Stochastic modeling of regulation of gene expression by multiple small RNAs

Charles Baker, Tao Jia, and Rahul V. Kulkarni

Phys. Rev. E **85**, 061915 — Published 20 June 2012

DOI: 10.1103/PhysRevE.85.061915

Stochastic modeling of regulation of gene expression by multiple small RNAs

Charles Baker,* Tao Jia,† and Rahul V. Kulkarni‡
Department of Physics,
Virginia Polytechnic Institute and State University,
Blacksburg, VA 24061

A wealth of new research has highlighted the critical roles of small non-coding RNAs (sRNAs) in diverse processes such as quorum sensing and cellular responses to stress. The pathways controlling these processes often have a central motif comprised of a master regulator protein whose expression is controlled by multiple sRNAs. However, the stochastic gene expression of a single target gene regulated by multiple sRNAs is currently not well understood. To address this issue, we analyze a stochastic model of regulation of gene expression by multiple sRNAs. For this model, we derive exact analytic results for the regulated protein distribution including compact expressions for its mean and variance. The derived results provide novel insights into the roles of multiple sRNAs in fine-tuning the noise in gene expression. In particular, we show that, in contrast to regulation by a single sRNA, multiple sRNAs provide a mechanism for independently controlling the mean and variance of the regulated protein distribution.

I. INTRODUCTION

Small non-coding RNAs (sRNAs), such as microRNAs (miRNAs) and prokaryotic small RNAs, are known to play a central role in diverse cellular pathways that bring about global changes in gene expression [1, 2]. In several cases, such global changes are coordinated by a master regulatory protein whose expression is controlled by multiple sRNAs [3–5]. Examples include regulation of the master regulator in bacterial quorum-sensing pathways by multiple sRNAs [6] and regulation of the alternative sigma factor σ^s by four distinct sRNAs, each of which responds to different environmental stresses [3]. In eukaryotes, recent experiments have shown that the messenger RNA (mRNA) *p21*, which plays a critical role in tumor suppression, is regulated by 28 different sRNAs [4, 5]. Despite its importance in coordinating critical cellular processes such as stress response, quorum sensing, and tumor suppression, the role of multiple sRNAs in regulating the expression of a single target gene is not yet well understood [4, 5]. In this work, we address this issue by analyzing a stochastic model that elucidates potential functional roles for this important regulatory motif.

Regulation of gene expression by sRNAs is a post-transcriptional process: sRNAs can bind to mRNAs and control protein production by altering mRNA stability or by regulating translational efficiency [1]. The intrinsic stochasticity of the underlying biochemical reactions can produce significant variation ('noise') in gene expression among individual cells in isogenic populations [7–11]. Although in some cases noise in gene expression can have deleterious effects and thus needs to be limited, in other cases such noise is utilized and indeed required by the cell, e.g. for processes leading to probabilistic cell-fate decisions [12, 13]. Furthermore, it has been argued that noise in gene expression could be advantageous under conditions of high stress, since variability in a population provides a bet-hedging strategy that can enable survival [14, 15]. Regulation of the noise in gene expression is thus essential for the proper functioning of several cellular processes. Since sRNAs regulate critical cellular processes, understanding their role in fine-tuning the noise in gene expression is of fundamental importance [13].

A quantitative understanding of the cellular functions of sRNAs is aided by the development of models, which can often produce insights that guide future experiments. In recent research, several models which include regulation by sRNAs have been developed [16–22]. Since many sRNAs are known to repress gene expression, most previous models have focused on regulation by irreversible stoichiometric degradation [16–20]. However, sRNAs can affect not only mRNA degradation rates but also protein production rates, and the corresponding biochemical reactions are, in general, reversible [1, 23]. Furthermore, not all sRNAs repress gene expression; there are sRNAs which are known to activate gene expression and even some which can switch from activating to repressing in response to cellular signals [23, 24]. To quantify the corresponding effects on stochastic gene expression, a general model which includes the different mechanisms of sRNA-based regulation needs to be analyzed. Such a model, for the case of a single sRNA regulator, has been developed in recent work [25]. Analysis of this model and its extension to multiple sRNAs thus provides a means of addressing outstanding questions about the impact of different modes of regulation by sRNAs on the noise in gene expression.

In this work, we generalize our previous model [25] to analyze the case of multiple sRNAs regulating a single mRNA target. Specifically, we derive exact analytic expressions for the generating function of a protein burst distribution resulting from the regulation of a single target by an arbitrary number of sRNAs and provide results relating the burst and steady-state distributions. For the

* cbaker@vt.edu
† tjia@vt.edu
‡ kulkarni@vt.edu

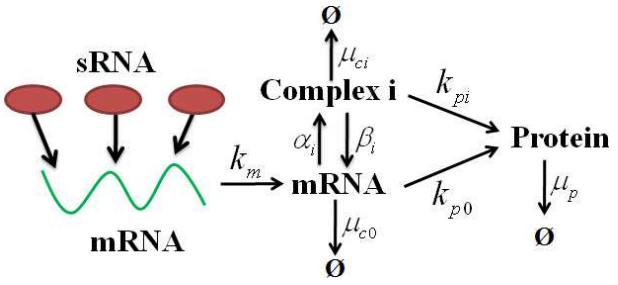

FIG. 1. (Color online) Schematic illustration of regulation of gene expression by multiple sRNAs. In the full reaction scheme, there are N different regulators and the kinetic scheme is shown for the i^{th} sRNA regulator. Note that for the mRNA to transition from one complex to another, it must first return to its unbound state before forming a new complex.

burst distribution, we obtain compact analytic expressions for both the mean and variance of the regulated protein distribution. We analyze these results for the case of a single regulator and derive insights into the effects of different modes of regulation by sRNAs. These results are then contrasted with features associated with regulation by multiple sRNAs. In particular, we show that, in contrast to regulation by a single sRNA, regulation by multiple sRNAs provides the cell with a mechanism to independently fine-tune both the mean and variance of the regulated protein distribution.

II. BACKGROUND AND MODEL FRAMEWORK

In the following, we consider protein production from an mRNA which is regulated independently by N sRNAs. The corresponding reaction scheme is shown in Fig. 1. In this scheme, the mRNA has $N+1$ possible states, with the states $i = 1, ..., N$ denoting mRNA bound to the i^{th} sRNA regulator to form complex i. For notational simplicity, we denote the unbound mRNA state as complex 0. An unbound mRNA is produced with rate k_m and forms a particular complex i with rate α_i. This complex can either dissociate with a rate β_i, decay with a rate μ_{c_i}, or initiate protein production with a rate k_{p_i}.

In the model, sRNA regulators are taken to be present in large amounts such that fluctuations in their concentration can be ignored, an assumption that holds for many known sRNAs in regulatory networks [26–28]. Correspondingly, we model the binding of an sRNA to an mRNA by the formation of a complex, such that the rate of complex formation, α_i, and complex dissociation, β_i, are taken to be constant. We further note that the proposed kinetic scheme applies to general post-transcriptional regulators, for example the regulators can be proteins instead of sRNAs. However, since our motivation derives from the known instances of multiple sRNAs regulating a common target, we will refer to the post-transcriptional regulators as sRNAs in this paper.

For the above reaction scheme, we derive the protein burst distribution, $P_{b,N}(n)$ produced by a single mRNA which is then generalized to include the case of transcriptional bursting. Previous work has shown how, in the limit of independent bursts such that the protein lifetime is much greater than the mRNA lifetime, the noise in burst distributions can be connected to the noise in steady-state protein distributions [29, 30]. We investigate this connection for the case of multiple sRNAs regulating a single target and provide expressions for the generating function of the protein steady-state distribution.

III. BURST DISTRIBUTIONS

We begin by defining the function $f_i(n,t)$ which denotes the probability that n proteins have been produced and the mRNA is in state i at time t. The time-evolution of these probabilities is governed by the master equation:

$$\begin{aligned}\frac{\partial f_0(n,t)}{\partial t} &= k_{p_0}(f_0(n-1,t) - f_0(n,t)) \\ &\quad -(\mu_{c_0} + \sum_{i=1}^{N}\alpha_i) f_0(n,t) + \sum_{i=1}^{N}\beta_i f_i(n,t) \\ \frac{\partial f_i(n,t)}{\partial t} &= k_{p_i}(f_i(n-1,t) - f_i(n,t)) \\ &\quad -(\mu_{c_i} + \beta_i) f_i(n,t) + \alpha_i f_0(n,t) \end{aligned} \tag{1}$$

The initial condition corresponds to creation of a single unbound mRNA and no proteins in the system at time $t = 0$, i.e. $f_0(0,0) = 1$. The protein burst distribution produced from a single mRNA, $P_{b,N}(n)$, can be determined using

$$P_{b,N}(n) = \int_0^\infty \sum_{i=0}^{N} f_i(n,t)\mu_{c_i} dt. \tag{2}$$

The above equation can be understood as follows: given that the mRNA is in state i at time t, the probability it decays in the following time interval $[t, t+dt]$ is $\mu_{c_i} dt$. The burst probability distribution, Eq. (2), is then obtained by first conditioning $P_{b,N}(n)$ on the mRNA being in the state i at time t and then degrading during the interval $[t, t+dt]$. The corresponding probability is $f_i(n,t)\mu_{c_i} dt$. Then, $P_{b,N}(n)$ is obtained by integrating over the time of decay and summing over the possible states prior to decay.

We now define dimensionless parameters which allow the results to be presented in a compact form. First we define $n_i = \frac{k_{p_i}}{\mu_{c_i}}$ for $i \geq 0$, which can be interpreted as the mean number of proteins produced by an unregulated mRNA with translational efficiency k_{p_i} and degradation rate μ_{c_i}. Now, if we let $\xi_i = \frac{k_{p_i}}{\beta_i + \mu_{c_i}}$, we can interpret ξ_i as the mean number of proteins produced during a single sojourn in state i where in the limit that the formation of

the complex becomes irreversible, i.e. $\beta_i \to 0$, this mean reduces to n_i. We then define weighting parameters $\omega_i = \frac{\alpha_i}{\beta_i+\mu_{c_i}}\left(\frac{\mu_{c_i}}{\mu_{c_0}}\right)$ for $i > 0$. Finally, by setting $\omega_0 = 1$ and $\xi_0 = 0$ we define 'weight functions' $\Omega_i(z) = \omega_i \frac{1}{1+\xi_i(1-z)}$.

Using the above definitions, we obtain the following exact expression for the generating function of the protein burst distribution for the case of a single mRNA (the derivation is presented in the Appendix).

$$G_{b,N}(z) = \frac{\sum_{i=0}^{N}\Omega_i(z)}{\sum_{i=0}^{N}\Omega_i(z) + \sum_{i=0}^{N}\Omega_i(z)n_i(1-z)} \tag{3}$$

The above expression for the generating function can be extended to include the case of transcriptional bursting. Specifically, instead of a single mRNA produced with rate k_m, we consider a burst of mRNAs arriving with the same rate. The number of mRNAs produced in a single burst can have an arbitrary distribution and we denote the corresponding generating function by $G_m(z)$. Assuming that each mRNA contributes independently to the production of proteins, the generating function for the protein burst distribution ($\overline{P}_{b,N}(n)$) is given by $\overline{G}_{b,N}(z) = G_m(G_{b,N}(z))$ (where $G_{b,N}(z)$ is the generating function for the burst of proteins produced from a single mRNA) [30].

Previous work [31] has shown that, for a range of models of gene expression, the probability of producing $m \geq 1$ mRNAs in a single burst can be represented by the conditional geometric distribution $P(m) = (1-p)^{m-1}p$ with mean mRNA burst size $m_b = 1/p$ ($p \leq 1$). Focusing on this case, we find that the generating function for the protein burst distribution is given by

$$\overline{G}_{b,N}(z) = \frac{p\sum_{i=0}^{N}\Omega_i(z)}{p\sum_{i=0}^{N}\Omega_i(z) + \sum_{i=0}^{N}\Omega_i(z)n_i(1-z)} \tag{4}$$

where for $p = 1$ we recover Eq. (3) (see the Appendix for a derivation of this result). Furthermore, for $N = 0$, i.e. the unregulated case, the generating function reduces to $\overline{G}_{b,0}(z) = p/(p + n_0(1-z))$ in agreement with previous work showing that the protein burst distribution is a geometric distribution with mean $n_0 = \frac{k_{p_0}}{\mu_{c_0}}$ [25, 32]. Eq. (4) provides the generalization of this result for the case of regulation by N sRNAs.

An important mechanism of regulation by sRNAs corresponds to the case that sRNA binding prevents ribosome access and thus blocks translation. For the case that all the regulators act to fully repress translation, i.e. $k_{p_i} = 0$ for $i > 0$, we have $\Omega_i(z) = \omega_i$. Correspondingly, the generating function reduces to the generating function of an unregulated system with a renormalized mean given by $\frac{n_0}{\sum_{i=0}^{N}\omega_i}$. This observation indicates that regulation by sRNAs which function by fully repressing translation is, in principle, reversible: for arbitrary concentrations of the sRNA regulator, by appropriately adjusting the parameter k_{p_0}, the regulated protein distribution in the presence of sRNAs can be made identical to the distribution for the unregulated case (i.e. prior to introduction of the sRNAs).

IV. MEAN AND NOISE OF PROTEIN BURST DISTRIBUTIONS

For the general case, using Eq. (4), we derive compact analytic expressions for the burst mean, n_{b_N}, and coefficient of variance, $\sigma^2_{b_N}/n^2_{b_N}$. The mean (scaled by the unregulated mean) is given by

$$\frac{n_{b_N}}{n_{b_0}} = 1 + F_N \tag{5}$$

and the noise strength (squared coefficient of variance) is given by

$$\frac{{\sigma_{b_N}}^2}{{n_{b_N}}^2} = 1 + \frac{1}{n_{b_N}} + Q_N \tag{6}$$

where

$$F_N = \frac{\sum_{i=0}^{N}\omega_i\,(n_i - n_0)}{\sum_{i=0}^{N}\omega_i n_0} \tag{7}$$

and

$$Q_N = \frac{p\sum_{i,j=0}^{N}\omega_i\omega_j(\xi_i - \xi_j)(n_i - n_j)}{\left(\sum_{i=0}^{N}\omega_i n_i\right)^2} \tag{8}$$

Note that the signs of F_N and Q_N characterize the impact of the sRNAs on the regulated protein distribution. Specifically, the unregulated case has mean n_{b_0}; thus, $F_N < 0$ corresponds to repression whereas $F_N > 0$ corresponds to activation. Similarly, an unregulated protein burst distribution with mean n_{b_N} has a squared coefficient of variance $1 + 1/n_{b_N}$; thus, when $Q_N < 0$ we have noise reduction whereas $Q_N > 0$ corresponds to increased noise strength (relative to an unregulated burst distribution with the same mean). Furthermore, we note that the above expressions illustrate that the parameter n_i characterizes sRNA i as a repressor or an activator; namely, sRNA i is an activator if $n_i > n_0$ and sRNA i is a repressor if $n_i < n_0$.

We now focus on using Eq. (5) and Eq. (6), to elucidate interesting features for the case of regulation by a single sRNA, i.e. $N = 1$. Note that all of the variables in the expressions for the mean and noise strength are always positive (or zero) except for the term $(n_1 - n_0)$. Thus, the sign of F_1 and Q_1 is determined completely by $\Delta_{10} = n_1 - n_0$. When $\Delta_{10} > 0$ the mean and noise strength are higher than their unregulated values. Similarly, when $\Delta_{10} < 0$ both the mean and noise strength are lower than the corresponding unregulated values (except for the case $\xi_i = 0$ for which the noise strength is identical to an unregulated distribution with the same mean). In either case, we note that for a single sRNA

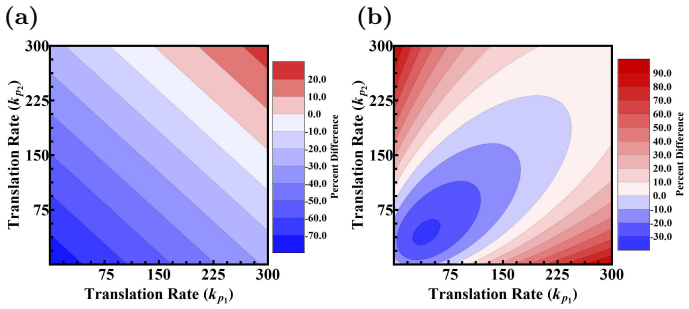

FIG. 2. (Color) Contour plots for the percent change in the mean and noise strength of a two regulator system from the corresponding unregulated values as a function of k_{p_1} and k_{p_2}. (a) Mean: Plot of $M(k_{p_1}, k_{p_2}) = F_2 \cdot 100\%$. Note that along the contour $M(k_{p_1}, k_{p_2}) = -20\%$ the noise strength changes from less than -5% to over 70%. (b) Noise Strength: Plot of $S(k_{p_1}, k_{p_2}) = \frac{Q_2}{1+1/n_{b_2}} \cdot 100\%$. Note that $S(k_{p_1}, k_{p_2})$ contains contours that sweep out a large portion of the plotted (k_{p_1}, k_{p_2}) state space. By proportionally changing the k_p values corresponding to the two regulators, the noise strength can be varied while maintaining the same mean value. The parameters used were $p = 1$, $k_{p_0} = 50$, $\mu_{c_0} = 1$, $\mu_{c_1} = 4.5$, $\mu_{c_2} = 4.5$, $\beta_1 = 1$, $\beta_2 = 0.5$, $\alpha_1 = 2$ and $\alpha_2 = 2$.

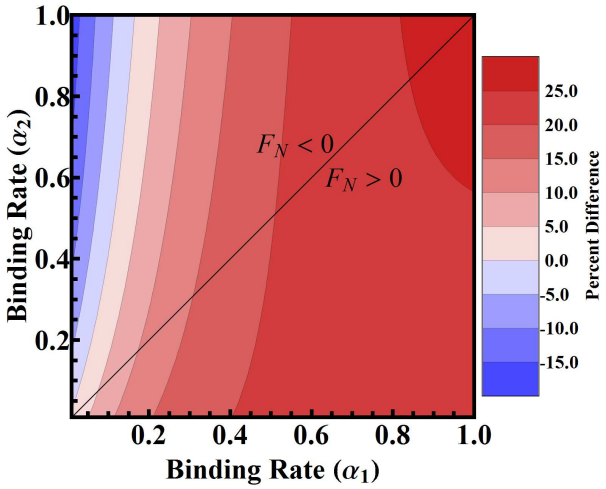

FIG. 3. (Color) Contour plot of the percent change in noise strength of a two regulator pathway from its corresponding unregulated value as a function of α_1 and α_2, i.e. $S(\alpha_1, \alpha_2) = \frac{Q_2}{1+1/n_{b_2}} \cdot 100\%$. The change in mean from the unregulated to regulated pathways, F_N, is positive below and negative above the line $\alpha_1 = \alpha_2$. The parameters used were $k_{p_0} = 50$, $p = 1$, $k_{p_1} = 200$, $k_{p_2} = 72.5$, $\mu_{c_0} = 1$, $\mu_{c_1} = 2.725$, $\mu_{c_1} = 2.725$, $\beta_1 = 0.15$ and $\beta_2 = 0.15$.

regulator present at high concentrations, there is a coupling between the mean and noise strength of the regulated burst distribution such that both cannot be tuned independently, e.g. a decrease in the mean cannot be associated with an increase in the noise strength. Note that in the limit of low sRNA concentration, where the model assumptions do not apply, additional ways of tuning of noise and mean may be possible.

In contrast to the case of regulation by a single sRNA, in the case of regulation by multiple sRNAs, the mean and noise of the protein distribution can be tuned independently. The deviation of the mean from its unregulated value depends solely upon terms of the form $\Delta_{i0} = n_i - n_0$. On the other hand, considering the general form of the noise strength for $N > 1$, we have terms of the form $\tilde{\Delta}_{ij} = (\xi_i - \xi_j)(n_i - n_j)$ that contribute to the deviation from the corresponding unregulated value. Thus, for appropriately chosen parameters, two or more sRNAs can be used to tune both the mean and variance of the regulated protein distribution as discussed below.

Consider the case of regulation by two sRNAs that are maintained at some fixed cellular concentrations. A mRNA target for these sRNAs can arise from the evolution of appropriate sRNA binding sites on the mRNA sequence. For the mRNA target, we assume that the parameters k_{p_1} and k_{p_2} can be tuned based on changes in the sequence and location of the sRNA binding sites. The corresponding variation in the mean and noise strength is shown in Fig. 2. Note that by maintaining a linear relationship between k_{p_1} and k_{p_2}, the mean of the regulated protein distribution can be left unchanged; however, the noise strength can be tuned over a large range. For example, for some choices of the parameters, the mean can be fixed and the noise strength can be varied by over 100% relative to the unregulated distribution (see Fig. 2). In this context, it is interesting to note that it has been observed that several sRNAs have a minimal effect on the mean levels of their regulatory targets. For such targets, sRNAs could be functioning primarily as modulators of noise while giving rise to only a minimal change in mean levels due to regulation [13]. Our results provide quantitative insight into how such regulation can be implemented using multiple sRNA regulators.

The results obtained also illustrate how changing sRNA concentrations can be used to modulate the noise in gene expression. For our model, changes in the concentration of the sRNA regulators effectively alter the binding rates (α_i) to the mRNA. From Eq. (5), we see that for two regulators, by choosing one of the regulators to be a repressor and the other to be an activator, the mean of the regulated protein distribution can be increased ($F_N > 0$) or decreased ($F_N < 0$) by adjusting the relative concentrations of the two regulators. Furthermore, by changing the concentrations of the regulators such that their relative concentration is fixed, the mean of the regulated protein distribution is left unchanged, whereas the variance can be tuned over a range of values. This insight is particularly relevant, given that noise can be advantageous to a cell. In particular, noise in gene expression is known to be especially important in response to stress; a response which is often governed by pathways involving regulation by multiple sRNAs.

V. PROTEIN STEADY-STATE DISTRIBUTIONS

Since sRNAs primarily impact protein production, the previous sections focused on their regulatory effects on protein burst distributions. However, it is also of interest to model how multiple sRNAs impact the protein steady-state distributions, given that steady-state distributions are generally easier to measure experimentally. Using results from previous work [29, 30, 33], the expressions obtained above for burst protein distributions can be used to find corresponding steady-state results. In particular, for arbitrary burst arrival processes, applying results from queueing theory, the steady-state mean is given by $n_s = k_m \tau_m n_b$ where τ_m is the average decay time of mRNAs [29, 34]. Given that the arrival of protein bursts is a Poisson process with rate k_m with each burst producing n proteins where n is drawn from a probability distribution with generating function $\overline{G}_{b,N}(z)$ given by Eq. (4), the generating function of the protein steady-state distribution is given by

$$\overline{G}_{s,N}(x) = \exp\left(\frac{k_m}{\mu_p}\int_1^x \left(\frac{A(y)}{B(y)}\right) dy\right) \tag{9}$$

for

$$A(y) = \sum_{i=0}^{N} \Omega_i(y)\frac{n_i}{p} \tag{10}$$

$$B(y) = \sum_{i=0}^{N} \Omega_i(y) + \sum_{i=0}^{N} \Omega_i(y)\frac{n_i}{p}(1-y) \tag{11}$$

Now, as shown in the Appendix, the integrand in Eq. (9) can be rewritten as a polynomial of degree N divided by a polynomial of degree $N+1$. These polynomials have real coefficients, so a solution to the integral of $G_{s,N}(x)$ can always be found using the method of partial fractions for any given system of regulators. Using Eq. (9), the steady-state mean and squared coefficient of variance are found to be

$$n_{s_N} = \left(\frac{k_m}{\mu_p}\right) n_{b_N} \tag{12}$$

and

$$\frac{\sigma^2_{s_N}}{n^2_{s_N}} = \frac{1}{n_{s_N}} + \frac{1}{k_m/\mu_p}\left[1 + \frac{Q_N}{2}\right] \tag{13}$$

Consider the case of a single post-transcriptional regulator acting via pure repression, i.e. $Q_1 = 0$. The steady-state mean of the protein distribution is given by $n_{s_1} = k_m n_{b_1}/\mu_p$ and the noise is given by $\sigma^2_{s_1}/n^2_{s_1} = \frac{1}{k_m/\mu_p} + \frac{1}{n_{s_1}}$. For sufficiently high mean, $n_{s_1} \gg 1$, the noise is approximately $\sigma^2_{s_1}/n^2_{s_1} \approx \frac{1}{k_m/\mu_p}$. Thus, for a system with protein decay rate μ_p, the steady-state noise can be tuned at the transcriptional level by adjusting the rate of mRNA arrival, k_m, and the mean can be tuned at the post-transcriptional level by changing the value of n_{b_1}. This indicates the potential for specific pathways to achieve fine-tuning of mean and noise through not only pure post-transcriptional regulation but through a combination of transcriptional and post-transcriptional mechanisms.

VI. DISCUSSION

Modulation of gene expression noise can provide a multitude of selective advantages in specific contexts and our model provides a means of gaining insight into molecular mechanisms for achieving such control using multiple post-transcriptional regulators. In the present work, we have obtained closed-form solutions for the generating function of the protein burst distribution (including transcriptional bursting) with regulation by an arbitrary number of sRNAs. Using this generating function, the mean and noise of the protein burst distribution were investigated analytically. In particular, our analysis shows that by adjusting the concentrations of multiple sRNA regulators, a cell can initiate finely-tuned responses to external stimuli, in particular to alter the noise in protein distributions without significantly affecting the mean. The burst expressions were then connected to their steady-state counterparts and indicated the potential for combinations of post-transcriptional and transcriptional to achieve some level of fine-tuning of protein distributions. This suggests that the cell may utilize a combination of post-transcriptional and transcriptional mechanisms to tune protein distributions, with post-transcriptional regulators offering the advantage of faster response to stimuli than transcriptional regulators. This could explain the ubiquity of sRNA regulators in cellular stress response, a process for which gene expression noise is known to be critical and quick tuning of protein levels could be vital to cell survival. In a broader context, the fine control afforded to cells via regulation by multiple sRNAs could be useful in signal integration when multiple environmental stimuli are present simultaneously.

In future work, the interplay between transcriptional and post-transcriptional regulators will be studied in greater detail. In particular, previous work [11] has shown how the noise variance can be tuned independently from the mean based on regulation at the promoter level. While this work obtained results valid for general promoters, the effect of post-transcriptional regulation was not considered. In contrast, the present work includes the effects of post-transcriptional regulation but only considers promoters for which bursts arrive according to a Poisson process. It would thus be of interest in future work to combine both approaches and gain further insight into how the combination of transcriptional and post-transcriptional regulation can be used to fine-tune the noise in protein distributions. Overall, this work provides a basis for future work investigating phenomena

such as signal integration and response in more complex pathways, thus opening new avenues for understanding and modeling stochastic gene expression in a wider class of regulatory networks.

ACKNOWLEDGMENTS

The authors acknowledge funding support from the NSF through award PHY-0957430.

Appendix A: Derivation of the Generating Function for Proteins Produced from a Single mRNA

The general strategy for finding the generating function, $G_{b,N}(z)$, involves taking the Laplace transform of the master equation, Eq. 1, and then taking the limit of the resulting transformed equation to find the desired relations. Specifically,

$$G_{b,N}(z) = \lim_{s\to 0} \sum_{i=0}^{N} \mu_{cI} F_i(z,s) \tag{A1}$$

where we define $F_i(z,s) = \sum_{n=0}^{\infty} z^n L_i(n,s)$ and $L_i(n,s) = \int_0^\infty e^{-st} f_i(n,t)\mathrm{d}t$ is the Laplace transform of $f_i(n,t)$. Now, by taking the Laplace transform of the master equation we find

$$\begin{aligned} sL_0(n,s) &= f_0(n,0) + k_{p_0}(L_0(n-1,s) - L_0(n,s)) \\ &\quad -(\mu_{c_0} + \sum_{i=1}^{N} \alpha_i) L_0(n,s) + \sum_{i=1}^{N} \beta_i L_i(n,s) \\ sL_i(n,s) &= f_i(n,0) + k_{p_i}(L_i(n-1,s) - L_i(n,s)) \\ &\quad -(\mu_{c_i} + \beta_i) L_i(n,s) + \alpha_i L_0(n,s) \end{aligned} \tag{A2}$$

By taking the unilateral z-transform of these expressions and substituting the initial condition $f_0(0,0) = 1$ we find

$$\begin{aligned} F_0(z,s) &= 1 \Big/ \Big(s + \mu_{c_0}\Omega_0(z)(1 + n_0(1-z)) \\ &\quad + \sum_{i=1}^{N} \frac{\mu_{c_0}\omega_i(1+n_i(1-z))}{1 + \xi_i(1-z) + \frac{s}{\mu_{c_i}+\beta_i}} \Big) \\ F_i(z,s) &= \frac{\omega_i \mu_{c_0} F_0(z,s)}{\mu_{c_i} + \mu_{c_i}\xi_i(1-z) + \frac{\mu_{c_i} s}{\mu_{c_i}+\beta_i}} \end{aligned} \tag{A3}$$

Substituting these expressions into Eq. (A1), we obtain the generating function Eq. (3).

Appendix B: Derivation of the Generating Function for Protein Bursts with Transcriptional Bursting

As discussed in the main text, $\overline{G}_{b,N}(z) = G_m(G_{b,N}(z))$ where $G_m(z)$ is the generating function for the production of mRNAs in a burst and $G_{b,N}(z)$ is the generating function for the burst of proteins produced by a single mRNA. For the considered case of geometrically distributed bursts, $G_m(z) = pz/(1 - z(1-p))$ and thus $\overline{G}_{b,N}(z) = pG_{b,N}(z)/(1 - G_{b,N}(z)(1-p))$. Thus, plugging in the result from Eq. (3) derived above we find

$$\overline{G}_{b,N}(z) = \frac{p\sum_{i=0}^{N}\Omega_i(z)}{p\sum_{i=0}^{N}\Omega_i(z) + \sum_{i=0}^{N}\Omega_i(z) n_i (1-z)} \tag{B1}$$

Appendix C: Steady-State Generating Function

As discussed in [30, 33], for a Poisson arrival of bursts, the steady-state protein level is given by

$$\overline{G}_S(x) = \exp\left(\frac{k_b}{\mu_p} \int_1^x \frac{\left(\overline{G}_{b,N}(y) - 1\right) dy}{y-1} \right) \tag{C1}$$

which upon rearrangement yields

$$\overline{G}_S(x) = \exp\left(\frac{k_b}{\mu_p} \int_1^x \frac{p\sum_{i=0}^{N}\Omega_i(y) n_i dy}{\sum_{i=0}^{N}\Omega_i(y)(p + n_i(1-y))} \right) \tag{C2}$$

If we define the function

$$h_i(y) = \prod_{k=0, k\neq i}^{N} (1 + \xi_k(1-y)) \tag{C3}$$

then, by multiplying the numerator and denominator of the steady-state generating function by $h_i(y)$ we can see that this expression reduces to

$$\overline{G}_S(x) = \exp\left(\frac{k_b}{\mu_p} \int_1^x \frac{P(y)}{Q(y)} dy \right) \tag{C4}$$

for

$$P(y) = p\sum_{i=0}^{N} \omega_i n_i h_i(y) \tag{C5}$$

and

$$Q(y) = \sum_{i=0}^{N} \omega_i h_i(y)(p + n_i(1-y)) \tag{C6}$$

where $P(y)$ is in general a polynomial of degree N and $Q(y)$ is in general a polynomial of degree $N+1$.

[1] L. S. Waters and G. Storz, Cell **136**, 615 (2009).
[2] D. Jost, A. Nowojewski, and E. Levine, BMB reports **44**, 11 (2011).
[3] C. L. Beisel and G. Storz, FEMS Microbiology Reviews **34**, 866 (2010).
[4] S. Wu, S. Huang, J. Ding, Y. Zhao, L. Liang, T. Liu, R. Zhan, and X. He, Oncogene **29**, 2302 (2010).
[5] M. E. Peter, Oncogene **29**, 2161 (2010).
[6] D. Lenz, K. Mok, B. Lilley, R. Kulkarni, N. Wingreen, and B. Bassler, Cell **118**, 69 (2004).
[7] M. Kaern, T. C. Elston, W. J. Blake, and J. J. Collins, Nature Reviews. Genetics **6**, 451 (2005).
[8] J. Paulsson, Physics of Life Reviews **2**, 157 (2005).
[9] N. Maheshri and E. K. O'Shea, Annual Review of Biophysics and Biomolecular Structure **36**, 413 (2007).
[10] A. Raj and A. van Oudenaarden, Cell **135**, 216 (2008).
[11] A. Sanchez and J. Kondev, Proc. Natl. Acad. Sci. USA **105**, 5081 (2008).
[12] R. Losick and C. Desplan, Science **320**, 65 (2008).
[13] E. Hornstein and N. Shomron, Nature Genetics **38 Suppl**, S20 (2006).
[14] D. Fraser and M. Kaern, Molecular Microbiology **71**, 1333 (2009).
[15] A. Eldar and M. B. Elowitz, Nature **467**, 167 (2010).
[16] E. Levine, Z. Zhang, T. Kuhlman, and T. Hwa, PLoS Biology **5**, e229 (2007).
[17] N. Mitarai, A. M. C. Andersson, S. Krishna, S. Semsey, and K. Sneppen, Physical Biology **4**, 164 (2007).
[18] P. Mehta, S. Goyal, and N. S. Wingreen, Molecular Systems Biology **4**, 221 (2008).
[19] K. S. Fröhlich and J. Vogel, Current Opinion in Microbiology **12**, 674 (2009).
[20] V. P. Zhdanov, Physics Reports **500**, 1 (2011).
[21] M. Komorowski, J. Miekisz, and A. M. Kierzek, Biophys J **96**, 372 (2009).
[22] M. Osella, C. Bosia, D. Cora, and M. Caselle, PLoS Comput Biol **7** (2011).
[23] A. S. Flynt and E. C. Lai, Nature Reviews Genetics **9**, 831 (2008).
[24] S. Vasudevan, Y. Tong, and J. Steitz, Science **318**, 1931 (2007).
[25] T. Jia and R. V. Kulkarni, Physical Review Letters **105**, 018101 (2010).
[26] C. Ragan, M. Zuker, and M. A. Ragan, PLoS Comput Biol **7**, e1001090 (2011).
[27] D. P. Bartel, Cell **116**, 281 (2004).
[28] J. S. Mattick, Science **309**, 1527 (2005).
[29] T. Jia and R. Kulkarni, Physical Review Letters **106**, 058102 (2011).
[30] V. Elgart, T. Jia, A. T. Fenley, and R. Kulkarni, Physical Biology **8**, 046001 (2011).
[31] P. J. Ingram, M. P. H. Stumpf, and J. Stark, PLoS Comp Biol **4** (2008).
[32] O. G. Berg, J Theor Biol **71**, 587 (1978).
[33] S. Iyer Biswas, *Applications of Methods of Non-equilibrium Statistical Physics to Models of Stochastic Gene Expression*, Ph.D. thesis, Ohio State University (2009).
[34] V. Elgart, T. Jia, and R. Kulkarni, Phys. Rev. E. **82**, 021901 (2010).