolution of a chemical reacting system that involves delayed reactions [17, 18] and which was proved by Cai to be an exact stochastic simulation method of Barrio et al., which is a modification 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 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: \( O_0, O_P, O_{D1}, O_K, O_K^2, P, M, E, D^2, R, R^2, R^2 \) 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 \( \beta \neq 0 \) 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 \( \chi_i(t) \) with the rejection method of Barrio et al. described in Appendix F. Each trajectory \( \chi_i(t) \) 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., \( \{C_k(n)\} \) are the coefficients obtained from the fast Fourier transform of the sampled trajectory \( \chi_i(t) \), the average power spectrum is given by the formula:

\[ C(n) = \frac{1}{N} \sum_{k=1}^{N} |\chi_k(n)|^2 \quad \text{for} \quad n = 0, 1, 2, \ldots, 299, \]

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 range of the promoters \( P_{\text{lac}}, P_{\text{lacO}-1}, P_{\text{lacO}-2}, \) and \( P_{\text{lacO}-1} \). The error is less than 7.7% in the four cases. Recall that the parameter \( \alpha \) 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 \( P_{\text{lacO}-1} \) for which we assume that \( \beta = 3 \).

3.1. Regulatory range of the lac promoter \( P_{\text{lac}} \)

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 \( P_{\text{lac}} \) and a concentration of 30 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.

\[ k_a = 7 \times 10^{10} \text{M}^{-1} \text{s}^{-1}, \quad k_d = 6.2 \times 10^{-4}/\text{s}, \]
\[ k_b = 2 \times 10^{10} \text{M}^{-1} \text{s}^{-1}, \quad k_p = 0.0038/\text{min}, \]
\[ [P] = 830\text{mpb}, \quad [I] = 0.0038/\text{min}, \quad [J] = 10^{10} \text{mpb}, \]  
\[ n(1 + \beta) = 1. \]

As we said in the introduction, Oehler, Eismann et al. also measured in vivo a regulatory range of 18 for the promoter \( P_{\text{lac}} \) 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 \([R] = 4 \text{mpb}\) 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 $P_{A\text{laczO}}$

Lutz and Bujard measured in vivo a regulatory range of 350 for the promoter $P_{A\text{laczO}}$, when controls the luciferase reporter gene in a pZ vector [1, p. 1205]. This promoter is build from the bacteriophage 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 promoter [21, p. 1260]. We obtain the value 372 from equation (5) after substituting the value of the parameters below; see Appendix B.

$$k_A = 7 \times 10^5 / M/s, \quad k_B = 6.2 \times 10^{-4} / s, \quad \beta = 0.09 / s, \quad \alpha = 0.033 / s,$$

$$[R] = 90 \text{ milleHertz}, \quad [P] = 8300 \text{ milleHertz}.$$ (14)

$$k_A = 6 \times 10^2 / M/s, \quad k_B = 6.2 \times 10^{-4} / s, \quad \beta = 0.09 / s, \quad \alpha = 0.033 / s,$$

$$[R] = 90 \text{ milleHertz}, \quad [P] = 8300 \text{ milleHertz}.$$ (15)

3.3. Regulatory range of promoter $P_{L\text{tetO}}$

Lutz and Bujard measured in vivo a regulatory range of 620 for the promoter $P_{L\text{tetO}}$, 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 promoter [21, p. 1260]. We obtain the value 595.9 from equation (5) after substituting the value of the parameters below; see Appendix C.

$$k_A = 4.81 \times 10^2 / M/s, \quad k_B = 6.2 \times 10^{-4} / s, \quad \beta = 0.09 / s, \quad \alpha = 0.033 / s,$$

$$[R] = 90 \text{ milleHertz}, \quad [P] = 8300 \text{ milleHertz}.$$ (16)

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 milleHertz for IPTG’s concentrations of 100mM and 150mM, respectively. 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 milleHertz 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 $P_{\text{LlacO}^{-1}}$ and the concentration of IPTG is: a) 0 mbp, b) 50 mbp, c) 100 mbp, d) 150 mbp, e) 200 mbp, f) 250 mbp. 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 100 mbp and 150 mbp, respectively.

The intensity 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 $P_{\text{LlacO}^{-1}}$; 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 $P_{\text{LlacO}^{-1}}$ 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 $P_{\text{LlacO}^{-1}}$ has been substituted by $P_{\text{lac}}$ or $P_{\text{A1lacO}^{-1}}$. 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 $P_{\text{LlacO}^{-1}}$ is substituted by $P_{\text{A1lacO}^{-1}}$, 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 $P_{\text{LlacO}^{-1}}$ is substituted by $P_{\text{lac}}$. In Fig. 4 we present four average power spectra calculated with the promoter $P_{\text{lac}}$ 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.
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_{A1}^{lacO-1}, P_{LlacO-1}, P_{LtetO-1}, and P_{N25tetO-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_{A1}^{lacO-1}, P_{LlacO-1}, and P_{LtetO-1}, when one substitutes the given parameters by the corresponding chemical kinetic constants measured in vitro by several authors. In particular, the probability \( P(\Omega_c^R) \) that a polymerase forms a closed complex with the respective promoter can be expressed as the Hill function introduced in (4); i.e.,

\[
P(\Omega_c^R) = \frac{[P]}{[P] + m[R]}
\]

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):

\[
\begin{align*}
    P_{lac} & \quad P_{A1}^{lacO-1} & \quad P_{LlacO-1} & \quad P_{LtetO-1} \\
    m & \quad 3.5 \times 10^3 & \quad 1.4 \times 10^4 & \quad 5.9 \times 10^3 & \quad 7.0 \times 10^3 \\
\end{align*}
\]

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-1} is substituted by P_{A1}^{lacO-1}. Recall that Stricker et al. oscillator is based on a pZ vector, which contains a copy of the lac I structural genes controlled by the promoter P_{LlacO-1}. Moreover, no oscillations must be observed if the promoter P_{LlacO-1} is
substituted by \( P_{A\text{lAcO}-1} \). 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 \( P_{A\text{lAc}} \) [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 \( \text{LacI} \) is \( 4 \text{mpb} \) in Oehler et al. experiments, while the concentration is \( 90 \text{mpb} \) in Lutz and Bujard experiments.

7. Oehler, Amouyal et al. also measured in vivo a regulatory range greater than \( 12 \times 10^3 \) for a synthetic \( \text{lac} \) promoter that includes two functional \( \text{lac} \) 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 \( \text{lac} \) 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 \( P_{A\text{lAcO}-1} \) and \( P_{\text{lAcO}-1} \) the \( \text{lac} \) 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 vivo for the promoter \( P_{A\text{lAc}} \)

1. Complex (repressor \( \text{LacI} \))/operator formation rate constant \( k_A \); Riggs et al. measured \( 7(\pm0.9) \times 10^3 /\text{M/s} \) in vitro [22, p. 408]. We take \( k_A = 7 \times 10^3 /\text{M/s} \).
2. Complex (repressor \( \text{LacI} \))/operator dissociation rate constant \( k_d \); Riggs et al. measured \( 6.2(\pm1.3) \times 10^{-7} /\text{s} \) in vitro [22, p. 404]. We take \( k_d = 6.2 \times 10^{-7} /\text{s} \).
3. Complex polymerase/(promoter \( P_{A\text{lAc}} \)) formation rate constant \( k_B \). Lanzer and Bujard state the rate \( 2(\pm1.3) \times 10^9 /\text{s} \) [23, p. 8975]. We take \( k_B = 2 \times 10^9 /\text{s} \).
4. Complex polymerase/(promoter \( P_{A\text{lAc}} \)) 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 \( k_B = \frac{\ln2}{3.5 \text{ minutes}} = 0.0033 /\text{s} \).
5. Transcription initiation rate \( k_m \). The polymerase opens the DNA in the second transcription step. Malan et al. measured an opening rate of \( 0.003 /\text{s} \) for the \( \text{lac} \) wild-type promoter in presence of CRP-cAMP [25, p. 891], so we take \( k_m = 0.003 /\text{s} \).
6. Number of repressors \( \text{LacI} \) 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 \( P_{A\text{lAcO}-1} \)

1. Complex repressor/operator formation rate constant \( k_A \); see Appendix A.
2. Complex repressor/operator dissociation rate constant \( k_d \); see Appendix A.
3. Complex polymerase/promoter \( P_{A\text{lAcO}-1} \) formation rate constant \( k_B \). Lanzer and Bujard measured \( 6(\pm0.5) \times 10^3 /\text{M/s} \) for the promoter \( P_{A\text{lAcO}-1} \) with a \( \text{lac} \) operator located between the −10 and −33 hexamers [23, p. 8975]. We take \( k_B = 6 \times 10^3 /\text{s} \).
4. Complex polymerase/promoter \( P_{A\text{lAcO}-1} \) dissociation rate constant \( k_B \). Brunner and Bujard measured \( k_B = 9.2 \times 10^{-8} /\text{s} \) for the promoter \( P_{A\text{lAcO}-1} \) [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 \( 0.09 /\text{s} \) for the bacteriophage promoter \( P_{P1} \) [28, p. 1618], so we take \( k_m = 0.09 /\text{s} \).
6. Number of repressors \( \text{LacI} \) per cell \([R]\). Müller et al. state that the mutant promoter \( \text{lac}^{\text{D}} \) induces between 80 and 100 repressors/cell [21, p. 1260]. We take \([R] = 90 \) repressors/cell.
7. Number of polymerases per cell \([P]\). We obtain \([P] = 8300 \) polymerases/cell; see Appendix A.
8. Appendix C: Parameters measured in vitro for the promoter P_{lacO-1}

1. Complex repression/operator formation rate constant \( k_A = 7 \times 10^8 / M/s \); see Appendix A.
2. Complex repression/operator dissociation rate constant \( k_{D} = 6.2 \times 10^{-3} / s \); see Appendix A.
3. Complex polymerase/promoter P_{lacO-1} formation rate constant \( k_B \). Lanzer and Bujard measured 1.4(\pm 0.2) \times 10^{10} / M/s for the promoter \( P \) with a lac operator located between the \(-10 \) and \(-33 \) hexamers [23, p. 8975]. We take \( k_B = 1.4 \times 10^{10} / M/s \).
4. Complex polymerase/promoter P_{lacO-1} dissociation rate constant \( k_{D} \). Brunner and Bujard measured \( k_{D} = 6.8 \times 10^{-2} / s \) for the promoter \( P \) [27, p. 3141].
5. Transcription initiation rate \( k_m \). McClure observed a delay of 68s in the opening of complex polymerase/promoter \( P \) [29, p. 5636]. We take \( k_m = \frac{1}{68} \).
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_{lacO-1}

1. Complex (repressor TetR)/operator equilibrium binding constant \( k_A / k_D \). Hillen et al. measured a value between 10^{12} / M and 10^{15} / M [13, p. 715]. We take \( k_A / k_D = 5 \times 10^{12} / M \).
2. Complex (repressor TetR)/operator dissociation rate constant \( k_D \). Hillen et al. measured a half-life of 12 min. [13, p. 715]. We obtain \( k_D = \frac{0.693}{12} \approx 9.62 \times 10^{-2} / s \).
3. Complex polymerase/promoter formation rate constant \( k_B = 1.4 \times 10^{10} / M/s \); see Appendix C.
4. Complex polymerase/promoter dissociation rate constant \( k_{D} = 6.8 \times 10^{-2} / s \); see Appendix C.
5. Transcription initiation rate \( k_m = \frac{1}{60} \); 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 \( \mu \text{m}^3 \) [30]. We take a mean volume of 3 \( \mu \text{m}^3 \), so that we get the conversion

\[
1 \text{M} = \frac{6.0221 \times 10^{23} \text{molecules}}{\text{M} \text{bacteria}} \times 3 \times 10^{-15} \mu \text{m}^3
\]

\[
= 1.8 \times 10^{6} \text{mpb.}
\]

- Number of lacI structural genes per cell. There is only one lacI gene in the corresponding pL vector. Stricker et al. (2008) state that there are between one and three pL vectors per cell [4], so we take two structural lacI genes per cell.

- Translation initiation rate for the lactose repressor monomer \( k_I \). Chen et al. state the value \( k_I = 10 / \text{min} \) [31].

- Dimerization and tetramerization formation rate constants \( k_D \) and \( k_E \). We use the values \( k_D = 0.018 / \text{mpb} / \text{min} \) given by Stricker et al. in [4] (suppl. material).

- Complex (inducer IPTG)/(repressor LacI) formation rate constant \( k_L \). Damany et al. measured in vitro the value \( k_L = 4.6 \times 10^{7} / \text{M} / \text{s} \) [32].

- Dimerization and tetramerization dissociation rate constants \( k_{D} \) and \( k_{E} \). They seem to be negligible in comparison to \( k_D = k_E \), so that we take \( k_D = k_{E} = 0 / \text{mpb}/\text{min} \).

- Complex (inducer IPTG)/(repressor LacI) dissociation rate constant \( k_F \). It seems to be negligible in comparison to \( k_D \), so that we also take \( k_F = 0 / \text{mpb}/\text{min} \).

- Degradation rate constant for the lacI mRNA: \( \gamma_{l} \). Varmus et al. measured a half-life time of approximately one minute [33, p. 2264], so that \( \gamma_{l} \approx \frac{0.693}{10} \).

- Degradation rate constants for the repressor LacI: \( \gamma_{L} \), \( \gamma_{D} \), and \( \gamma_{E} \). 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$^{-1}$ [31].

- **Time delay in the translation of the lac mRNA molecule** $\tau_{cl}$. We assume that $\tau_{cl} = 0$, 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** $\tau_t$. 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 $\tau_t \approx \frac{360 \text{ amino acids}}{12 \text{ residues per second}} = 30 \text{ s}$.  

### 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:

$$X \xrightarrow{a_j(X)} X + v(j), \quad j = 1, 2, \ldots, m,$$  

where $a_j(X) \geq 0$ denotes the propensity of the $j$-th chemical reaction, and $v(j) \in \mathbb{N}^n$ 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 pseudo-code 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 $\bar{a} := \sum_{j=1}^{m} a_j$, where $a_j := a_j(X(t))$ is the propensity function for the $j$-th reaction in (17). Generate the waiting time $\tau = -\ln(a_1)/\bar{a}$ from a standard uniform random variable $a_1$ in the interval $[0, 1]$.  
3. Generate a standard uniform random variable $a_2$ in the interval $[0, 1]$ and choose the index $\mu = 1, 2, \ldots, m$ such that $\sum_{j=1}^{\mu} a_j < a_2 \leq \sum_{j=1}^{\mu+1} a_j$.
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)

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

(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 $\bar{a} := \sum_{j=1}^{m} a_j$, where $a_j := a_j(X(t))$ is the propensity function for the $j$-th reaction in (17) and (18). Generate the waiting time $\tau = -\ln(a_1)/\bar{a}$ from a standard uniform random variable $a_1$ in the interval $[0, 1]$. If there is no delayed reaction (18) to finish in the time interval $[t, t+\tau)$, find the time $t_\delta \in [t, t+\tau)$ of the first delayed reaction to finish, discard $\tau$, set the state vector $X(t_\delta) = X(t) + v[0]$, update the time $t = t_\delta$, 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 $a_2$ in the interval $[0, 1]$ and choose the index...
\[ \mu = 0, 1, \ldots, m \text{ such that } \sum_{i=0}^{t} a_i < \mu \leq \sum_{i=0}^{t+1} a_i. \] If \( \mu = 0 \), add the ending time \( t + t_1 \) to the list of delayed reactions that are in process of finishing. If \( \mu > 0 \), set the state vector \( X(t + \tau) = X(t) + \mu \).

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 TetrR/O and AraC/I1-I2 regulatory elements, Escherichia coli towards a synthetic pathway using a metabolic toggle switch, in: Metabolic flux redirection from a central metabolic pathway toward a synthetic pathway using a metabolic toggle switch, Methods Enzymol 23 (2014), 175–184.

[2] S. Orteh, E.R. Eismann, H. Kramer and B. Müller-Hill, The three operators of the lac operon cooperate in repression, EMBO J 4 (1985), 973–979.

[3] S. Orteh, M. Amoroso, P. Kielhof, B. von Wüstenberg and B. Müller-Hill, Quality and position of the three lac operators of E.coli define efficiency of repression, EMBO J 3 (1984), 3148–3155.

[4] L.M. Bond, J.P. Peterson, N.A. Becker, J.D. Kahn and L.J. Muller III, Gene expression by minimal lac loops in vivo, Nucleic Acids Res 32 (2004), 8072–8082.

[5] M. Lewis, The lac repressor, C.R.Biol 320 (2005) 521–546.

[6] S. Müller, S. Orteh and B. Muller-Hill, Repression of lac promoters as a function of distance, phase and quality of an auxillary lac operator, J Mol Biol 285 (1999), 21–29.

[7] W. Hillen, C. Gatz, L. Altschmied, K. Schollmeier and M. Meier, Control of expression of the Tn10-encoded tetracycline resistance genes, J Mol Biol 160 (1983), 797–821.

[8] Uniprot.org: Tetracycline repressor protein TetR, http://www.uniprot.org/uniprot/P99X64

[9] C.G. Zamora-Chimal and E.S. Zeron, A simple model for Lutz and Bujard's controllable promoters.