Journal of Computational Biology: http://mc.manuscriptcentral.com/liebert/jcb

Determining the Transcription Rates Yielding Steady State Production of mRNA in the lac Genetic Switch of Escherichia coli

Journal:	*Journal of Computational Biology*
Manuscript ID	Draft
Manuscript Type:	Original Paper
Keyword:	biochemical networks, RNA, GENE EXPRESSION, Statistical models
Manuscript Keywords (Search Terms):	Escherichia coli, kinetic model, mRNA, steady state, transcription
Abstract:	In order to elucidate the regulatory dynamics of the gene expression activation and inactivation, an in silico biochemical model of the lac circuit in Escherichia coli was used to evaluate the transcription rates which yield the steady state mRNA production in active and inactive states of the lac circuit. This result can be used in synthetic biology applications to understand the limits of the genetic synthesis. Since most genetic networks involve many interconnected components with positive and negative feedback control, intuitive understanding of their dynamics is often difficult to obtain. Although the kinetic model of the lac circuit considered involves only a single positive feedback, the developed computational framework can be used to evaluate supported ranges of other reaction rates in genetic circuits with more complex regulatory networks. More specifically, the inducible lac gene switch in E. coli is regulated by unbinding and binding of the inducer-repressor complexes to or from the DNA operator to switch the gene expression on and off. The dependency of mRNA production at steady state on different transcription rates and the repressor complexes has been studied by computer simulations in the Lattice Microbe software. Provided that the lac circuit is in active state, the transcription rate is independent of the inducer-repressor complexes present in the cell. In inactive state, the transcription rate is dependent on the specific inducer-repressor complex bound to the operator that inactive the gene expression. We found that the repressor complex with the largest affinity to the operator yields the smallest range of the feasible transcription rates to yield the steady state while the lac circuit is in inactive state. On the other hand, the steady state in active state can be obtained for any value of the transcription rate.

Determining the Transcription Rates Yielding Steady State Production of mRNA in the lac Genetic Switch of Escherichia coli

Komlan Atitey

College of Engineering, Swansea University, Swansea SA1 8EN, UK

E-mail: 885242@swan.ac.uk

Pavel Loskot, Corresponding author

College of Engineering, Swansea University, Swansea SA1 8EN, UK

E-mail: p.loskot@swan.ac.uk

Telephone: +44 1792 602619

Fax: +44 1792 295676

Paul Rees

College of Engineering, Swansea University, Swansea SA1 8EN, UK

E-mail: p.rees@swan.ac.uk

Keywords

Escherichia coli; kinetic model; mRNA; steady state; transcription

Abstract

In order to elucidate the regulatory dynamics of the gene expression activation and inactivation, an in silico biochemical model of the lac circuit in Escherichia coli was used to evaluate the transcription rates which yield the steady state mRNA production in active and inactive states of the lac circuit. This result can be used in synthetic biology applications to understand the limits of the genetic synthesis. Since most genetic networks involve many interconnected components with positive and negative feedback control, intuitive understanding of their dynamics is often difficult to obtain. Although the kinetic model of the lac circuit considered involves only a single positive feedback, the developed computational framework can be used to evaluate supported ranges of other reaction rates in genetic circuits with more complex regulatory networks. More specifically, the inducible lac gene switch in E. coli is regulated by unbinding and binding of the inducer-repressor complexes to or from the DNA operator to switch the gene expression on and off. The dependency of mRNA production at steady state on different transcription rates and the repressor complexes has been studied by computer simulations in the Lattice Microbe software. Provided that the lac circuit is in active state, the transcription rate is independent of the inducer-repressor complexes present in the cell. In inactive state, the transcription rate is dependent on the specific inducer-repressor complex bound to the operator that inactive the gene expression. We found that the repressor complex with the largest affinity to the operator yields the smallest range of the feasible transcription rates to yield the steady state while the lac circuit is in inactive state. On the other hand, the steady state in active state can be obtained for any value of the transcription rate.

1 Introduction

The gene expression is commonly controlled at the transcription initiation stage by the transcription factors dedicated for that purpose. These transcription factors function either by repressing the transcription inactivation (negative control), or activation (positive control), or both (Jacob and Monod (1961)). It is well-known that the DNA transcription is realized by a multi-subunit DNA-dependent RNA polymerase (RNAP) (Burgess (1969; 1971)). The recruitment of RNAP to the DNA promoter, and its isomerization to a competent open complex are the first two important regulatory steps in the gene transcription (McClure (1980); Ptashne and Gann (1997); Mekler et al. (2002)). However, the rate-limiting steps governing multi-step processes of transcription initiation and the kinetics of transition between active and inactive states of the gene expression have not been elucidated (Tang et al. (2009)). The cell DNA, in general, contains more than one operator, and only a fraction of those are blocked by the bounded repressor at any given time (Reznikoff et al. (1974); Ohler et al. (1990)). The synthesized mRNA serves as a template for the translation into a protein. The transcription termination is essential for the accurate gene expression and removal of RNAP at the end of the transcription unit (Ananya et al. (2016)). The transcript termination efficiency is strongly dependent on the rate of transcription (McDowell et al. (1994)). Although the transcription is a complex multi-step process with the transcription initiation often representing the rate limiting step, here, the whole transcription is represented as a single step having a certain transcription rate. More detailed model of transcription based on the molecular structures is presented in (zuo and Steitz (2016)). Moreover, in E. coli, the transcription and translation are usually not coupled as the translation occurs on free mRNA diffused to the ribosome-rich cytoplasm away from the DNA nucleoid region (Bakshi et al. (2012)). The ribosomes may reduce the chances of RNAP backtracking, pausing and even stalling during the elongation, so the overall time of elongation is greatly reduced (Proshkin et al. (2010)). In addition, the first trailing ribosome appears to assist the

RNA transcription, so it may be responsible for a precise match between the rates of transcription and translation (Proshkin et al. (2010)).

The recent theoretical and experimental studies investigating the behavior of the lac genetic circuit in E. coli assume that it is a stochastic system which randomly fluctuates between active and inactive states (Mettetal et al. (2005); Stamatakis and Mantzaris (2009)). For instance, the stochastic analyses of the inducible lac genetic switch in E. coli for well-stirred and spatially resolved models under the slow and fast-growth conditions were performed in (Roberts et al. (2011)). In the E. coli lac operon, a separate regulatory gene (*Lac I*) encodes the lac repressor which forms a positive feedback loop controlling the operon (Ma (2004); Russell et al. (2008); Esmaeili et al. (2015)) (Figure 1). The bacteria uses variety of mechanisms to direct RNAP to specific promoters in order to activate the transcription in its response to the environmental signals (Chen et al. (2010); Lee et al. (2012)). The initiation of transcription of the lac operon in E. coli by RNAP (more precisely, RNAP type II) is inhibited by binding of the lac repressor to a DNA operator site which overlaps with the lac promoter (Straney and Crothers (1985); Schlax et al. (1995); Davis et al. (2005)). In the absence of glucose, but in the presence of external lactose (exlact), the cellular metabolism of exlact is performed with the enzymes encoded by the *lacZ*, *lacY*, and *lacA* genes in the lac operon (Figure 1). Exlact acts as an inducer (e.g. allolactose), and can inactivate the repressor (Kalisky et al. (2007); Russell et al. (2008); Basan et al. (2015); Ray et al. (2016)). The binding between the inducer (monomer I, or dimer I_2) and the repressor (R_2) produces one of the three repressor species complexes (R_2, IR_2 or I_2R_2) (Figure 2) binding to and unbinding from the operator which modulates the rate of transcription (Roberts et al. (2011)). However, since binding and unbinding affinities of these repressor complexes are not accurately known, their stoichiometry is currently subject to debate (Oehler et al. (2006)). In bacteria, the transcript is terminated either by an intrinsic termination or by the Rho-dependent termination

(Ananya et al. (2016)). The intrinsic termination is mediated by signals directly encoded within the DNA template and the nascent RNA. The Rho-dependent termination relies upon the adenosine triphosphate-dependent RNA translocase Rho which binds the nascent RNA, and dissociates the elongation complex (Ananya et al. (2016)). The transcription termination can be enforced by NusA protein which binds and inhibit RNAP (Qayyum et al. (2016)). The degradation of mRNA in bacteria is driven by the ribonuclease (RNase) which renders the mRNA molecule incapable of acting as a template for further protein synthesis (Jain (2002); Wang et al. (2009)). An important role of RNase in the mRNA decay was first suggested by the studies of the total mRNA turnover, and subsequently confirmed by many other studies examining the breakdown of individual messages (Jain (2002); Kushner (2002)). More recent genomic analyses using microarrays have established that RNase is a major participant in the mRNA turnover process (Bernstein et al. (2004)). In addition, the initiation of the mRNA degradation is followed by the complete breakdown of mRNA to mononucleotides (Deutscher (2006)).

Figure 1: The E. coli lac operon. The enzymes taking part in lactose metabolism are encoded by the genes lacZ, lacY and lacA. The separate regulatory gene lacI encodes the lac repressor which controls the operon regulation. The promoter binds RNAP, and the repressor complexes bind the operator. The transcription unit which extends from the transcription initiation site to the transcription termination site contains the operon genes.

Figure 2: The inducer-repressor complexes. (A) Activation of the repressor R to the activated repressor R_2. (B) The activated repressor R_2 binds the inducer monomer (allolactose, I) to produce the repressor complex IR_2. (C) The activated repressor R_2 binds the inducer dimer (allolactose I_2) to produce the repressor complex I_2R_2. (D) The repressor complex IR_2 binds the inducer monomer (allolactose) to produce the repressor complex I_2R_2.

Mary Ann Liebert, Inc., 140 Huguenot Street, New Rochelle, NY 10801

2 Methods

2.1 Modeling and simulation framework

A complete well-stirred kinetic model of the lac genetic switch presented in (Roberts et al. (2011)) was adopted to carry out in silico experiments (Figure 3 and supplementary file). All biochemical reactions considered are reversible, so they have both the forward and the reverse rate. The kinetic model is simulated using the Lattice Microbe software (Roberts et al. (2013)). The latest version 2.3 was downloaded and compiled with a GPU support on the Fedora 25 Linux workstation. In brief, the Lattice Microbe software numerically solves the chemical master equation of a biochemical reaction network given initial concentrations and the reaction rate constants by generating a specified number of independent stochastic trajectories of the species counts. In order to obtain statistically meaningful data for evaluating the lac circuit dynamics, the time evolutions of mRNA were collated over at least 100 independent simulation realizations.

Figure 3: The minimum steady-state count of mRNA molecules synthesized while the circuit is in inactivate state. (A) The basal mRNA synthesized in the lac circuit locked in inactive state with the rate of transcription 7.87e-3s^{-1}. (B) The mRNA synthesized in active state of the lac circuit with the same transcription rate 7.87e-3s^{-1}. a, b, and c represent the circuit models containing only the inducer-repressor species R_2, IR_2 and I_2R_2, respectively.

Our in silico experiments utilize controlled binding and unbinding of the repressor complexes to and from the operator. In particular, by enforcing either active or inactive state of the gene expression and by assuming a specific inducer-repressor species, our aim is to determine a range of transcription rates which can be supported by the lac genetic circuit in order to maintain the steady state mRNA synthesis. We define steady-state mRNA production as having relatively small variations about the mean mRNA count (Figure 5.1) in (Erban et al. (2007)). It should be noted that the transcription rate

can have two different meanings. By default, we assume that it is a rate parameter of the transcription reaction, but it can also express the number of mRNA molecules synthesized per unit of time. The determined ranges of supported transcription rates are indicative of binding and unbinding affinities of the inducer-repressor species to and from the operator. We propose a simple formula to quantify the repression efficiency as a measure of binding affinity of the repressor complexes to the operator. Even though the binding affinity of the repressor to the operator modulated by the inducer concentration can be expressed accurately by utilizing the constants of the dissociation and association processes (Roberts et al. (2011)), our formula is simpler, and it depends only on the determined maximum and minimum values of the supported reaction rates. Since the dynamics of the cell response occur over the cell lifetime (Nath and Koch (1970)), all simulations were performed over one hour of the E. coli cell time. The time of complete mRNA degradation is assumed to be negligible compared to the duration of inactive state. The time in all figures is expressed in seconds.

2.2 Determining the feasible transcription rates in inactive state

The initiation and completion of inactive state of the lac genetic circuit does not occur spontaneously. The inactive state resumes by binding of the inducer-repressor species to the operator. Different inducer-repressor species can display vastly differing binding and unbinding rate constants (Xue and Yeung (1995); Lu et al. (1998); Zhuang et al. (2000)). The transcription in inactive state is inefficient, so its rate is greatly reduced, although it remains non-zero (Russell et al. (2008); Abhyudai (2013)). The steady state mRNA count synthesized in inactive state of the lac genetic circuit can be followed by a full degradation of the mRNA synthesized as a response after the basal mRNA synthesis has ended (Mettetal et al. (2005)). On the other hand, the transcription rate abruptly increases when the repressor complex unbinds from the operator, and the lac circuit switches to

active state. The transcription rates in active and inactive state can be measured in vivo and in vitro under the precise environmental conditions. In in silico models, the transcription rate is a time constant parameter which is not precisely known. This uncertainty can be evaluated as a range of reaction rate values which can be supported by the genetic circuit and its regulatory mechanism. In particular, we observed that the range of transcription rates in active and inactive state cannot be arbitrary in order to maintain the steady state mRNA production modulated by the lac circuit regulation. This has been confirmed by our numerical experiments. Hence, our aim is to determine the interval $[k_{tr\min}, k_{tr\max}]$ of the supported transcription rates k_{tr} which can be supported in active and inactive state of the lac circuit. We vary otherwise constant transcription rate values in the kinetic model while all other reaction rates have their default value while observing whether the mRNA production reaches a steady state. Furthermore, we investigated how the supported transcription rates are affected by the specific inducer-repressor complexes. Therefore, we modified the full kinetic model of the lac switch from (Roberts et al. (2011)) to enforce that there is only one inducer-repressor species present in the cell by removing the corresponding chemical reactions from the kinetic model. The models containing either only species R_2, IR_2, or I_2R_2 are denoted as Model 1, 2 and 3, respectively. The binding and unbinding reaction rates of the inducer-repressors species to and from the operator are given in Table 1 including their default reaction rates. The default transcription rate is denoted as $k_{tr\mathrm{d}}$.

2.2.1 Determining the minimum feasible transcription rate in inactive state

Even in uninduced state (i.e., in the absence of the repressor inducer in the cell), approximately one mRNA molecule (Figure 3) can be transcribed in the cell until the cell division (Tropp (2011); Abhyudai (2013)). Such so-called basal synthesis can occur, since binding of the repressor to the operator is never infinitely strong (Tropp (2010)). The repressor can temporarily, for a short period of time, come off the operator before it rebinds again while the cellular crowding keeps the repressor in

Mary Ann Liebert, Inc., 140 Huguenot Street, New Rochelle, NY 10801

the vicinity of the operator. During such events, the already bound RNAP may initiate the transcription (Solomon et al. (2005)). Hence, we can assume that the minimum transcription rate $k_{tr\min}^{inac}$ in inactive state corresponds to the transcription rate of the basal synthesis equal to one mRNA molecule transcribed over the cell lifetime (Figure 3).

2.2.2 Determining the maximum feasible transcription rate in inactive state

The inactive state is completed when the operator is cleared, so RNAP can much more readily bind to the promoter and initiate the transcription across the operator. Hence, the transcription rate in active state is significantly increased. By comparing the mRNA synthesis in inactive and active transcription states in Figure 4, we have devised the following two step procedure to determine the maximum supported transcription rate in inactive state of the lac circuit.

Figure 4: The maximum steady-state count of mRNA molecules synthesized while the circuit is in inactivate state. (A) The basal mRNA synthesized in the lac circuit locked in inactive state with the transcription rate 0.13e-1s^{-1}. (B) The mRNA synthesized in active state of the lac circuit with the same transcription rate 0.13e-1s^{-1}. a, b, and c represent the circuit models containing only the inducer-repressor species R_2, IR_2 and I_2R_2, respectively.

1) Initialization: Search the initial transcription range R_0.

Let $R_0 \rightarrow [k_{tr\min}^{inac}, k_{tr\max}^{inac}(0)]$ where the initial value $k_{tr\max}^{inac}(0) = (1+\alpha)k_{tr\min}^{inac}$, and for the lac circuit considered, we assumed α=*20%*. Generate 100 uniformly spaced and sorted values of the transcription rates k_{tr} from R_0. Starting from the smallest sample of transcription rate, we determine whether a complete mRNA degradation occurs during the protein lifetime, and so the basal mRNA

synthesis no longer occurs. If the full mRNA degradation did not occur, we consider the next larger sample of k_{tr}, otherwise we set $k_{tr\max}$ to be the current value of k_{tr} and continue to the next step.

2) Iterations: Search the sequence of the transcription ranges R_n.

Let $R_n \to [k_{tr\max}^{inac}(\mathrm{n}-1), k_{tr\max}^{inac}(\mathrm{n})]$ where $k_{tr\max}^{inac}(n) = k_{tr\min}^{inac} + n\alpha k_{tr\max}^{inac}(0)$, and again, α=*20%*. For every *n=1,2,…*, generate 100 uniformly spaced and sorted samples of the transcription rates from R_n. Starting from the smallest sample of k_{tr}, we determine the first value of k_{tr} when the full mRNA degradation occurs during the cell lifetime, and then set $k_{tr\max}$ to be equal to this value, otherwise the next large sample of k_{tr} is considered. This process is repeated for all three inducer-repressor species considered independently. The obtained maximum transcription rates in inactive state of the lac circuit corresponding to the species R_2, IR_2 and I_2R_2, respectively, are denoted as $k_{tr\max}^{inac1}$, $k_{tr\max}^{inac2}$ and $k_{tr\max}^{inac3}$.

However, the value of $k_{tr\max}$ determined by this procedure is random, so it shows small fluctuations (a variance) due to the intrinsic noise of the simulated stochastic kinetic model (Cox et al.(2008)). Hence, we assume that the maximum transcription rate $k_{tr\max}$ supported in inactive state lies in the interval $[k_{tr\max}^{inac} - \varepsilon, k_{tr\max}^{inac} + \varepsilon]$ where $k_{tr\max}^{inac}$ is the average value obtained by repeating the measurement procedure above, and the value of ε*>0* has been set, so that the interval contains all the measured values of the maximum supported transcription rates in inactive state. In other words, the value of ε represents the range of transcription rates due to the basal mRNA synthesis which are followed by the complete degradation of the mRNA synthesized.

2.3 Determining the feasible transcription rates in active state

We claim that when the lac circuit is in active state, the transcription rate must be at least:

Mary Ann Liebert, Inc., 140 Huguenot Street, New Rochelle, NY 10801

$$k_{tr\min}^{ac} \geq k_{tr\max}^{inac} + \varepsilon \quad (1)$$

where the value of ε has been determined previously. In other words, if the transcription rate in active state would be smaller than the limiting value given in (1), then, with very high probability, the mRNA synthesis cannot reach steady state, for any value of the degradation rate. The ranges of feasible transcription rates in inactive and active states of the lac circuit are depicted in Figure 5. It should be noted that, for any genetic circuit, we always have, $k_{tr\min}^{ac} > k_{tr\max}^{inac}$ (Russell et al. (2008); Abhyudai (2013)). Moreover, the previous experimental results such as (Skinner et al. (2013)) and (So et al. (2011)) reported that, in active state, the lac circuit in E. coli can produce up to 50 mRNA molecules which simultaneously exist in the cell. Consequently, we can assume that the maximum transcription rate $k_{tr\max}^{ac}$ which can be observed in the lac circuit in active state corresponds to the maximum synthesis of 50 mRNA molecules before the degradation. However, we observed that the steady-state mRNA production in active state can be reached for any value of the transcription rate considered. This can be important in synthetic biology applications aiming to synthesize more than 50 mRNA molecules in the cell.

Figure 5: The range of feasible transcription rates in inactive and active state. The values $k_{tr\min}^{inac}$ and $k_{tr\max}^{ac}$ represent the minimum and the maximum feasible transcription rates in inactive and active state, respectively. The value k_{trmax} is the average maximum transcription rate measured in inactive state.

2.4 Binding affinities of the inducer-repressor species inducing the transcription repression

The repressor R_2 can be deactivated by binding of the inducer (e.g., allolactose) monomer I or dimer I_2 (Yagil and Yagil (1971)) (Figure 2):

$$R_2 + I \xrightarrow{k_{11}} R_2 I\,,\ R_2 + I_2 \xrightarrow{k_{12}} R_2 I_2\,,\ R_2 I + I \xrightarrow{k_{13}} R_2 I_2\,.$$

The binding of the repressor complex R_2I_n, $n=0,1,2$, to the operator O is the reaction (Yagil and Yagil (1971)):

$$O + R_2 I_n \xrightarrow{k} O R_2 I_n\,.$$

Furthermore, there are multiple operators in the cell DNA (Oehler et al. (2006)), so the total number of operators is a sum $O_{\text{tot}} = O_{\text{bounded}} + O_{\text{unbounded}}$ of the number of operators inactivated by the repressor complex and the number of repressor-free operators. The fraction of unbounded operators was shown to be a function of the inducer concentration $[I]$ (Oehler et al. (2006)):

$$f([I]) \approx O_{unbounded} / O_{tot} \qquad (2)$$

and the full repression of transcription can only occur when there are no inducers bound to the repressors (Oehler et al (2006)). Inspired by the dependency (2), we can deduce binding affinity of the inducer-repressor species to the operator in inactive state of the lac circuit from the observed feasible ranges of transcription rates which were determined in the previous subsection. In particular, we define the efficiency of transcription repression (TR) as:

$$TR = k_{tr\,\min}^{inac} \Big/ k_{tr\,\max}^{inac}\,. \qquad (3)$$

The TR values can be calculated for the different inducer-repressor species to allow a comparison of their binding affinities to the operator.

3 Results

In order to calculate the *TR* values which are indicative of the binding and unbinding affinities of the three inducer-repressor species considered, we first determined the ranges of transcription rates in inactive state of the lac genetic circuit. We also evaluate the steady state mRNA synthesis in active state of the lac circuit for various values of transcription rates.

3.1 Inferring binding affinity of the inducer-repressor species

The first important finding is that all three inducer-repressor species yield the same minimum transcription rate in inactive state of the lac circuit. Hence, the feasible transcription rate $k_{tr\min}^{inac}$ is independent of the particular inducer-repressor species considered. Using the formula (3), we confirmed that inducer reduces binding affinity of the repressor to the operator. Specifically, we observed that the effect of the dimer I_2 in reducing binding affinity of the repressor is larger than that of the monomer *I*. The measured transcription rate for synthesizing one molecule of mRNA in inactive state yields the minimum rate $k_{tr\min}^{inac} = 2^{-4} k_{trd} = 7.87 \times 10^{-3} s^{-1}$. This value was observed for all three inducer repressor species considered in Models 1, 2 and 3, respectively. Some examples of the minimum mRNA counts synthesized while the lac circuit is in inactive state are shown in Figure 3.

The maximum feasible transcription rates $k_{tr\max}^{inac}$ in inactive state corresponding to the three inducer-repressor species considered are given in Figure 4. Figure 6 then shows a full degradation of the mRNA synthesized in inactive state over the cell lifetime, and also the start of the mRNA synthesis when the lac circuit switches to active state. These results justify why the measured transcription rate k_{trmax} is the rate limit which separates active and inactive state of the lac circuit. In particular, the measured maximum transcription rates for Model 1 and 2 are $k_{tr\max}^{1} = 0.013 s^{-1}$ and $k_{tr\max}^{2} = 0.033 s^{-1}$,

respectively. By evaluating over 100 simulations, we determined the transcription rate uncertainty parameter ε to be $\varepsilon_1=0.008s^{-1}$ for Model 1, and $\varepsilon_2=0.012s^{-1}$ for Model 2, and thus, $k_{tr\max}^{inac1}=(0.13\pm0.08)\times10^{-1}s^{-1}$, and $k_{tr\max}^{inac2}=(0.33\pm0.12)\times10^{-1}s^{-1}$. In case of Model 3, we were unable to observe a complete degradation of the mRNA synthesized, even when the transcription rate was substantially increased, so we concluded that the maximum supported transcription rate in inactive state of the lac circuit is not limited by its regulatory mechanism, and have that, $k_{tr\max}^{inac3}$? $k_{tr\max}^{inac2}$? $k_{tr\max}^{inac1}$.

Figure 6: The minimum count of mRNA molecules synthesized in activate state of the lac circuit. (A) The full degradation of mRNA synthesized during the basal mRNA transcription in the lac circuit locked in inactive state and having the transcription rate 0.2e-1s^{-1}. (B) The mRNA synthesized at steady state in active state of the lac circuit with the same transcription rate 0.21e-1s^{-1}. a, b, and c represent the circuit models containing only the inducer-repressor species R_2, IR_2 and I_2R_2, respectively.

The measured rates $k_{tr\max}^{inac1}$ and $k_{tr\max}^{inac2}$ are used to compute the transcription repression TR_1 and TR_2 defined in (3) for Model 1 and Model 2, respectively. For Model 3, we cannot calculate the value of TR_3, although we can compare it to the values of TR_1 and TR_2 as shown in Table 2. The limiting transcription rates which were obtained independently for each inducer-repressor species can be combined to obtain an equivalent limiting value of the maximum transcription rate in inactive state for the full kinetic model of the lac circuit containing all three inducer-repressor species, i.e.:

$$k_{tr\max}^{inac}=\min_{i=1,2,3}k_{tr\max}^{inac,i}=k_{tr\max}^{inac1}=(0.13\pm0.08)\times10^{-1}s^{-1}. \quad (4)$$

This confirms that the inducer reduces binding affinity of the repressor to the operator. Furthermore, we observe from Figure 4 and Figure 6 that the basal synthesis of mRNA in inactive state can produce at most two molecules of mRNA.

3.2 Transcription activation efficiency of different inducer-repressor species

Unbinding of the inducer-repressor complex from the operator abruptly increases the transcription rate (Jain (2002); Russell et al. (2008)). The inducer-repressor species which are not bound to the free operator do not affect the mRNA synthesis in active state of the lac circuit. Assuming the results (1) and (4), we obtain the minimum transcription rate in active state of the lac circuit to be equal to:

$$k_{tr\min}^{ac} = k_{tr\max}^{inac} + \varepsilon = 0.013 + 0.008 = 0.021 s^{-1}.$$

This minimum transcription rate in active state can produce at least two mRNA molecules (Figure 6). On the other hand, assuming that the maximum number of mRNA molecules synthesized at steady state when the lac circuit is in active state is about 50 (Skinner et al. (2013)), the corresponding transcription rate is independent of the particular inducer-repressor species in the cell, and we get, $k_{tr\max}^{ac} = k_{trd} \times 2^2 = 0.504 s^{-1}$. Finally, having established the interval $\left[k_{tr\min}^{ac}, k_{tr\max}^{ac}\right]$ of the feasible transcription rates in active state for all three inducer-repressor species, we generate 100 uniformly distributed random samples of transcription rates from this interval. For each of these samples, we obtained the steady state mean count of mRNA for all three models considered (Figure 7). These results are summarized in Table 3. In addition to the mRNA counts corresponding to the minimum transcription rate $k_{tr\min}^{ac} = 0.021 s^{-1}$ and the maximum transcription rate $k_{tr\max}^{ac} = 0.504 s^{-1}$, respectively, Table 3 also presents 4 different mRNA counts corresponding to 4 selected transcription rates. We observe that, for all transcription rate samples considered, the mRNA counts in all three circuit

models are approximately equal. This confirms that the inducer-repressor species do not constitute a factor regulating the efficiency of transcription activation when the lac circuit is already in active state.

Figure 7: The steady-state count of mRNA synthesized in active state of the lac circuit. (A) The minimum feasible transcription rate 7.87e-3 s^{-1}, and (B) the maximum feasible transcription rate 0.504 s^{-1}. a, b, and c represent the circuit models containing only the inducer-repressor species R_2, IR_2 and I_2R_2, respectively.

4 Discussion

Many studies demonstrated that the protein-DNA interactions are central to control the gene expression in all forms of life (Oehler et al. (2006); Munro et al. (2016)). In E. coli, the genes for transcription regulation are negatively regulated (Jacob and Monod (1961); Xu et al. (2011)). The lac repressor in E. coli is constitutively expressed and binds to the upstream cis-activated operator. It subsequently blocks the transcription of genes which are necessary for the cell to digest lactose as the energy source. In this case, the negative regulation is relieved in the presence of a particular effector (e.g., allolactose) which acts as an inducer, and binds to the repressor to activate the expression of genes necessary for lactose metabolism. Understanding how the effector (inducer) molecules alter the binding properties of the repressor at the molecular level is essential for establishing a detailed understanding and modeling of the gene regulations (Daber et al. (2007)). The interactions between the inducer and the repressor yield different inducer-repressor complexes having different binding affinities to the operator (Roberts et al. (2011)). In this paper, we obtained the feasible transcription rates which can be supported by the lac genetic circuit. The limiting values of transcription rates in inactive state were used to deduce the binding efficiencies of the three inducer-repressors species in inactive state of the lac circuit. The inducer-repressor species were considered independently by modifying the full kinetic model to enforce the existing of only one inducer-repressor complex in the cell. In active state, the transcription efficiency can be measured as the steady-state of mean count of mRNA. We confirmed that the efficiency of the transcription initiation is independent of the specific inducer-repressor complex, so these complexes do not modulate the transcription while the circuit is in active state.

In general, the rate of transcription is affected by the ambient temperature (Kondo et al. (1993)), the growth media and its viscosity, the type and level of the inducer, concentration of NTPs, ribosome control, binding affinities of transcription factors, molecular crowding and diffusion rates, and other factors (Chung et al. (2017)). Consequently, the in vitro and in vivo transcription rates using the same molecular machinery may be quite different. The DNA generally contains regulatory sequences causing transcript pausing and eventually its termination which is one of the main transcription regulation factors (Tolic-Norrelykke et al. (2004)). Thus, the transcription elongation rates is affected by the matching DNA sequence (codons) as well as the presence of metallic cations influence RNAP grip to the DNA, so different genes can be transcribed at different rates. The DNA looping plays an important role in the gene regulation as it modulates the kinetics of the protein binding and unbinding (Vanzi et al. (2006)). The efficiency of termination is strongly dependent on the elongation rate (Tolic-Norrelykke et al. (2004)). However, even elongation of the same gene can experience relatively large differences in the elongation velocities at each pass which is attributed to different RNAP conformational states (Tolic-Norrelykke et al. (2004); Fuchs et al. (2014); Mejia et al. (2014)). The RNAP conformational changes modulating its activity and contacts with RNA substrate can be a bacterial resistance mechanism to adapt to the stress factors (Esyunina et al. (2016)). In addition, the overall elongation rate also strongly depends on the distribution and duration of the random pausing events which makes the elongation rate in each transcription cycle to be also random (Tolic-Norrelykke et al. (2004)). Furthermore, the maximum transcription rate is likely limited by the transcription fidelity (Mejia et al. (2014)). The elongation efficiency can be measured by how close the actual elongation rate is to the maximum observable elongation rate (Singh and Padgett (2009)).

Exploring the transcription rates which can be supported by the regulatory circuit in the lac genetic switch can facilitate understanding of the regulatory design of the gene expression (Bridges et al. (2005); Sharabiani et al. (2005); Daber et al. (2007)). The allolactose effector is a lactose isomer

Mary Ann Liebert, Inc., 140 Huguenot Street, New Rochelle, NY 10801

which acts as an inducer by binding and inactivating the repressor (Yang et al. (2015)). Regardless of the number of inducer molecules bound to the repressor, the repressor complex can always bind to the operator (Swint-Kruse (2004); Daber et al. (2007)). However, the number of inducer molecules bound to the repressor is an important transcription factor in inactive state. The stoichiometry of the inducer-repressor binding is currently subject to debate (Oehler et al. (2006)) and it is unclear whether affinity of IR_2 to the operator is of the same order as that of R_2 (Reznikoff et al. (1974)). Our numerical experiments revealed that the inducer-free receptor (R_2) is the most effective transcription repression factor among the three inducer repressors species (R_2, IR_2, and I_2R_2) in the regulation model of the lac genetic circuit. We measured the transcription rates in inactive state to deduce the binding affinity of the repressor complex to the operator. This confirmed the previous results about the inactivation effects of the inducer (Smith and Hanawalt (1969); Yang et al. (2015)). Furthermore, we extended the previous works by defining a range of transcription rates which are observed in the presence of the specific inducer-repressor species.

In synthetic biology applications, the biochemical components are assembled or used in new ways in order to produce the desired biological activity. One such practical application is controlling the regulation of the gene expression (Oehler et al. (2006)). In general, the transcriptional regulation allows the cell to allocate its valuable resources towards the production of desirable proteins in order to optimize the response to changes in the environmental conditions, and also to control the stoichiometry of enzymes, and to avoid futile cycles, producing underutilized enzymes, build-up of undesired or even toxic metabolites while not inhibiting the assembly of desired macromolecular structures (Oehler et al. (2006)). The RNA is recognized as a powerful biomolecule for controlling and engineering the cellular functions (Lei (2014); McKeague et al. (2016); Chappell et al. (2017)). Therefore, the capability to vary the reaction rate can be an effective method to control the level of

mRNA synthesis in the cell which affects other aspects of the gene expression (Chappell et al. (2013)). Our numerical experiments indicate that the basal transcription in inactive state can produce at most two mRNA molecules. Hence, irrespective of the actual value of the transcription rate, the lac circuit always produces mRNA at all times, so the corresponding protein can always be synthesized. We determined the minimum and maximum transcription rate values which are supported in active and inactive state, respectively (Figure 8). We found that the mRNA production in active state of the lac genetic circuit yields identical statistics for all three repressor complexes R_2, IR_2, and I_2R_2 considered. The mRNA steady-state count for the lac circuit comprising only one inducer-repressor complex can be then used to assess the efficiency of transcription for that particular inducer-repressor complex. Since the transcription rate is a good indicator of the transcription efficiency, we can conclude that the transcription efficiency in active state is independent of the particular inducer-repressor species present in the cell.

Figure 8: The range of the feasible transcription rates in inactive and active state of the lac genetic circuit. The red steps represent the number of mRNA molecules synthesized per unit of time (s^{-1}). The transcription rates in the interval [*0.00787, 0.021*] (s^{-1}) define the basal synthesis of mRNA in inactive state of the lac circuit while active state is only achievable when the transcription rate is at least 2.2e-3s^{-1}.

In general, the genetic circuits can be usually programmed to produce the desired response to a selected combination of the environmental signals at their input (Kobayashi et al. (2004)). The regulation of transcription elongation and transcription termination has been discussed in (Washburn and Gottesman (2015)). The environment-responsive promoters can be used to regulate the transcription via genetic switches (Khalil and Collins (2010)). The promoter function can be regulated by the ligand-inducible transcription factors. For example, the predictable control of transcription rates in on and off states can be achieved by modifying the natural E. coli promoters

Mary Ann Liebert, Inc., 140 Huguenot Street, New Rochelle, NY 10801

(Chen et al. (2017)). However, the bistable genetic switch with a single autocatalytic promoter considered here is less robust than the switch design with two coupled promoters and repressors in (Gardner (2000)), and the former is also more difficult to tune experimentally. The bistability conditions of the genetic switch in (Gardner (2000)) are dependent on the rates of synthesis of the two genes as verified using the actual E. coli plasmid. More generally, the synthetic transcription factors which usually exploit the ligand-controlled bacterial enzymes can be used to program the cell functional responses. In more complex eukaryotic cells, it is desirable to control the gene expression at several levels with transcription (especially at the initiation) and post-transcription control which also include the mRNA and protein stability (Aulander (2013)). However, the completely synthetic genetic circuits may create another problem of unintended interactions with other host systems (Lee et al.(2010)).

5 Conclusion

We measured the feasible transcription rates in inactive and active states of the lac circuit in E. coli. Modulating the transcription rate remains the principal factor controlling the mRNA production. In the lac switch model considered, the values of transcription rates which are affected by binding and unbinding of the inducer-repressor complexes can be used to deduce their binding affinities to the operator. We showed that repressor R_2 has the largest binding affinity, so it affects the transcription in inactive state the most. Consequently, the range of supported transcription rates in inactive state corresponding to R_2 is narrower than the supported transcription rate ranges corresponding to the other two inducer-repressor complexes. We did not find that unbinding rates of different inducer-repressor species affect the transcription rates in active state, so we conjecture that the inducer-repressor species do not act as transcription factors of the lac circuit when it is in active state.

Supplementary file

This file briefly describes the Lattice Microbe software, and also lists chemical reactions and the associated reaction rates comprising the kinetic model of the lac genetic switch considered in this paper.

Acknowledgements

The work of KA was supported by the Zienkiewicz Scholarship at Swansea University.

Authors' contributions

KA performed the analyses and drafted the manuscript. PL devised the concept, supervised the work, and prepared the manuscript for submission. PR provided valuable suggestions. All authors read and approved the final manuscript.

Author Disclosure Statement

No competing financial interests exist.

References

Abhyudai, S., Vargas, C. A., and Karmakar, R. (2013). "Stochastic analysis and inference of a two-state genetic promoter model", in: *American Control Conference (ACC).* (Washington, DC, USA IEEE).

Ananya, R., Bellecourt, M. J., and Landick, R. (2016). Mechanisms of Bacterial Transcription Termination: All Good Things Must End. *Annual Review Biochemistry* 85**,** 319-347.

Aulander, S., Fussenegger, M. (2013). From gene switches to mammalian designer cells: present and future prospects. *Cell press* 31(3).

Bakshi, S., Siryaporn, A., Goulian, M. et al. (2012). Superresolution imaging of ribosomes and RNA polymerase in live Escherichia coli cells. *MolecularMicrobiology* 85(1)**,** 21-38

Basan, M., Hui, S., Okano, H. et al. (2015). Overflow metabolism in Escherichia coli results from efficient proteome allocation. *Nature* 528(99).

Bernstein, J.A., Lin, P., Cohen, S.N. et al. (2004). Global analysis of Escherichia coli RNA degradosome function using DNA microarrays. *PNAS* 101(9)**,** 2758–2763.

Bridges, D., Fraser, M.E., and Moorhead, G.B. (2005). Cyclic nucleotide binding proteins in the Arabidopsis thaliana and Oryza sativa genomes. *BMC Bioinformatics* 6(6)

Burgess, R.R. (1969). Separation and Characterization of the Subunits of Ribonucleic Acid Polymerase. *The journal of Biological Chemestry* 244(November 25)**,** 6168-6176.

Burgess, R.R. (1971). RNA Polymerase. *Annual Review of Biochemistry* 40**,** 711-740.

Chappell,J., Takahashi, M.K., Meyer, S. et al. (2013). The centrality of RNA for engineering gene expression. *Biotechnology Journal* 8**,** 1379–1395.

Chappell, J., Westbrook, A., Verosloff, M. et al. (2017). Computational design of small transcription activating RNAs for versatile and dynamic gene regulation. *Nature Communication*. doi: 10.1038/s41467-017-01082-6.

Chen, J., Darst, S.A., and Thirumalai, D. (2010). Promoter melting triggered by bacterial RNA polymerase occurs in three steps. *PNAS* 107(28)**,** 12523-12528.

Chen, Y., Ho, J.M., Shis, D.L. et al. (2017). Tuning the dynamic range of bacterial promoters regulated by ligand-inducible transcription factors. *Nature Communication* 9(64).

Chung, S.Y., Lerner, E., Jin, Y. el al. (2017). The effect of macromolecular crowding on singleround transcription by E. coli RNA polymerase. *bio-archive journal.*

Cox, C.D., McCollum, J.M., Allen, M.S. et al. (2008). Using noise to probe and Characterize gene circuits. *PNAS* 105(31)**,** 10809-10814.

Daber, R., Stayrook, S., Rosenberg, A. et al. (2007). Structural analysis of lac repressor bound to allosteric effectors. *Elsevier; Journal of Molecular Biology* 370(4)**,** 609-619.

Davis, C.A., Capp, M.W., Record, M.T. et al. (2005). The effects of upstream DNA on open complex formation by Escherichia coli RNA polymerase. *PNAS* 102(2)**,** 285-290.

Deutscher, M.P. (2006). Degradation of RNA in bacteria: comparison of mRNA and stable RNA. *Nucleic Acids Research* 34(2)**,** 659–666.

Erban, R., Chapman, J., and Maini, P. (2007). A practical guide to stochastic simulations of reaction-diffusion processes. *Cornell University Library***,** 1-35.

Esmaeili, A., Davison, T., Wu, A., et al. (2015). PROKARYO: an illustrative and interactive computational model of the lactose operon in the bacterium Escherichia coli. *BMC Bioinformatics*. doi: 10.1186/s12859-015-0720-z.

Esyunina, D., Turtola, M., Pupov, D. et al. (2016). Lineage-specific variations in the trigger loop modulate RNA proofreading by bacterial RNA polymerases. *Nucleic Acids Research* 44(3).

Fuchs, G., Voichek, Y., Benjamin, S. et al. (2014). 4sUDRB-seq: measuring genomewide transcriptional elongation rates and initiation frequencies within cells. *BMC Genome Biology* 15(R69).

Gardner, T.S., Cantor, C.R., Collins, J.J. (2000). Construction of a genetic toggle switch in Escherichia coli. *Letters to Nature* 403(20)**,** 339-342.

Jacob, F., and Monod, J. (1961). Genetic Regulatoty Mechanism in the Synthesis of Proteins. *Elsevier, Journal of Molecular Biology* 3(3)**,** 318-356.

Jain, C. (2002). Degradation of mRNA in Escherichia coli. *IUBMB Life* 54**,** 315–321.

Kalisky, T., Dekel, E., and Alon, U. (2007). Cost-benefit theory and optimal design of gene regulation functions. *Physical Biology* 4**,** 229-245.

Khalil, A.S., and Ollins, J.J. (2010). Synthetic biology: applications come of age. *Nature Reviews Genetics* 11.

Kobayashi, H., Kaern, M., Araki, M. et al. (2004). Programmable cells: Interfacing natural and engineered gene networks. *PNAS* 101(22)**,** 8414-8419.

Kondo, T., Strayer, C. A., Kulkarni, R. D. et al. (1993). Circadian rhythms in prokaryotes: luciferase as a reporter of circadian gene expression in cyanobacteria. *Proc Natl Acad Sci* 90(12)**,** 5672-5676.

Kushner, S.R. (2002). mRNA Decay in Escherichia coli Comes of Age. *Journal of Bacteriology* 187(17)**,** 4658–4665.

Lee, D.H., Bae, J.E., Lee, J.H. et al. (2010). Quantitative Detection of Residual E. coli Host Cell DNA by Real-Time PCR. *Journal of Microbiology Biotechnology* 20(10)**,** 1463-1470.

Lee, D.J., Minchin, S.D., and Busby, S.J. (2012). Activating transcription in bacteria. *Annual Review of Microbiology* 66**,** 125-152.

Lei, S. Q., and Arkin, A. P. (2014). A versatile framework for microbial engineering using synthetic noncoding RNAs. *Nature Reviews: Microbiology* 12.

Lu, H.P. Xun, L., and Xie, X.S. (1998). Single-Molecule Enzymatic Dynamics. *Science* 282(5395).

Ma, H-W., Buer, J., and Zeng, A-P. (2004). Hierarchical structure and modules in the Escherichia coli transcriptional regulatory network revealed by a new top-down approach. *BMC Bioinformatics* 5(199).

McClure, W.R. (1980). Rate-limiting steps in RNA chain initiation. *PNAS* 77(10)**,** 5634-5638.

McDowell, J.C., Roberts, J.W. Jin, D.J. et al. (1994). Determination of Intrinsic Transcription Termination Efficiency by RNA Polymerase Elongation rate. *Science* 266**,** 822-825.

McKeague, M., Wong, R.S., and Smolke, C.D. (2016). Opportunities in the design and application of RNA for gene expression control. *Nucleic Acids Research* 44(7)**,** 2987–2999.

Mejia, Y.X., Nudler, E., and Bustamantea, C. (2014). Trigger loop folding determines transcription rate of Escherichia coli's RNA polymerase. *PNAS* 112(3)**,** 743-748.

Mekler, V., Kortkhonjia, E., Mukhopadhyay, J. et al. (2002). Structural Organization of Bacterial RNA Polymerase Holoenzyme and the RNA Polymerase-Promoter Open Complex. *Cell* 108(5)**,** 599–614. doi: Org/10.1016/S0092-8674(02)00667-0.

Mettetal, J., Muzzey, D., Pedraza, J.M. et al. (2005). Predicting stochastic gene expression dynamics in single cells. *PNAS* 103(19)**,** 7304–7309. doi: 10.1073/pnas.0509874103.

Munro, P.D., Ackers, G.K., and Shearwin, K.E. (2016). Aspects of protein–DNA interactions: a review of quantitative thermodynamic theory for modelling synthetic circuits utilising LacI and CI repressors, IPTG and the reporter gene lacZ. *Biophysical Reviews* 8(4)**,** 331–345.

Nath, K. and Koch, A.L. (1970). Protein Degradation in Escherichia coli. *Journal of Biological chemistry* 245(Issue of June 10)**,** 2889-2900.

Oehler, S. Alberti, S., and Muller-Hill, B. (2006). Induction of the lac promoter in the absence of DNA loops and the stoichiometry of induction. *Nucleic Acids Research* 34(2)**,** 606-612.

Oehler, S., Eismann, E.R., Kramer, H. et al. (1990). The three operators of the lac operon cooperate in repression. *The EMBO Journal* 9(4)**,** 973-979.

Proshkin, S., Rahmouni, A.R., Mironov, A. et al. (2010). Cooperation Between Translating Ribosomes and RNA Polymerase in Transcription Elongation. *Science* 328.

Ptashne, M., and Gann, A. (1997). Transcriptional activation by recruitment. *Nature* 386, 569-577.

Qayyum, M.Z., Dey, D., and Sen, R. (2016). Transcription elongation factor NusA is a general antagonist of Rho-dependent termination in Escherichia coli. *Journal of Biological chemistry* 291.

Ray, J.C., Wckersheim, M.L., Jalihal, A.P. et al. (2016). Cellular Growth Arrest and Persistence from Enzyme Saturation. *Plos computational Biology*, 1-21.

Roberts, E., Magis A., Ortiz J.O. et al. (2011). Noise Contributions in an Inducible Genetic Switch: A Whole-Cell Simulation Study. *Plos computational Biology* 7(3), 1-21.

Roberts, E., Stone, J.E., and Luthey-Schulten, Z. (2013). Lattice Microbes: high-performance stochastic simulation method for the reaction-diffusion master equation. *Journal of Computational Chemestry* 34, 245-255. doi: 10.1002/jcc.23130.

Reznikoff, W.S., Winter, R.B., and Hurley, C.K. (1974). The Location of the Repressor Binding Sites in the lac Operon. *Proc. Nat. Acad.* 71(6), 2314-2318.

Russell, W., Hertz Starr and Mc Millan (2008). *Biology: the dynamic science, First Edition.* USA: Yolanda Cossio.

Schlax, P.J., Capp, M.W., and Record, M.T. (1995). Inhibition of Transcription Initiation by lac Repressor *Journal of Molecular Biology* 245, 331-350.

Sharabiani, M.T., Siermal, M., Lehtinen, T. et al. (2005). Dynamic covariation between gene expression and proteome characteristics. *BMC Bioinformatics*, 1-18.

Singh, J., and Padgett, R.A. (2009). Rates of in situ transcription and splicing in large human genes. *Nature Structural and Molecular Biology* 16(11), 1128-1133.

Skinner, S.O., Sepulveda, L.A., Xu, H. et al. (2013). Measuring mRNA copy-number in individual Escherichia coli cells using single-molecule fluorescent in situ hybridization (smFISH). *National Institutes of Health* 8(6), 1100-1113.

Smith, K.C. and Hanawalt, P.C. (1969). *Molecular Photobiology: Inactivation and Recovery.* United Kingdom: Academic Press, Inc. (London) LTD.

So, L., Ghosh, A., Zong, C. et al. (2011). General properties of transcriptional time series in Escherichia coli. *Nature Genetics* 43(6), 554-560.

Solomon, E.P., Berg,, L.R., Martin, D.W. (2005). *Biology, seventh edition.* USA: Thomson Brooks/Cole.

stamatakis, M., and Mantzaris, N.V. (2009). Comparison of Deterministic and Stochastic Models of Lac Operon Genetic Network. *Biophysical Journal* 96, 887-906.

Straney, D.C., and Crothers, D.M. (1985). Intermediates in Transcription Initiation from the E. coli lac UV5 Promoter. *Cell* 43, 449-459.

Swint-Kruse, L. (2004). Using Networks To Identify Fine Structural Differences between Functionally Distinct Protein States. *Journal of Biochemistry* 43(34)**,** 10886-10895

Tang, G-Q., Roy, R., Bandwar, R.P. et al. (2009). Real-time observation of the transition from transcription initiation to elongation of the RNA polymerase. *PNAS* 106(52)**,** 22175–22180.

Tolic-Norrelykke, S.F., Engh, A.M., Landick, R. et al. (2004). Diversity in the Rates of Transcript Elongation by Single RNA Polymerase Molecules. *The Journal of Biological Chemistry* 279(5)**,** 3292-3299.

Tropp, B.E. (2010). *Molecular Biology: Genes to Proteins Fourth Edition.* Jones and Bartlett Learning.

Tropp, B.E. (2011). *Molecular Biology: Genes to Proteins, third edition.* Canada: Jones and Bartlett.

Vanzi, F., Broggio, C., Sacconi, L. et al. (2006). Lac repressor hinge flexibility andDNAlooping: single molecule kinetics by tethered particle motion. *Nucleic Acids Research* 34(12)**,** 3409-3420.

Wang, R., Jin, G., Zhang, X. et al. (2009). Modeling post-transcriptional regulation activity of small non-coding RNAs in Escherichia coli. *BMC Bioinformatics***,** 1-13. doi: 10.1186/1471-2105-10-S4-S6.

Washburn, R.S., and Gottesman, M.E. (2015). Regulation of Transcription Elongation and Termination. *Biomolecules* 2015(5)**,** 1063-1078.

Xu, J., Liu, S., Chen, M. et al. (2011). Altering residues N125 and D149 impacts sugar effector binding and allosteric parameters in Escherichia coli lactose repressor. *Biochemistry Journal* 50 (42)**,** 9002-9013. doi: 10.1021/bi200896t.

Xue, Q., and Yeung, E.S. (1995). Differences in the chemical reactivity of individual molecules of an enzyme. *Nature* 13.

Yagil, G., and Yagil, E. (1971). On the Relation between Effector Concentration and the Rate of Induced Enzyme Synthesis. *elsevier biophysical journal* 11(1)**,** 11-27. doi: 10.1016/S0006-3495(71)86192-1.

Yang, X., Chen, B., Cai, Y. et al. (2015). Understanding the lac operon with GeneAct. *Journal of Computational Biology and Drug Design* 8(2)**,** 168-188.

Zhuang, X., Bartley, L.E., Babcock, H.P. et al. (2000). A Single-Molecule Study of RNA Catalysis and Folding. *Science* 288**,** 2048-2051

Zuo, Y., and Steitz, T.A. (2016). A structure-based kinetic model of transcription. *Point of view* 8(1)**,** 1-8.

Table 1: The lac operon regulation. The regulatory reactions in inactive and active states of the lac circuit and the corresponding rate values (Roberts et al. (2011)). These rate constants are assumed as the default values (supplementary file).

Reactions of lac operon regulation	Stochastic rates
Gene inactive state	K_{inac} $(M^{-1}s^{-1})$
$R_2 + O \longrightarrow R_2O$	$k_{ron} = 2.43e^{+06}$
$IR_2 + O \longrightarrow IR_2O$	$k_{iron} = 1.21e^{+06}$
$I_2R_2 + O \longrightarrow I_2R_2O$	$k_{i2ron} = 2.43e^{+04}$
Gene active state	K_{acti} (s^{-1})
$R_2O \longrightarrow R_2 + O$	$k_{roff} = 6.30e^{-04}$
$IR_2O \longrightarrow IR_2 + O$	$k_{iroff} = 6.30e^{-04}$
$I_2R_2O \longrightarrow I_2R_2 + O$	$k_{i2roff} = 3.15e^{-01}$

Table 2: The measurements of the maximum transcription rates and of the transcription repression for different inducer-repressor complexes. The measured maximum transcription rates in three lac circuit kinetic models in inactive state. The transcription repression factors are computing using the expression (3). Models 1, 2 and 3 contain only the repressor complexes R_2, IR_2, and I_2R_2, respectively.

	Model 1	**Model 2**	**Model 3**
$k_{tr\max}(\mathrm{s}^{-1})$	$k^1_{tr\max} = 0.013 \pm 0.008$	$k^2_{tr\max} = 0.033 \pm 0.012$	$k^3_{tr\max}$? $k^2_{tr\max}$
TR	0.60	0.23	$TR_3 = TR_2$

Table 3: The mRNA steady-state abundances for the selected values of transcription rate. The mRNA steady-state counts in three models of the lac circuits in active state. Models 1, 2 and 3 contain only the repressor complexes R_2, IR_2, and I_2R_2, respectively.

Transcription rate (s^{-1})	Model 1	Model 2	Model 3
0.21e-1	2	2	2
0.31e-1	4	3	4
0.63e-1	6	6	5
1.26e-1	12	13	12
2.52e-1	25	24	24
5.04e-1	49	48	50

Supplementary Material

Determining Transcription Rates Yielding Steady State Production of mRNA in the lac Genetic Circuit of E. coli

Komlan Atitey, Pavel Loskot*, Paul Rees

*** Correspondence:** Pavel Loskot: p.loskot@swan.ac.uk

1 Lattice Microbe software

The software is freely downloadable from [S1]. We compiled the version 2.3 from the source code to ensure its full compatibility with the operating system and the hardware and software installed. The simulations in Lattice Microbe can be configured and run from a terminal command line, or the users can use the Python interface which is provided with the software. The User Guide and the tutorial Manual can be obtained from [S1]. One of the neat features of the software is its automatic discovery and usage of the available computing resources on the multi-core or multi-processor systems with one or more GPU (Graphical Processing Unit) cards.

The main objective of the Lattice Microbe software is to efficiently numerically solve the chemical master equation or the reaction-diffusion master equation. A network of chemical reactions can be imported from a SBML (Systems Biology Markup Language) file, or the reactions can be inserted individually, for example, using a Python script. Starting from the initial species concentrations, and for the specified rates of chemical reactions, the software generates stochastic trajectories of the species counts. In order to limit the volume of output data produced from the simulations, the

trajectories are further sub-sampled before they are added to the data file. The input parameters and output data are stored in a single HDF5 (hierarchical data filesystem, version 5) file. Matlab as well as Python can read and write the HDF5 files; this can be used to both configure the simulations as well as to process the simulation outputs. Lattice Microbe offers several numerical solvers to choose from. The default solver is based on the Gillespie's algorithm, so it solves the chemical master equation exactly. Furthermore, Lattice Microbe strongly supports parallelization of simulations. For instance, in case of well-stirred simulations which are completely described by a network of chemical reactions, the users can specify how many independent trajectories of the species counts should be generated at the same time. This feature facilitates the statistics of the simulation outputs, for example, to obtain a time-evolution of the species count distribution.

On the other hand, among the features we found missing in Lattice Microbe are: (1) support for time-varying reaction rates, and (2) recoding the species counts into the output file based on certain events rather than at regular time intervals. For instance, the simulation could be stopped automatically when the steady-state count has been reached.

2 Kinetic model of the lac genetic switch

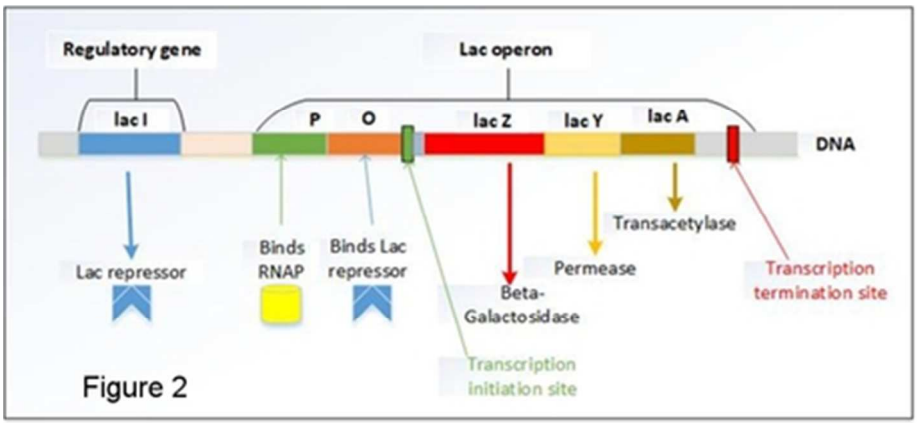
Table S1 lists all reactions in a biochemical network of the lac genetic switch. The chemical kinetic model is adopted from [S2, Table 1] ([16] in the main text). The rates in Table S1 are considered to be the default values in our numerical experiments. In the three models considered in our work, we only modified the reactions pertaining to the inducer-repressor interactions. Specifically, Models 1, 2 and 3 only contain the species R_2, IR_2 and I_2R_2, respectively, and the rates of other reactions are set as described in the main text.

Table S1 Kinetic model of the lac circuit

Reaction	Forward rates			Reverse rates		
Lac operon regulation						
$R_2+O \leftrightarrow R_2O$	k_{ron}	2.43e+06	$M^{-1}s^{-1}$	k_{roff}	6.30e-04	s^{-1}
$IR_2+O \leftrightarrow IR_2O$	k_{iron}	1.21e+06	$M^{-1}s^{-1}$	k_{iroff}	6.30e-04	s^{-1}
$I_2R_2+O \leftrightarrow I_2R_2O$	k_{i2ron}	2.43e+04	$M^{-1}s^{-1}$	k_{i2roff}	3.15e-01	s^{-1}
Transcription, translation, and degradation						
$O \rightarrow O+mY$	k_{tr}	1.26e-01	s^{-1}			
$mY \rightarrow mY+Y$	k_{tn}	4.44e-02	s^{-1}			
$mY \rightarrow \emptyset$	k_{degm}	1.11e-02	s^{-1}			
$Y \rightarrow \emptyset$	k_{degp}	2.10e-04	s^{-1}			
Inducer-repressor interactions						
$I+R_2 \leftrightarrow IR_2$	k_{ion}	2.27e+04	$M^{-1}s^{-1}$	k_{ioff}	2.00e-01	s^{-1}
$I+IR_2 \leftrightarrow I_2R_2$	k_{i2on}	1.14+04	$M^{-1}s^{-1}$	k_{i2off}	4.00e-01	s^{-1}
$I+R_2O \leftrightarrow IR_2O$	k_{iopon}	6.67e+02	$M^{-1}s^{-1}$	k_{iopoff}	1.00e+00	s^{-1}
$I+IR_2O \leftrightarrow I_2R_2O$	k_{i2opon}	3.33e+02	$M^{-1}s^{-1}$	$k_{i2opoff}$	2.00e+00	s^{-1}
Inducer transport						
$I_{ex} \leftrightarrow I$	k_{id}	2.33e-03	s^{-1}	k_{id}	2.33e-03	s^{-1}
$Y+I_{ex} \leftrightarrow YI$	k_{yion}	3.03e+04	$M^{-1}s^{-1}$	k_{yioff}	1.20e-01	s^{-1}
$YI \rightarrow Y+I$	k_{it}	1.20e+01	s^{-1}			

[S1] http://www.scs.illinois.edu/schulten/lm/

[S2] Roberts E, Magis A, Julio O, Baumeister W and Luthey-Schulten Z. Noise Contributions in an Inducible Genetic Switch: A Whole-Cell Simulation Study. Plos computational Biology 2011, 7(3):1-21.



38x17mm (300 x 300 DPI)

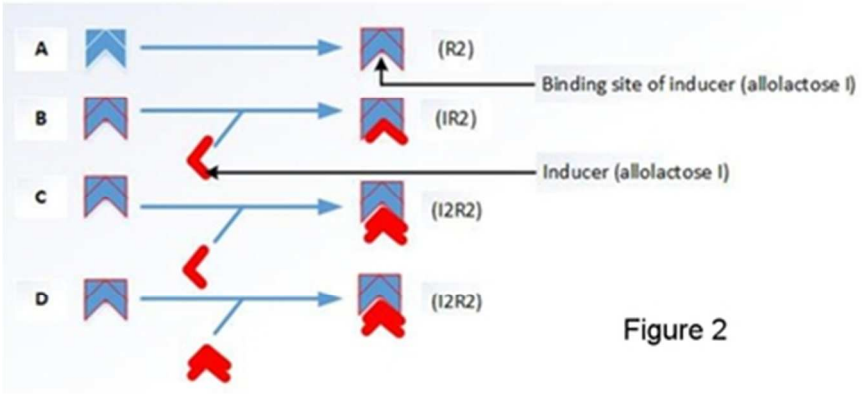
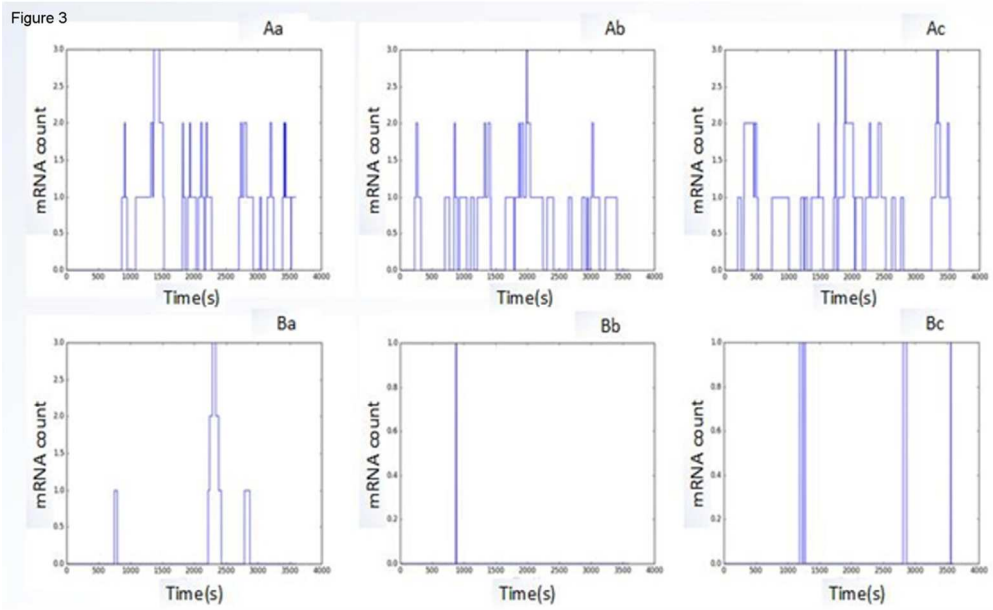
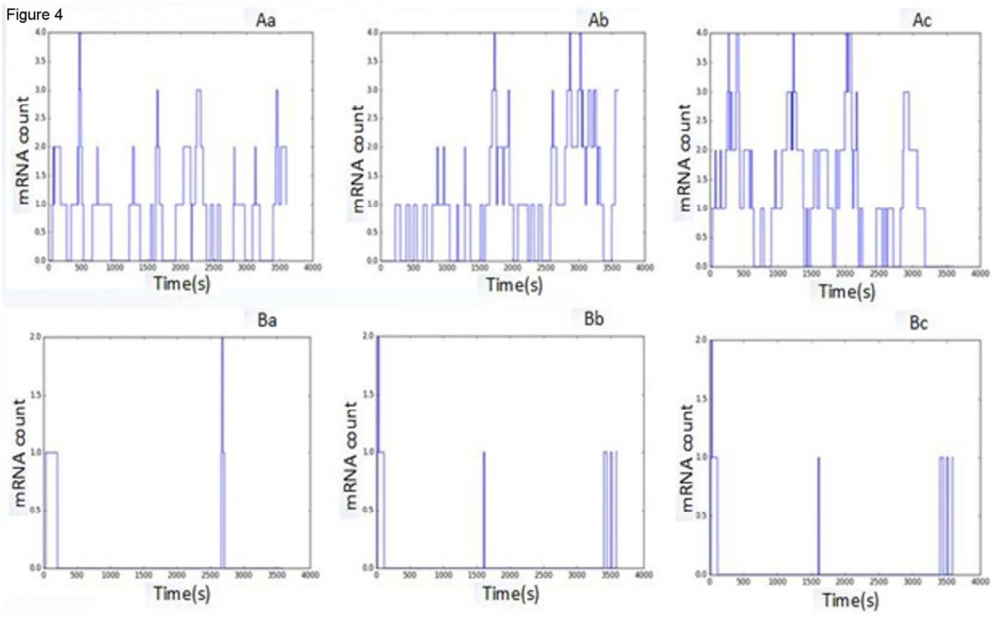


Figure 2

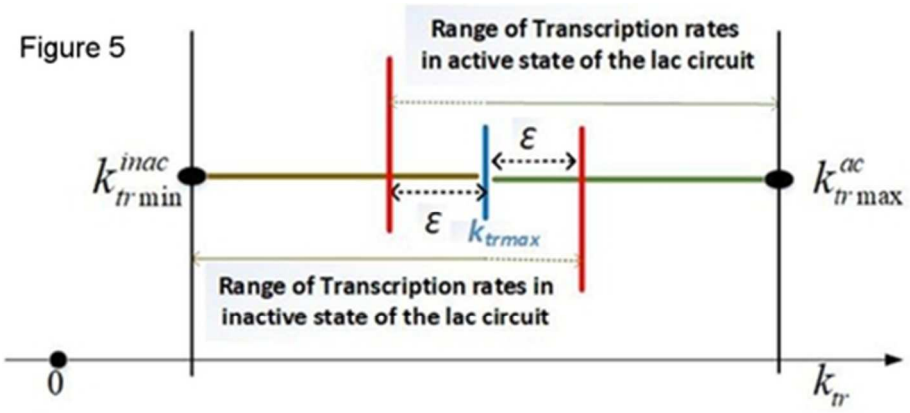
36x16mm (300 x 300 DPI)



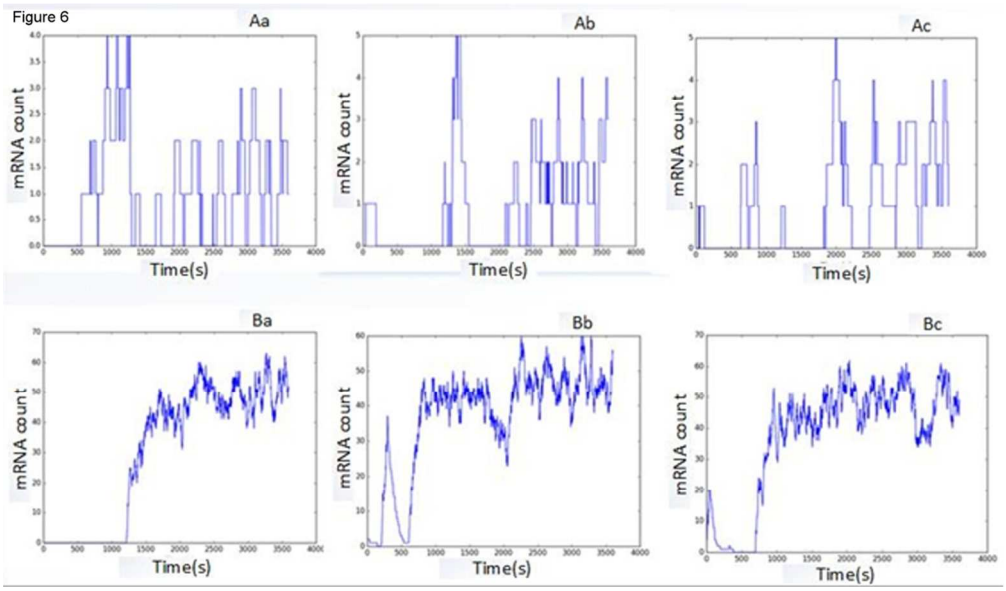
107x65mm (300 x 300 DPI)



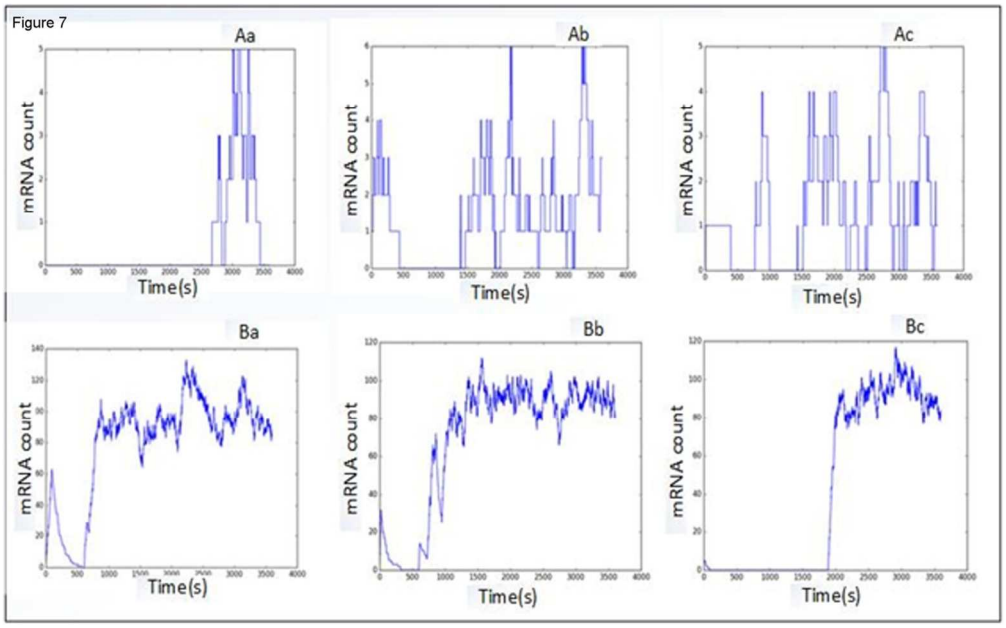
105x64mm (300 x 300 DPI)



38x17mm (300 x 300 DPI)



102x59mm (300 x 300 DPI)



111x69mm (300 x 300 DPI)

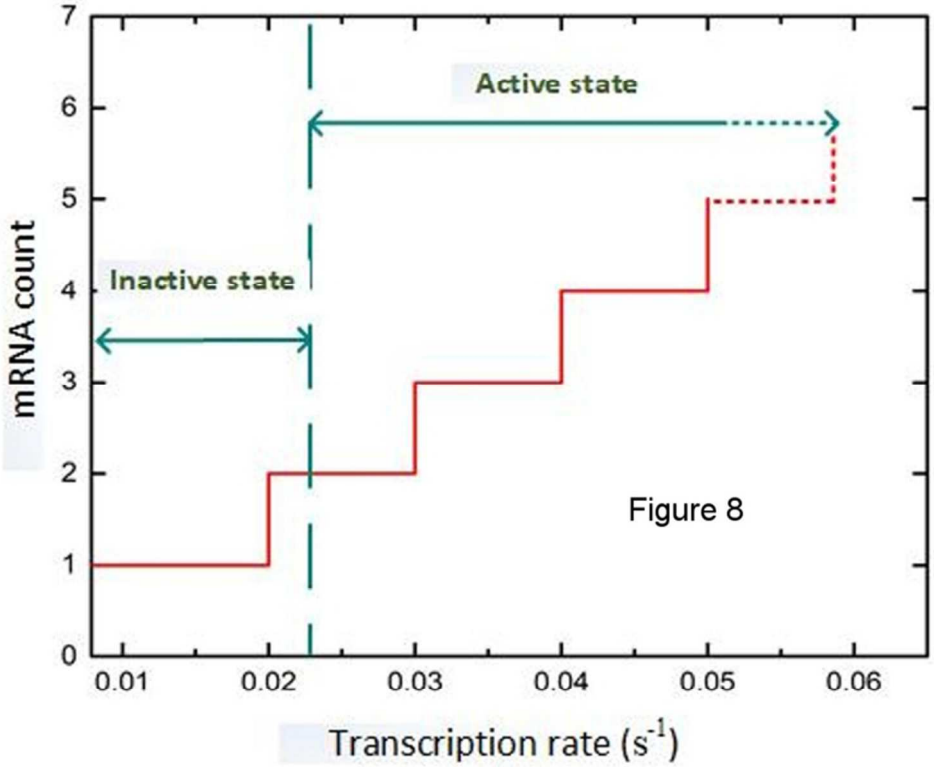


Figure 8

69x55mm (300 x 300 DPI)