

Delayed Stochastic Model of Transcription at the Single Nucleotide Level

ANDRE S. RIBEIRO, *OLLI-PEKKA SMOLANDER, *TIINA RAJALA,
*ANTTI HÄKKINEN, and OLLI YLI-HARJA

ABSTRACT

We present a delayed stochastic model of transcription at the single nucleotide level. The model accounts for the promoter open complex formation and includes alternative pathways to elongation, namely pausing, arrest, misincorporation and editing, pyrophosphorolysis, and premature termination. We confront the dynamics of this detailed model with a single-step multi-delayed stochastic model and with measurements of expression of a repressed gene at the single molecule level. At low expression rates both models match the experiments but, at higher rates the two models differ significantly, with consequences to cell-to-cell phenotypic variability. The alternative pathway reactions, due to, for example, causing polymerases to collide more often on the template, are the cause for the difference in dynamical behaviors. Next, we confront the model with measurements of the transcriptional dynamics at the single RNA level of an induced gene and show that RNA production, besides its bursting dynamics, also exhibits pulses (2 or more RNAs produced in intervals smaller than the smallest interval between initiations). The distribution of occurrences and amplitudes of pulses match the experimental measurements. This pulsing and the noise at the elongation stage are shown to play a role in the dynamics of a genetic switch.

Key words: elongation regulation, stochastic, traffic, transcription dynamics.

1. INTRODUCTION

STOCHASTIC FLUCTUATIONS IN GENE EXPRESSION have a significant role at the single-cell level (McAdams and Arkin, 1997; Arkin et al., 1998; Elowitz et al., 2002; Ozbudak et al., 2002), for example, in phenotypic expression and differentiation pathway selection (Süel et al., 2006). The relevance of accurately accounting for the noise in gene expression, especially in transcription (Elowitz et al., 2002), is enhanced by the low number of transcription events and its regulators (transcription factors and promoters) (Beckskei et al., 2005; Bernstein et al., 2002; Bon et al., 2006). Thus, the favored approach to simulate gene expression is the stochastic simulation algorithm (Gillespie, 1977; McAdams and Arkin, 1998).

While the first stochastic models assumed transcription to be instantaneous (McAdams and Arkin, 1997), it takes a considerable time for an RNA polymerase (RNAP) to create an RNA strand. This time interval

Computational Systems Biology Research Group, Department of Signal Processing, Tampere University of Technology, Tampere, Finland.

*These three authors contributed equally to the work.

depends on the gene length and, thus, varies from gene to gene. Therefore, recent models introduced time delays in the appearance of the products of gene expression (Ribeiro et al., 2006). Although models using only non-delayed reactions can mimic gene expression fluctuations (Raser and O’Shea, 2004), these studies focused on steady state dynamics, where usually delayed and non-delayed models match after a transient (Zhu et al., 2007). Models of complex gene networks (e.g., with feedback loops) require delayed reactions to accurately reproduce the dynamics (Gaffney and Monk, 2006; Bratsun et al., 2005).

Also, most models of gene expression and gene networks assume that transcription regulation occurs only during initiation, via its rate constant and transcription factors (Kierzek et al., 2001). However, during elongation, the RNAP is in constant kinetic competition with other regulatory pathways (Greive et al., 2005), and regulatory mechanisms act at this stage in both prokaryotes and eukaryotes (Davenport et al., 2000; Core and Lis, 2008).

Recently, the real-time expression of a repressed lac promoter was directly monitored in individual *E. coli* cells in an epifluorescence microscope with a single-protein resolution (Yu et al., 2006). It was found that proteins are produced in bursts, with the distribution of bursts fitting a Poisson distribution, while the number of proteins per burst follows a geometric distribution (Yu et al., 2006). The delayed stochastic model of gene expression, that models transcription and translation as multiple time delayed reactions, reproduces the observed kinetics (Zhu et al., 2007).

This gene was kept strongly repressed, thus its expression was sparse in time. When less repressed, its expression increases and several alternative reaction pathways, not previously considered, play a relevant role. Delayed stochastic simulations and detailed models of transcription, where the template strand is explicitly modeled, were shown to match except when two-body effects (collisions between polymerases or “traffic”) are important (Roussel and Zhu, 2006b). Thus, the importance of modeling, at the elongation stage, events such as pauses or arrests since these affect traffic significantly and their effects have not yet been studied.

A detailed model of transcription, similar to the one proposed in Roussel and Zhu (2006a), but including backtracking, was recently analyzed (Voliotis et al., 2008). This study focused on the distribution of elongation times and showed the relevant role of backtracking at the elongation stage. This model did not include, for example, RNA polymerase halting (Greive et al., 2005), promoter complex formation (McClure, 1980) or premature termination.

The alternative pathways that can occur at the elongation stage play a role in transcription regulation (Greive et al., 2005; Landick, 2006). For example, their occurrence can amplify traffic events between preceding RNAP molecules. Finally, the time duration of promoter binding and open complex formation, that can last from a few seconds to several minutes before elongation begins (McClure, 1980) and thereby, affect the dynamics of even the simplest gene networks (Ribeiro, 2007), have not yet been accounted for in models of transcription at the single nucleotide level (Roussel and Zhu, 2006b; Kosuri et al., 2007; Voliotis et al., 2008).

We propose a detailed model of transcription, at the single nucleotide level, based on the model proposed in Roussel and Zhu (2006), but that incorporates promoter occupancy time, pausing, arrest, misincorporation and editing, pyrophosphorolysis, premature termination, and accurately accounts for the region occupied by an RNAP when on the DNA template. A variable time delayed reaction models the formation of the promoter-RNAP complex (McClure, 1980), to account for the time that the RNAP is not moving while occupying the promoter, preventing further transcription initiations. Since most experimental measurements of transcriptional dynamics are from *Escherichia coli*, we model transcription in *E. coli*.

First, we show that when transcription is sparse in time, the detailed model here proposed matches the delayed model proposed in Zhu et al. (2007), here referred to as the “reduced model,” and that both models accurately match measurements of gene expression at the single molecule level (Yu et al., 2006). Next, we compare the dynamics of the two models when the promoter is not repressed. Finally, we show that the detailed model matches the experimental measures of transcription at the single RNA level (Golding et al., 2005), while the reduced model is unable to capture some of the dynamics observed, and that noise of events at the elongation stage affect the dynamics of a simple gene network, the toggle switch, and therefore, are relevant.

2. MODEL OF TRANSCRIPTION

The dynamics of simulations (Ribeiro and Lloyd-Price, 2007) is driven by the delayed Stochastic Simulation Algorithm (Roussel and Zhu, 2006b), based on the original SSA (Gillespie, 1977). Delayed events

in reactions are represented as, for example, $A \rightarrow B + C(\tau)$. When occurring at moment t , B is instantaneously produced at t and C placed on a waitlist until it's released, at $t + \tau$ seconds. τ can be drawn from a distribution each time the reaction occurs.

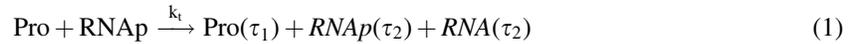
Gene expression has two main steps: transcription and translation. Transcription is the copying of a gene in the DNA strand by an RNA polymerase (RNAP) into an RNA molecule which is then translated into a protein. The RNAP unwinds and reads the DNA, making the RNA step by step, adding matching nucleotides while going through the DNA strand (Alberts et al., 2002).

Transcription has three main phases: initiation, elongation, and termination. Initiation consists of the RNAP attaching to a promoter and unwinding a portion of the DNA double helix to expose the template DNA strand. After this, the RNAP speeds up and elongation begins. The nucleotide sequence of the growing RNA chain is determined by the base sequence of the template DNA. Behind the region where ribonucleotides are added, the RNA chain is displaced and the DNA double helix is reformed. In termination, at the end of the gene, a single-stranded RNA molecule is released.

The RNAP, at nucleotide n_i , can go forward along the DNA template, but can also stop due to pausing or arrest (Davenport et al., 2000), or go backwards due to editing to correct a misincorporation, or by pyrophosphorolysis (Greive et al., 2005). In addition, the RNAP can prematurely fall off from the DNA strand (premature termination) (Uptain et al., 1997). The probability of occurrence of each of these pathways is often sequence-specific and can depend on the presence of specific transcription factors (Greive et al., 2005).

2.1. Reduced delayed model of transcription

The reduced model of transcription consists of a multi-delayed reaction of transcription (reaction 1) (Ribeiro et al., 2006). To model the experiment in Yu et al., (2006) one also needs reactions of repression (reaction 2) and un-repression (reaction 3) of the gene promoter region, Pro, by a repressor molecule, X, since this gene was kept under strong repression (Yu et al., 2006) (reactions 2 and 3 will be identical when using the detailed model of transcription). Rate constants (in s^{-1}) are set to: $k_t = 0.01$, $k_{rep} = 1$, and $k_{unrep} = 0.1$. Time delays (in s) are set to: $\tau_1 = 40$, $\tau_2 = 90$, and $\tau_3 = 2$. Initially, Pro = 1, RNAP = 40, RNA = 0, and X = 100 (Zhu et al., 2007).



2.2. Detailed model of transcription at the single nucleotide level

During elongation the RNAP occupies ~ 25 nucleotides (Greive et al., 2005). Thus, at nucleotide n , the RNAP occupies nucleotides $[n - \Delta, n + \Delta]$, where $\Delta = 12$ is the number of nucleotides occupied by an RNAP to the left and right of its center in the template. Hence, the active centers of two consecutive RNAP's cannot be less than 25 nucleotides apart.

The number of RNAP's free for transcribing at any moment is estimated to be 28 per cell (Bremer et al., 2003). This amount set here differs from Zhu et al., (2007). However, the initiation rate constants also differ so that the propensities of transcription initiation are equal in the two models. It's noted that the amount of RNAP's never limited transcription in any simulation.

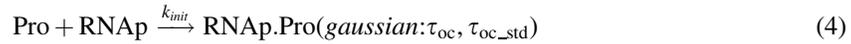
The following abbreviations are used: U_n is an unoccupied nucleotide at position n of the chain. O_n is a nucleotide at position n occupied by an RNAP. A_n is a nucleotide occupied and activated by an RNAP. Pro is the gene promoter region and $\alpha = 10$, is the number of nucleotides that form the region of elongation initiation. When at these first 10 nucleotides, the RNAP has a slower elongation rate (Hansen and McClure, 1980).

2.2.1. Promoter binding and open complex formation. When the RNAP detects the promoter in the DNA strand, it binds and forms the closed complex. Next, it unwinds the DNA double helix creating the open complex, revealing the template DNA strand and allowing the RNAP to begin forming the RNA strand from nucleotide n_1 .

The promoter binding (closed complex formation) and the open complex formation are rate-limiting steps of transcription (McClure, 1980) and, during this stage, that can last from a few seconds to several minutes

(McClure, 1980), the RNAP does not move to start transcribing. We model this by introducing a delay, τ_{oc} , on the promoter release (reaction 1), drawn from a Gaussian distribution with a mean of 40 s and standard deviation of 4 s, according to experimental measures of an active unrepresed lac promoter (Lutz et al., 2001).

While previous detailed models do not include this delay, it plays a relevant role, especially when $([RNAP] \cdot k_{init} < \tau_{oc}^{-1})$, for example, causing the distribution of intervals between initiation events to be Gaussian-like rather than exponential. Transcription initiation is modeled by reaction 4:



After the delay elapses and if the first 13 nucleotides are unoccupied, the RNAP can initiate elongation (reaction 5). When it does, the promoter becomes again available for reactions.

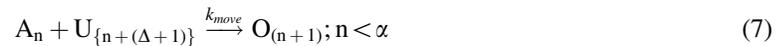


2.2.2. Elongation initiation. In the initial steps, the RNAP moves slower on the DNA template and the likelihood of pausing and premature termination are greater than in other elongation periods (Artsimovitch, 2000; Hansen and McClure, 1980; Kapanidis et al., 2006; Roberts, 2006). This is due to the different conformation of RNAP at this stage and, especially, due to the need of having a σ factor attached, required for promoter recognition (Mooney et al., 2005; Lewin, 2008). σ usually relaxes its hold from the RNAP after the first 10 nucleotides (Alberts et al., 2002; Hansen and McClure, 1980) and allows the RNAP to speed up to its normal elongation speed. Thus, we set the elongation rate at the first 10 nucleotides to five times smaller than for subsequent ones (the exact difference has not yet been experimentally accessed; reaction 8).

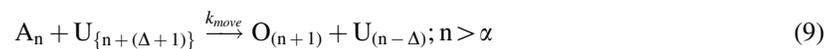
Since we assume a constant pool of nucleoside triphosphates (NTPs), reaction 6 and identical ones below are pseudo-first-order processes (Roussel and Zhu, 2006a).



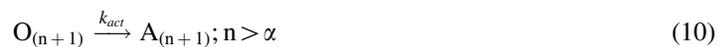
After transcribing the first nucleotide, the RNAP can move to the next one (reaction 7), activating and transcribing it (reaction 8). During this period, the RNAP is still occupying all initial nucleotides, so no preceding nucleotide is released.



2.2.3. Elongation. During elongation, nucleotides are added one at a time to the growing RNA molecule according to the DNA sequence. Elongation is divided into two steps. First, the RNAP moves from the activated nucleotide n to nucleotide $n+1$ (reaction 9).

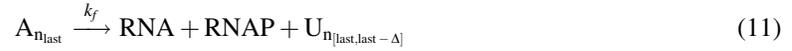


Once occupying nucleotide O_{n+1} , there are several competing pathways. The most likely to occur is activation (reaction 10), after which the RNAP can again move forwards. The activation step models the addition of the complementary nucleotide to the growing RNA molecule (Roussel and Zhu, 2006a).



The RNA chain elongation rate is usually inferred from the time interval between “beginning and end” states of transcription. Thus, the rate comprises events such as pausing and arrest, resulting in a lower “average” elongation rate than in reality (Darzacq et al., 2007). Thus, we set the elongation rate, $k_{act} = 75$ nt/s, to the upper range of values measured (Greive et al., 2005; Uptain et al., 1997). The elongation rate can vary with the growth rate of *E. coli* (Vogel and Jensen, 1994). The 75 nt/s is consistent with the duplication time of 55 minutes (Vogel and Jensen, 1994; Yu et al., 2006).

2.2.4. Termination. When the RNAP reaches the termination sequence, the transcription bubble collapses as the RNA-DNA hybrid disrupts and the RNAP and RNA are released. We modeled termination so that when the last nucleotide is activated and the mature RNA is released, the RNAP is released and the 12 last nucleotides are unoccupied (reaction 11). The rate for the transcript release, k_f , is 2 s^{-1} (Greive et al., 2008).



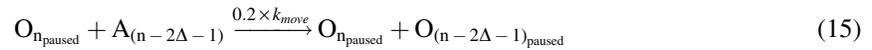
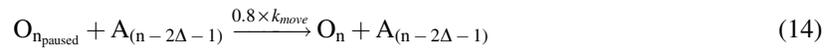
2.2.5. Pausing. Elongation in prokaryotes and eukaryotes is frequently interrupted by pauses (reaction 12) (Herbert et al., 2006; Landick, 2006; Shundrovsky et al., 2004), where the RNAP is halted at a nucleotide (Davenport et al., 2000). It is noted that pauses, especially due to their high frequency of occurrence, ought to be explicitly modeled, especially when more than one RNAP can be on the DNA strand.

Pausing is reversible and after a while the RNAP resumes its movement. This interval varies between pause events. Longer pauses, over 20 s, appear to occur at specific DNA template points. Most pauses last less than 5 s (Herbert et al., 2006). The rate of occurrence of pause is 0.55 s^{-1} (Greive et al., 2005; Neuman et al., 2003).

Notice that this reaction competes with the elongation reaction. The relative value between these two rates determines the fraction of times each occurs (Gillespie, 1977). Since $k_{\text{pause}} \sim k_{\text{act}}/136$, a pause event occurs, on average, every 136 activation events, which, in a template of 2445 nucleotides is significant, for example, for causing traffic when the promoter is expressing at a high rate.



The paused complex can be freed at any time, but on average is spontaneously released after 3 seconds (reaction 13, $d_{\text{pause}} = 3$) or sooner if there is a collision (reaction 14) with a following RNAP (Epshtein and Nudler, 2003; Greive et al., 2005). Sometimes, the collision causes the other RNAP to pause as well (reaction 15). This is set to have a probability of occurrence of 20% (reaction 15).



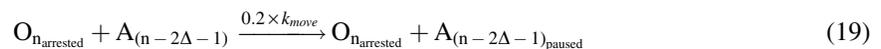
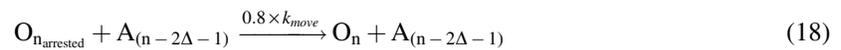
2.2.6. Arrest. An RNAP can stop irreversibly, if no external factor rescues it, via arrest (reaction 16). This reaction also competes with activation. Arrests are infrequent, with an estimated rate, k_{arrest} , of 0.00027 s^{-1} (Greive et al., 2005).



RNAP's can escape from this state with the help of transcription repair coupling factors Mfd or GreA/B proteins (Borukhov et al., 2005; Park et al., 2002). We set the rate of arrest escape to 0.01 s^{-1} ($d_{\text{arrest}} = 100$; reaction 17). Since, on average, an RNAP is arrested for 100 s, when these rare events occur, transcription is significantly affected.



The RNAP can escape arrest also due to the collision with a trailing RNAP (reaction 18) (Epshtein and Nudler, 2003). Sometimes this collision leaves the first RNAP still arrested and causes the colliding RNAP to pause (reaction 19). We set to 80% the fraction of collisions that free the arrested RNAP.



2.2.7. Misincorporation and editing. Misincorporation stands for the RNAP polymerizing an incorrect nucleotide at the growing transcript. Mismatches between the RNA and DNA hybrid weaken the stability of the elongation complex and slow down or halt elongation. This decrease in the binding affinity is presumed to give the RNAP proofreading capability (Greive et al., 2005). RNAP can cleave, with the help of elongation accessory proteins such as GreA, a short oligonucleotide sequence containing the misincorporated residue from the growing RNA chain, and correct the error (Erie et al., 1993). This process is called editing (modeled by reactions 20 and 21). Since during editing the RNAP can back-slide several nucleotides, we set in the reaction channel via which RNAs with errors are corrected, d_{correct} to 5 s (reaction 21), which corresponds to the average time necessary for such correction to be completed.

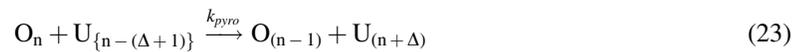
The rate at which the RNAP misincorporates nucleotides varies. Usually, it occurs at 0.05 to 1 min^{-1} (Erie et al., 1993; Greive et al., 2005); thus, $k_{\text{error correct}} = 0.00875 \text{ s}^{-1}$.



Not all errors due to misincorporations are corrected, and some RNAs can lead to proteins that are non-functional. Via reactions 20 and 21, errors are corrected, and the time to do so is accounted for. The cases where errors are not corrected are accounted for by reaction 22, which competes with the normal termination reaction (11). We estimated k_{mis} to be 0.05 s^{-1} , given the misincorporation rate per nucleotide and the number of nucleotides in the gene constructed in Yu et al., (2006).



2.2.8. Pyrophosphorolysis. Pyrophosphorolysis is the inverse of elongation, i.e., it's the sequential removal of nucleotides from the growing end of the RNA (reaction 23). It is estimated that the elongation is favored over pyrophosphorolysis by a factor of ~ 100 (Greive et al., 2005); thus, we set the rate constant of pyrophosphorolysis, k_{pyro} , to 0.75 s^{-1} .



2.2.9. Premature termination. The RNAP molecule can prematurely fall off the template DNA, failing to synthesize the full-length RNA (reaction 24). Although rare, this premature termination can affect regulation of gene expression (Grundy and Henkin, 2006). In general, the elongation complex is stable. The half-time for dissociation of the RNAP is 60 min (Lewin, 2008). Thus, we set the rate constant of premature termination to $k_{\text{prem}} = 0.00019 \text{ s}^{-1}$.



3. RESULTS AND DISCUSSION

We first validate our detailed model by confronting it with the reduced model (Zhu et al., 2007) and with experimental measurements of gene expression at the single protein level (Yu et al., 2006). This experiment measured the expression of the gene construct *tsr-venus*, which has 2445 nucleotides. To allow an exact count of the number of proteins and RNAs, the gene was strongly repressed. Thus, besides the reactions modeling gene transcription, we include reactions 2 and 3 that model the repression and un-repression of the promoter with a repressor molecule (Zhu et al., 2007).

In each simulation we measure the dynamics of the expression of a single *tsr-venus* gene. Thus, we refer to each simulation as a ‘‘cell.’’ Each ‘‘cell’’ has a lifetime of 55 minutes (Yu et al., 2006) and is initialized with 100 repressor molecules X (Zhu et al., 2007).

Figure 1 shows the distribution, from 1000 simulations, of the fraction of cells with a given number of completed transcription events (i.e., number of RNAs produced). The two models match with the

TABLE 1. VALUES OF THE RATE CONSTANTS USED IN THE DETAILED MODEL

Reactions	Parameter	Rate constant	Reference
Repression	k_{rep}	$1 s^{-1}$	Zhu et al., 2007
Unrepression	k_{unrep}	$0.1 s^{-1}$	Zhu et al., 2007
Initiation	k_{init}	$0.0145 s^{-1}$	www.openwetware.org
Elongation initiation	$k_{act\ first}$	$15 s^{-1}$	Estimated
Elongation	k_{move}, k_{act}	$75 s^{-1}$	Vogel et al., 1994
Termination	k_f	$2 s^{-1}$	Greive et al., 2008
Pausing	k_{pause}	$0.55 s^{-1}$	Greive et al., 2005
Release pause	$k_{pause\ release}$	$1/3 s^{-1}$	Greive et al., 2005
Arrest	k_{arrest}	$0.000278 s^{-1}$	Greive et al., 2005
Release arrest	$k_{arrest\ release}$	$1/100 s^{-1}$	Estimated
Editing	$k_{error\ correct}$	$0.00875 s^{-1}$	Greive et al., 2005
Release editing	$k_{edit\ release}$	$1/5 s^{-1}$	Estimated
Misincorporation	k_{mis}	$0.05 s^{-1}$	Estimated
Pyrophosphorolysis	k_{pyro}	$0.75 s^{-1}$	Greive et al., 2005
Premature termination	k_{prem}	$0.00019 s^{-1}$	Lewin et al., 2008
Open complex form.	τ_{oc}	$\mu = 40 s, \sigma = 4s$	Zhu et al., 2007

The estimated values fit experimental data from Yu et al., (2006) of single gene expression at the single protein level (Zhu et al., 2007) and for which the experimental values are not yet published.

measurements from (Yu et al., 2006). We use the two-sample Kolmogorov-Smirnov test (KS test) to assess the goodness of fit of simulation to experimental results. The p -values of reduced and detailed model distributions confronted with the experimental data are 0.87 and 0.95, respectively. Thus, under strong repression, both models match the experimental data very accurately and thus, are equivalent.

In both models, since $k_{rep} = 1$ and $k_{unrep} = 0.1$, and since X is 100, repression is 1000 times more propense than unrepression, and thus, only $\sim 0.1\%$ of the time should the promoter be available to transcribe. Since the cell lifetime is 3300 s, the expected interval that the gene is available is 3.3 s. Because $[RNAP].k_{init} = 0.4$, the expected number of transcription events during this time is ~ 1.3 . This quantity is slightly lower since the promoter release is delayed, which decreases the interval that the promoter is free to transcribe. In agreement, the three distributions have an average number of transcriptions per cell of ~ 1.2 (Fig. 1).

Since transcription events are rare, reduced and detailed model have identical dynamics. This is not expected to hold if the number of transcription events increases significantly, due to the emergence of traffic events and other alternative reaction pathways such as pause.

Thus, we now compare the transcriptional dynamics of the two models in the absence of repressor molecules X . In Figure 2, we plot the normalized distribution of the number of cells, out of 1000 cells, which produced a given amount of RNAs during their lifetime.

The transcriptional dynamics of the two models differ significantly. The distribution of the number of RNA molecules produced in the reduced model has a mean of 75.7 per cell whereas the detailed model has

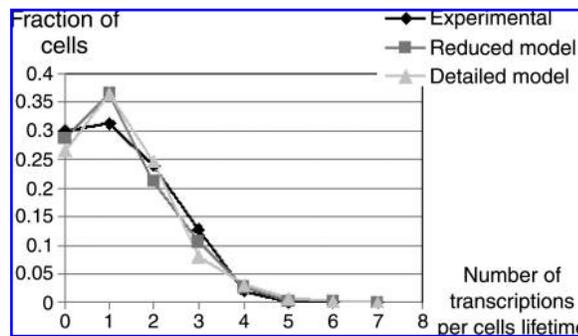


FIG. 1. Distributions of the fraction of cells with a given number of completed transcription events (i.e., number of RNAs produced) of the two models and experiments from Yu et al. (2006). The p -values of the KS test of reduced and detailed models distributions confronted with the experimental data are 0.87 and 0.95, respectively. For the two models, the distributions are from 1000 independent simulations each lasting 3300 s where genes are subject to repression.

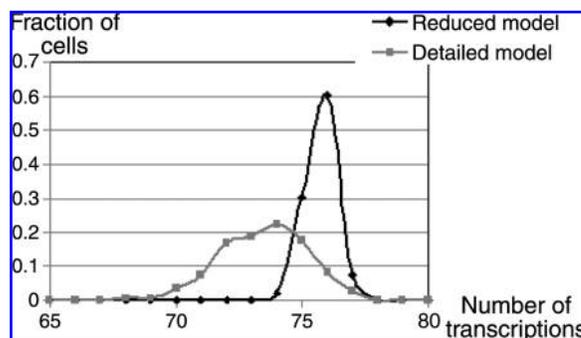


FIG. 2. Normalized distribution of the fraction of cells out of 1000 cells that transcribed a given number of RNA molecules during their lifetime of 3300s, for the reduced and detailed models, in the absence of repressor molecules. The cell-to-cell phenotypic variability of cells with a detailed model of transcription is much higher due to the stochasticity at the elongation stage.

73.4 (this quantity does not account for the 1.9 erroneous RNAs produced on average in each cell). The coefficients of variation (CV) are 0.008 and 0.024, respectively, differing significantly.

The difference in CV can be explained by the fact that the detailed model has more sources of stochasticity. In the reduced model there is only one stochastic event, transcription initiation, while the detailed model has additional noise sources, such as pauses, arrests, premature terminations, and traffic.

As said, a traffic event occurs (in the detailed model) when an RNAP molecule is impeded to move forward due to the presence of another RNAP molecule in the nucleotides ahead. Traffic can play a major role in transcriptional dynamics (Roussel and Zhu, 2006b). Thus, we measure traffic occurrences, i.e., when an RNAP molecule is inhibited from moving forward due to the presence of another RNAP occupying subsequent nucleotides. It is noted that while traffic can decrease the elongation rate, it can also have the opposite effect, by diminishing the duration of pauses and arrests (Epshtein and Nudler, 2003) due to collisions between RNAP's (reactions 14, 15, 18, and 19). The binned distribution of the number of cells with total number of traffic events, from a set of 1000 cells, is shown in Figure 3.

The repression to which the gene is subject to is controlled by the propensity of reaction 2 (Gillespie, 1977), which is determined by the product $k_{\text{rep}} \cdot [X]$. We can observe how the number of traffic events varies with the number of transcription events by repeating the previous simulation for cells with various values of k_{rep} .

As seen in Figure 4, traffic increases sharply when $k_{\text{rep}} < 0.2$. Thus, one could expect that the transcriptional dynamics of the two models match for $k_{\text{rep}} > 0.2$. We used the Kolmogorov-Smirnov test (KS test) to assess the similarity of the transcriptional dynamics of reduced and detailed model, for each value of k_{rep} . Usually, a threshold of 0.1 for the p-bias is set to accept or reject the null hypothesis. The two models matched only for $k_{\text{rep}} > 0.8$, meaning that traffic is not the only cause for differences in the dynamics of the

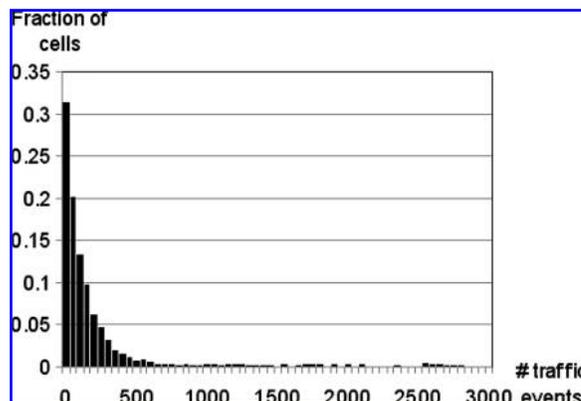


FIG. 3. Normalized binned distribution of the fraction of cells, out of 1000 cells, with a given number of traffic events during their lifetime of 3300s (the bin size is 50).

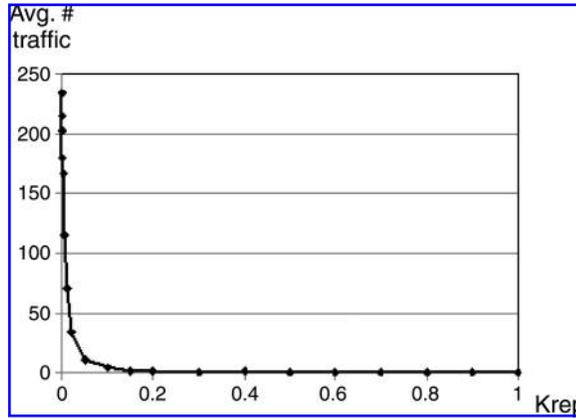


FIG. 4. Average number of traffic events in single cells, averaged over 1000 cells per data point, during their lifetime of 3300 s, for cells with promoters subject to various degrees of repression.

two models as previously assumed (Roussel and Zhu, 2006a). Even when traffic is not significant, as soon as the noise in the elongation phase due to the occurrence of pauses, arrests, premature terminations, etc starts playing a role, the dynamics of the two models diverge significantly (see Table 2 for the frequency of occurrence of these events).

One important difference between the dynamics of the two models, only partially due to traffic events in the detailed model, concerns the length of the intervals between transcription completions. From simulations of 1000 cells, where the repressor molecule is absent, we accessed that the mean time interval between completions is ~ 42.6 s (reduced model) and ~ 42.7 s (detailed model), but due to the stochasticity of the processes occurring during elongation, the standard deviations differ significantly, equaling 2.64 s in the reduced model and 18.9 s in the detailed model.

The amount of traffic caused by each alternative reaction pathway “separately” can be estimated by simulating cells where each alternative reaction pathway is removed, one at a time. For example, to remove pauses one sets $k_{pause} = 0$. In Table 3, we show the average number of traffic events observed in 1000 cells, with several alternative reaction pathways removed, one at a time.

As seen from Table 3, three alternative reaction pathways (pause, arrest and pyrophosphorolysis) contribute to traffic (the other alternative reaction pathways do not have a significant contribution), confirming that the number of traffic events is significantly underestimated if these alternative reaction pathways are not modeled.

Gene expression has been shown to occur by bursts, that is, several RNAs produced in short time intervals, followed by periods of inactivity (Berg, 1978) (Golding et al., 2005; Yu et al., 2006). This is accurately captured by reduced and detailed models (Zhu et al., 2007) given an accurate initiation rate of transcription.

TABLE 2. AVERAGE NUMBER OF EVENTS DURING CELLS’ LIFETIME, AVERAGED OVER 1000 CELLS, FOR VARIOUS VALUES OF k_{rep}

k_{rep}	RNA	Erroneous RNA	Premature terminations	Pauses	Arrests	Pyrophos.	Traffic events
0	73.4	1.91	0.22	697.7	0.347	948.011	234.5
0.1	11.0	0.26	0.04	103.5	0.058	142.077	4.71
0.2	5.9	0.15	0.03	55.7	0.025	76.286	1.149
0.3	4.1	0.10	0.02	39.0	0.018	53.466	0.181
0.4	3.1	0.09	0.01	29.2	0.018	40.235	1.283
0.5	2.5	0.06	0.01	23.9	0.008	32.801	0.159
0.6	2.0	0.05	0.01	19.1	0.007	25.984	0.026
0.7	1.7	0.04	0.004	16.4	0.014	22.437	0
0.8	1.63	0.039	0.003	15.4	0.012	20.751	0
0.9	1.435	0.041	0.004	13.7	0.005	18.519	0
1	1.252	0.035	0.008	12.1	0.009	16.361	0

TABLE 3. AVERAGE AND COEFFICIENT OF VARIATION OF THE NUMBER OF TRAFFIC EVENTS PER CELL (1000 CELLS SIMULATED)

<i>Reaction pathways</i>	<i>P0</i>	$\mu_{traffic}$	$cv_{traffic}$
Detailed model	0.12	204.13	1.75
$k_{pyro} = 0$	0.11	203.83	1.85
$k_{arrest} = 0$	0.15	171.30	2.02
$k_{pause} = 0$	0.70	41.461	3.98
$k_{pyro}, k_{arrest}, k_{pause} = 0$	0.99	9.8430	14.5

In the first column, P0 is the fraction of cells where traffic events did not occur.

However, if the gene is not subject to repression, when observed in detail, it's visible that the time series of RNA production of cells modeled with detailed and reduced models differ. Namely, in small time scales (~ 60 s), while in the reduced model the RNAs are produced in intervals equal or larger than the promoter time delay, the RNAs in the detailed model can appear in "pulses" that is, several RNAs completions can occur within an interval smaller than the promoter delay length.

The pulses are a result of the stochasticity at the elongation stage. For example, an RNAP might be delayed due to one or several events, and a following RNAP might approach it at the strand. The RNAP's will then eventually complete transcription separated by a very short time interval. Due to the minimum separation of 25 nucleotides, the minimum interval between completions is ~ 1 s, which is much smaller than the ~ 40 s which initially separate the two RNAP's due to the promoter delay.

This pulsing in transcription completions, combined with the stochastic nature of decay events, explains the higher variability in the RNA concentration in small time scales and thereby, proteins concentration, since RNA and protein concentrations correlate in *E. coli* (Golding et al., 2005). Therefore, the pulsing might play a role in cellular dynamics, especially in the contexts in which noise in gene expression plays a role. For example, differentiation is in some cases dependent on concentration thresholds (Kaern et al., 2005), who's variance depends on the size distribution of the bursts and, at a smaller time scale, on the size distribution of pulses. Stronger variance can allow overcoming thresholds otherwise unreachable.

To show the "pulsing" in the RNA concentrations captured by the detailed model, Figure 5 shows a detail of the time series of RNA production in single cells of the two models. The RNAs in the detailed model fluctuate further away from the expected value (attained assuming a constant and continuous production of RNAs over time) than for the reduced model, due to long intervals where no RNA is produced, compensated by the occurrence of pulses of completion of RNAs. Notice that we chose a time interval for which both cells have identical number of RNAs produced in the beginning and end of the time series, to show that this pulsing does not affect mean values of RNA concentration, only its variance in time.

We now compare the dynamics of the pulsing of RNA concentrations during cells' lifetime in the two models with experimental measurements reported in Golding et al., (2005). To model this experiment, in which RNA concentrations were measured at the single cell level in intervals of 60 s (cell lifetime is 6000 s) (Golding et al., 2005), we need, besides the reactions related to transcription, to introduce a decay reaction for the RNAs. Additionally, this gene has ~ 4000 nucleotides (Golding and Cox, 2004).

The average time between the productions of consecutive RNAs was measured to be ~ 150 s (Golding et al., 2005). From this, one can infer that the production rate of RNAs must be at least $(150 * [RNAP])^{-1} s^{-1}$, which equals $0.00024 s^{-1}$. Also, it was measured that the average number of RNAs in the cell after a long transient was ~ 10 (Golding et al., 2005). From this, the decay rate of RNA is set to $0.00067 s^{-1}$.

The initiation rate of transcription, k_{init} , to result in the desired production rate of RNAs must account for eventualities such as premature termination that cause the number of RNAs produced to be smaller than the number of transcription initiations.

Also, the model must account that, during the cell lifetime, the gene promoter region is only available for transcription $\sim 1/6$ of the time, being inactive the rest of the time (Golding et al., 2005). We added this alternative reaction pathway by, at randomly chosen moments, setting the promoter inactive, being available to express $1/6$ of the time.

The k_{init} that matches the measured RNA production rate is $0.00216 s^{-1}$ given which the detailed model matches the measured RNA production rate, RNA decay rate, and RNA quantities at equilibrium of ~ 10 molecules (Golding et al., 2005). The average time to reach the concentration at equilibrium (~ 3000 s) also matches the measurements (Golding et al., 2005).

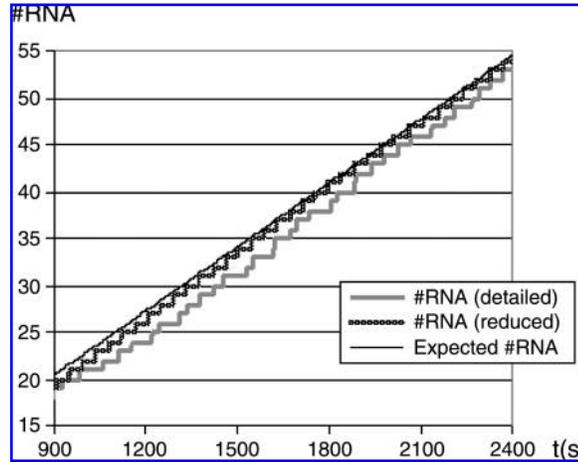


FIG. 5. Time series, from 900 s to 2400 s (1 data point per second), of the number of RNAs produced by a cell with the detailed model of gene expression and by another with the reduced model.

Both reduced and detailed model can accurately match the average number of RNAs produced, since the interval between initiations is long, thus traffic rarely occurs. But when comparing the number of occurrences and amplitude of pulses, the two models differ significantly, and only the detailed model can accurately match the observations.

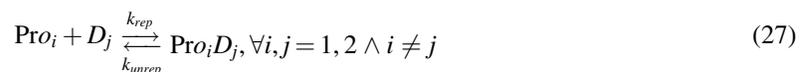
Because in the experiments the measurements were taken in intervals of 1 minute, we measure the RNA time series in the models with this sampling frequency. A “pulse” is defined as an increase in number of RNAs (by 2 or more) occurring in a time step and the pulse amplitude is the amount by which the number of RNAs increases in that interval.

Notice that in the reduced model, RNAs cannot be produced in intervals shorter than 40 s, thus, pulse amplitudes are never higher than 2, unlike in the detailed model. Also, the total number of pulses observed in the detailed model is much higher, matching the experimental measurements, due to the noise at the elongation stage. From 100 simulations of the detailed model, we obtained the average pulsing distribution during cell lifetime, and confronted it to average pulsing distribution obtained from experimental measurements (Fig. 6). As seen from Figure 6, the detailed model accurately captures the dynamics of pulsing.

These pulses can play a significant role in the dynamics of gene regulatory networks, especially considering that RNAs have a short lifetime, thus the importance of a modeling strategy that captures this dynamical feature.

To exemplify how the differences in the stochasticity in transcription of the two models of gene expression can affect the dynamics of gene regulatory networks, we simulate a toggle switch using the two modeling strategies. A toggle switch consists of two mutually repressive genes (Gardner et al., 2000).

Transcription reactions are those of the detailed and the reduced models and the gene length is set to 2445 nucleotides. The other reactions are identical for both toggle switches and follow the model of toggle switch used in Ribeiro et al., (2006). Namely, there are reactions for the translation of RNAs (reaction 25), dimerization and undimerization of the proteins (reaction 26), repression of the promoter of each gene by the other gene’s dimer (reaction 27) and decay of proteins (reactions 28 and 29). These reactions rate constants and delays are set to: $\tau_3 = 2$, $\tau_4 = 58$, $\tau_{5,avg} = 370$ and $\tau_{5,std} = 140$, $k_{tr} = 0.00042$ (Zhu et al., 2007), $k_{rep} = 0.5$, $k_{unrep} = 0.005$, $k_d = 10$ and $k_{ud} = 0.1$, and $p_d = 0.003$ (Zhu et al., 2007). Initially, there is one promoter of each gene, the number of RNAs is 28 (Bremer et al., 2003), and there are 100 ribosomes (rib) (Zhu et al., 2007).



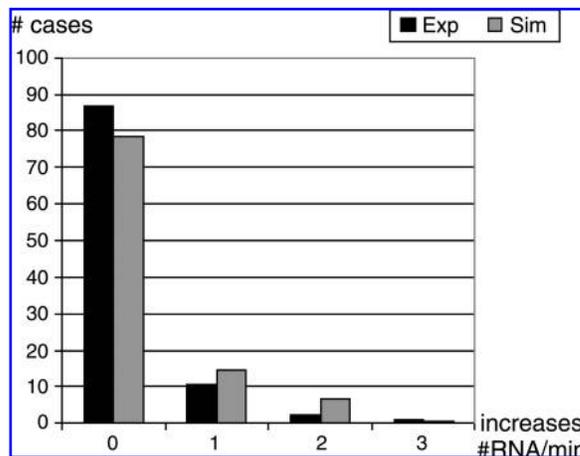


FIG. 6. Average distribution of increments of the RNA quantity, from one time step to the next, during cells' lifetime. Experimental data of three cells (Golding et al., 2005) and detailed model data of 100 cells.

$$Pro_i D_j \xrightarrow{p_d} Pro_i : \forall i, j = 1, 2 \wedge i \neq j \quad (28)$$

$$P_i \xrightarrow{p_d} \emptyset : i = 1, 2 \quad (29)$$

We set high rates for dimerization and decay of monomers, and weak decay for dimers. These two combined conditions result in short lifetimes of monomers and long lifetime of dimers, so that pulses in transcription are mostly responsible for the production of dimers which otherwise almost do not form. Less dimers results in a loss of stability of the toggle switch. Thus, counter-intuitively, the detailed model, with higher noise level at the transcriptional level, is the one whose toggle switch has more stable dynamics, i.e., toggles less frequently.

We simulated 1000 cells, each with a toggle switch, using reduced and detailed models of transcription. On average, the toggle switch with the detailed model has an average toggling frequency of $\sim 2.10^{-6} \text{ s}^{-1}$ while the one with the reduced model has an average toggling frequency of $\sim 7.10^{-6} \text{ s}^{-1}$. The major difference between the two models is the mean level of proteins dimers over time ($D1 + D2$). In the reduced model $(D1 + D2) \sim 13.1$ while in the detailed model $(D1 + D2) \sim 13.9$, explaining the higher toggling frequency in the reduced model. As an example, a time series of dimer amounts (Fig. 7) shows the bistable behavior of the model of toggle switch analyzed.

The difference in the toggling dynamics of the two switches show that events occurring at a fast time scale (proteins' concentration variation due to RNA pulsing) can affect the dynamics of events that occur at a much slower time scale (switching).

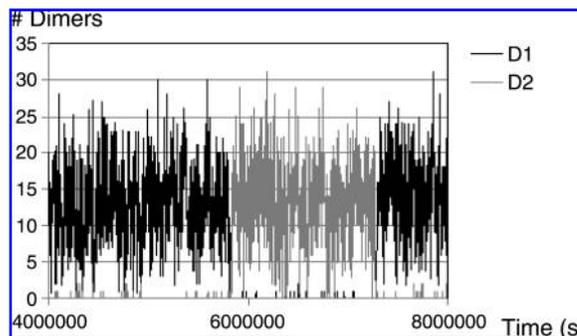


FIG. 7. Sample of a time series of the dimer quantities of a toggle switch whose genes' expression is modeled according to the detailed model.

4. CONCLUSION

We presented a new delayed stochastic model of transcription that models elongation explicitly, at the single nucleotide level, and thereby allows to model pause, arrest, misincorporation and editing, pyrophosphorolysis, and premature termination as explicit events. The model, whose parameter values are taken from experimental measurements, is validated by matching its transcriptional dynamics with a multi-delayed reduced model (Ribeiro et al., 2006) and measurements of gene expression at the single protein level (Yu et al., 2006) of a strongly repressed gene, since in this scenario the dynamics depends mostly on transcription initiation.

When the gene is less repressed, the dynamics of our detailed model and of the reduced delayed model do not match, indicating that the noise at the elongation stage cannot be neglected. Importantly, the differences in the dynamics of reduced and detailed model are not solely due to the emergence of traffic. Events such as pauses and arrests have a non-negligible role even when not sufficiently frequent to cause collisions between RNA polymerases. This result is of importance since single-step multi-delayed models of transcription were assumed accurate as long as there were no collisions between RNAP's in the strand (Roussel and Zhu, 2006a).

Importantly, our detailed model matched, at the single RNA level, the dynamics of transcription of an induced gene (Golding et al., 2005) while the reduced model failed to do so in some aspects, such as in the distribution of intervals between RNA completions. Also, the alternative reaction pathways incorporated in our detailed model have dynamical consequences not captured in previous models with explicit elongation (Roussel and Zhu, 2006a; Voliotis et al., 2008) especially due to causing collisions between RNAP's. The realistic time length for the promoter complex formation (McClure, 1980) here modeled is also nonnegligible, since it also plays a very significant role (e.g., in traffic intensity).

One important feature observed in the dynamics of an induced gene modeled with the detailed model is the existence of pulses in the time series of RNA production (RNA completions separated by a time interval smaller than the minimum interval between transcription initiations). The simulation results matched the average number of pulses and the distribution of pulse intensities obtained from measurements of gene expression at the single RNA level (Golding et al., 2005), showing that when gene expression is not strongly repressed, the detailed model dynamics is more accurate than the delayed single-step model.

Several works focusing on the dynamics of single gene expression have suggested that noise at the elongation stage might play a relevant role in the dynamics of gene regulatory networks (Golding et al., 2005; Voliotis et al., 2008; Ozbudak et al., 2002; Käern et al., 2005). Here, we investigated some of the consequences by comparing the dynamics of two toggle switches, where in one transcription is modeled by the reduced model and in the other, modeled by the detailed model.

We showed that, as suggested, the transcriptional noise arising at the elongation stage does affect transcriptional dynamics of single genes, even in the absence of traffic, and affects significantly the dynamics of the toggle switch, even though the frequency of toggling events is much smaller than the frequency of the fluctuations in RNA levels due to the noise at the elongation stage.

This modeling strategy provides means to investigate *in silico* the role that each of the alternative reaction pathways included plays in transcriptional dynamics. Here, we focused on how these alternative reaction pathways contribute to the emergence of traffic and to the variability in RNA time series of individual cells. More importantly, we showed that the dynamics of gene networks is affected by the noise arising at the elongation stage. Many other questions require future addressing, such as how the alternative reaction pathways (e.g., pause) might be used by cells to regulate gene expression.

ACKNOWLEDGMENTS

We thank Sunney Xie for the valuable information on the number of nucleotides of the Tsr-venus gene. The study was supported by the Academy of Finland (project no. 213462, Finnish Centre of Excellence program 2006-11).

DISCLOSURE STATEMENT

No competing financial interests exist.

REFERENCES

- Alberts, B., Johnson, A., Lewis, J., et al. 2002. *Molecular Biology of the Cell*. Garland Science, New York.
- Arkin, A., Ross, J., and McAdams, H. 1998. Stochastic kinetic analysis of developmental pathway bifurcation in phage λ -infected *E. coli* cells. *Genetics* 149, 1633–1648.
- Artsimovitch, I., and Landick, R. 2000. Pausing by bacterial RNA polymerase is mediated by mechanistically distinct classes of signals. *Proc. Natl. Acad. Sci. USA* 97, 7090–7095.
- Beckskei, A., Kaufmann, B.B., and van Oudenaarden, A. 2005. Contributions of low molecule number and chromosomal positioning to stochastic gene expression. *Nat. Genet.* 37, 937–944.
- Berg, O.G. 1978. A model for the statistical fluctuations of protein numbers in a microbial population. *J. Theor. Biol.* 71, 587–603.
- Bernstein, J.A., Khodursky, A.B., Lin, P.H., et al. 2002. Global analysis of mRNA decay and abundance in *Escherichia coli* at single-gene resolution using two-color fluorescent DNA microarrays. *Proc. Natl. Acad. Sci. USA* 99, 9697–9702.
- Bon, M., McGowan, S.J., and Cook, P.R. 2006. Many expressed genes in bacteria and yeast are transcribed only once per cell cycle. *FASEB J.* 20, 1721–1723.
- Borukhov, S., Lee, J., and Laptenko, O. 2005. Bacterial transcription elongation factors: new insights into molecular mechanism of action. *Mol. Microbiol.* 55, 1315–1324.
- Bratsun, D., Volfson, D., Tsimring, L.S., et al. 2005. Delay-induced stochastic oscillations in gene regulation. *Proc. Natl. Acad. Sci. USA* 102, 14593–14598.
- Bremer, H., Dennis, P., and Ehrenberg, M. 2003. Free RNA polymerase and modeling global transcription in *Escherichia coli*. *Biochimie* 85, 597–609.
- Core, L.J., and Lis, J.T. 2008. Transcription regulation through promoter-proximal pausing of RNA polymerase II. *Science* 319, 1791–1792.
- Darzacq, X., Shav-Tal, Y., de Turris, V., et al. 2007. *In vivo* dynamics of RNA polymerase II transcription. *Nat. Struct. Mol. Biol.* 14, 796–806.
- Davenport, R.J., Wuite, G.J., Landick, R., et al. 2000. Single-molecule study of transcriptional pausing and arrest by *E. coli* RNA polymerase. *Science* 287, 2497–2500.
- Elowitz, M.B., Levine, A.J., Siggia, E.D., et al. 2002. Stochastic gene expression in a single cell. *Science* 297, 1183–1186.
- Epshtein, V., and Nudler, E. 2003. Cooperation between RNA polymerase molecules in transcription elongation. *Science* 300, 801–805.
- Erie, D.A., Hajiseyedjavadi, O., Young, M.C., et al. 1993. Multiple RNA polymerase conformations and GreA: control of the fidelity of transcription. *Science* 262, 867–873.
- Gaffney, E.A., and Monk, N.A. 2006. Gene expression time delays and Turing pattern formation systems. *Bull. Math. Biol.* 68, 99–130.
- Gardner, T.S., Cantor, C.R., and Collins, J.J. 2000. Construction of a genetic toggle switch in *Escherichia coli*. *Nature* 403, 339–342.
- Gillespie, D.T. 1977. Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.* 81, 2340–2361.
- Golding, I., and Cox, E.C. 2004. RNA dynamics in live *Escherichia coli* cells. *Proc. Natl. Acad. Sci. USA* 101, 11310–11315.
- Golding, I., Johan Paulsson, J., Zawilski, S.M., et al. 2005. Real-time kinetics of gene activity in individual bacteria. *Cell* 123, 1025–1036.
- Greive, S.J., and von Hippel, P.H. 2005. Thinking quantitatively about transcriptional regulation. *Nat. Rev. Mol. Cell. Biol.* 6, 221–232.
- Greive, S.J., Weitzel, S.E., Goodarzi, J.P., et al. 2008. Monitoring RNA transcription in real time by using surface plasmon resonance. *Proc. Natl. Acad. Sci. USA* 105, 3315–3320.
- Grundy, F.J., and Henkin, T.M. 2006. From ribosome to riboswitch: control of gene expression in bacteria by RNA structural rearrangements. *Crit. Rev. Biochem. Mol. Biol.* 41, 329–338.
- Hansen, U.M., and McClure, W.R. 1980. Role of the sigma subunit of *Escherichia coli* RNA polymerase in initiation. II. Release of sigma from ternary complexes. *J. Biol. Chem.* 255, 9564–9570.
- Herbert, K.M., La Porta, A., Wong, B.J., et al. 2006. Sequence-resolved detection of pausing by single RNA polymerase molecules. *Cell* 125, 1083–1094.
- Kaern, M., Elston, T., Blake, W., et al. 2005. Stochasticity in gene expression: from theories to phenotypes. *Nat. Rev. Genet.* 6, 451–464.
- Kapanidis, A.N., Margeat, E., Ho, S.O., et al. 2006. Initial transcription by RNA polymerase proceeds through a DNA-scrunching mechanism. *Science* 314, 1144–1147.
- Kierzek, A.M., Zaim, J., and Zielenkiewicz, P. 2001. The effect of transcription and translation initiation frequencies on the stochastic fluctuations in prokaryotic gene expression. *J. Biol. Chem.* 276, 8165–8172.

- Kosuri, S., Kelly, J.R., and Endy, D. 2007. TABASCO: a single molecule, base-pair resolved gene expression simulator. *BMC Bioinform.* 8, 480.
- Landick, R. 2006. The regulatory roles and mechanism of transcriptional pausing. *Biochem. Soc. Trans.* 34, 1062–1066.
- Lewin, B. 2008. *Genes IX*. Jones and Bartlett Publishers, New York.
- Lutz, R., Lozinski, T., Ellinger, T., et al. 2001. Dissecting the functional program of *Escherichia coli* promoters: the combined mode of action of Lac repressor and AraC activator. *Nucleic Acids Res.* 29, 3873–3881.
- McAdams, H.H., and Arkin, A. 1997. Stochastic mechanisms in gene expression. *Proc. Natl. Acad. Sci. USA* 94, 814–819.
- McAdams, H.H., and Arkin, A. 1998. Simulation of prokaryotic genetic circuits. *Annu. Rev. Biophys. Biomol. Struct.* 27, 199–224.
- McClure, W.R. 1980. Rate-limiting steps in RNA chain initiation. *Proc. Natl. Acad. Sci. USA* 77, 5634–5638.
- Mooney, R.A., Darst, S.A., and Landick, R. 2005. Sigma and RNA polymerase: an on-again, off-again relationship? *Mol. Cell* 20, 335–345.
- Neuman, K.C., Abbondanzieri, E.A., Landick, R., et al. 2003. Ubiquitous transcriptional pausing is independent of RNA polymerase backtracking. *Cell* 115, 437–447.
- Ozbudak, E.M., Thattai, M., Kurtser, I., et al. 2002. Regulation of noise in the expression of a single gene. *Nat. Genet.* 31, 69–73.
- Park, J.S., Marr, M.T., and Roberts, J.W. 2002. *E. coli* transcription repair coupling factor (Mfd protein) rescues arrested complexes by promoting forward translocation. *Cell* 109, 757–767.
- Raser, J.M., and O’Shea, E.K. 2004. Control of stochasticity in eukaryotic gene expression. *Science* 304, 1811–1814.
- Ribeiro, A.S., Zhu, R., and Kauffman, S.A. 2006. A general modeling strategy for gene regulatory networks with stochastic dynamics. *J. Comput. Biol.* 13, 1630–1639.
- Ribeiro, A.S., and Lloyd-Price, J. 2007. SGN Sim—a stochastic genetic networks simulator. *Bioinformatics* 23, 777–779.
- Ribeiro, A.S. 2007. Effects of coupling strength and space on the dynamics of coupled toggle switches in stochastic gene networks with multiple-delayed reactions. *Phys. Rev. E* 75, 061903.
- Roberts, J.W. 2006. RNA polymerase, a scrunching machine. *Science* 314, 1097–1098.
- Roussel, M.R., and Zhu, R. 2006a. Stochastic kinetics description of a simple transcription model. *Bull. Math. Biol.* 68, 1681–1713.
- Roussel, M.R., and Zhu, R. 2006b. Validation of an algorithm for delay stochastic simulation of transcription and translation in prokaryotic gene expression. *Phys. Biol.* 3, 274–284.
- Shundrovsky, A., Santangelo, T.J., Roberts, J.W., et al. 2004. A single-molecule technique to study sequence-dependent transcription pausing. *Biophys. J.* 87, 3945–3953.
- Süel, G., Garcia-Ojalvo, J., Liberman, L., et al. 2006. An excitable gene regulatory circuit induces transient cellular differentiation. *Nature* 440, 545–550.
- Uptain, S.M., Kane, C.M., and Chamberlin, M.J. 1997. Basic mechanisms of transcript elongation and its regulation. *Annu. Rev. Biochem.* 66, 117–172.
- Voliotis, M., Cohen, N., Molina-Paris, C., et al. 2008. Fluctuations, pauses and backtracking in DNA transcription. *Biophys. J.* 94, 334–348.
- Vogel, U., and Jensen, K.F. 1994. The RNA chain elongation rate in *Escherichia coli* depends on the growth rate. *J. Bacteriol.* 176, 2807–2813.
- Yu, J., Xiao, J., Ren, X., et al. 2006. Probing gene expression in live cells, one protein molecule at a time. *Science* 311, 1600–1603.
- Zhu, R., Ribeiro, A.S., Salahub, D., and Kauffman, S.A. 2007. Studying genetic regulatory networks at the molecular level: delayed reaction stochastic models. *J. Theor. Biol.* 246, 725–745.

Address reprint requests to:

Dr. Andre S. Ribeiro
 Computational Systems Biology Research Group
 Department of Signal Processing
 Tampere University of Technology
 P.O. Box 553
 FI-33101 Tampere, Finland

E-mail: andre.sanchesribeiro@tut.fi

This article has been cited by:

1. Jarno Makela, Jason Lloyd-Price, Olli Yli-Harja, Andre S Ribeiro. 2011. Stochastic sequence-level model of coupled transcription and translation in prokaryotes. *BMC Bioinformatics* **12**:1, 121. [[CrossRef](#)]
2. Olli-Pekka Smolander, Meenakshisundaram Kandhavelu, Henrik Mannerstrom, Eero Lihavainen, Shanmugapriya Kalaichelvan, Shannon Healy, Olli Yli-Harja, Matti Karp, Andre S Ribeiro. 2011. Cell-to-cell diversity in protein levels of a gene driven by a tetracycline inducible promoter. *BMC Molecular Biology* **12**:1, 21. [[CrossRef](#)]
3. Andre S. Ribeiro, Antti Häkkinen, Shannon Healy, Olli Yli-Harja. 2010. Dynamical effects of transcriptional pause-prone sites. *Computational Biology and Chemistry* **34**:3, 143-148. [[CrossRef](#)]
4. Andre S. Ribeiro, Antti Häkkinen, Henrik Mannerström, Jason Lloyd-Price, Olli Yli-Harja. 2010. Effects of the promoter open complex formation on gene expression dynamics. *Physical Review E* **81**:1. . [[CrossRef](#)]