Post-transcriptional bursting in genes regulated by small RNA molecules

Guillermo Rodrigo*
Institute for Integrative Systems Biology, CSIC, Universidad de Valencia, 46980 Paterna, Spain

(Received 19 June 2017; revised manuscript received 5 February 2018; published 2 March 2018)

Gene expression programs in living cells are highly dynamic due to spatiotemporal molecular signaling and inherent biochemical stochasticity. Here we study a mechanism based on molecule-to-molecule variability at the RNA level for the generation of bursts of protein production, which can lead to heterogeneity in a cell population. We develop a mathematical framework to show numerically and analytically that genes regulated post transcriptionally by small RNA molecules can exhibit such bursts due to different states of translation activity (on or off), mostly revealed in a regime of few molecules. We exploit this framework to compare transcriptional and post-transcriptional bursting and also to illustrate how to tune the resulting protein distribution with additional post-transcriptional regulations. Moreover, because RNA-RNA interactions are predictable with an energy model, we define the kinetic constants of on-off switching as functions of the two characteristic free-energy differences of the system, activation and formation, with a nonequilibrium scheme. Overall, post-transcriptional bursting represents a distinctive principle linking gene regulation to gene expression noise, which highlights the importance of the RNA layer beyond the simple information transfer paradigm and significantly contributes to the understanding of the intracellular processes from a first-principles perspective.

DOI: 10.1103/PhysRevE.97.032401

I. INTRODUCTION

Living organisms of ranging complexity are continuously subjected to environmental changes that force them to make decisions. In this regard, they have evolved to base their behavior on intricate regulatory networks that allow expressing determinant genes according to dynamic signals [1]. However, the intracellular processes are inherently stochastic, which results in complex dynamic patterns [2,3]. In some cases, decisions mainly rely on deterministic processes, such as the expression of LacZ in the presence of lactose or IPTG in E. coli [4], whereas in others they rely on stochastic processes, such as competence (DNA uptake) in Bacillus [5]. Therefore, for a precise understanding of how such decisions are made and ultimately how living organisms perform, we need to characterize theoretically and experimentally the mechanisms that lead to dynamic gene expression as well as how such mechanisms can be shaped by evolution [6,7].

With the advent of experimental techniques with single-cell resolution, several mechanisms of stochastic nature have been recognized [8–12]. One of these mechanisms, which has been extensively studied in recent years, is transcriptional bursting [12–17]. In this mechanism the promoter of a gene (cis-regulatory region at the DNA level) presents multiple states in terms of transcription activity (typically simplified to two, on and off), which ends in the generation of bursts of protein production and, in turn, of cell-to-cell variability. The promoter switches stochastically from one state to another, a process that is generally attributed to the action of a transcription factor (TF) or the chromatin [17]. Broadly, this switching between functional states might also occur at other regulatory layers, in particular at the post-transcriptional level. Indeed, recent experimental results with single-molecule resolution are starting to uncover time-dependent differences in the translation rates of equal messenger RNAs (mRNAs) [18,19]. However, it is still unknown to what extent mRNA-to-mRNA variability can have an impact on the cell (or the organism) and what theoretical models are useful to describe this biochemical process.

In this work we study the dynamic expression of a gene homogenously transcribed, but with a two-state cis-regulatory region at the RNA level (the leader region). This is a quite different regulatory model with respect to previous models, as the bursting process involves ribosomes rather than polymerases, a phenomenon previously overlooked. We develop a general mathematical framework applicable in both prokaryotic and eukaryotic contexts, where the translation rate (0 in the off state and \( k_p \) in the on state) is approached by a random telegraph process \( \zeta \), characterized by the kinetic constants of activation and deactivation \( (k_{ao} \text{ and } k_{off}, \text{ respectively}) \) [12]. As a case study, we apply this framework to a bacterial gene that is controlled by a small RNA (sRNA) [20]. Here the ribosome-binding site (RBS) is trapped in a strong RNA structure in the off state and it is released upon interaction between the sRNA and mRNA in the on state [Fig. 1(a)]. Notably, through a comprehensive analysis with stochastic differential equations and molecular interaction thermodynamics, we are able to derive analytical noise expressions and provide a genotype-phenotype map.

We expect the high application of this theoretical framework, which can be extended beyond riboregulation, to rationalize dynamic gene expression data derived from experiments with single-molecule resolution of translation, once a technology based on a protein-tagging system (called SunTag) able to recruit fluorescent proteins fused to an antibody (with great
but also gene expression noise [24,25]. RNAs in not only fine-tuning gene expression levels [22,23], addition, this study also serves to point out the role of regulatory signal molecules and/or intrinsic stochastic fluctuations. In to the recognition of changes in activity due to the effect of 

\[ \langle \xi_t(t) \rangle = 0 \text{ and } \langle \xi_0(t) \rangle = \delta(t) \text{ for } i = m, p. \] 

In this work, we discard, for simplicity and without loss of generality, the effect of extrinsic noise [8]. Furthermore, the statistics for the telegraph process are \( \langle \xi(t,x,s) \rangle = \langle x \rangle p_{on} \) and \( \langle \xi(0,x,s) \xi(t,x,s) \rangle = \langle x \rangle p_{on}(1-p_{on}) \delta(t/\tau) \), with \( p_{on} = k_{on}(s)/k_{on}(s+k_{off}) \) and \( \tau = 2/(k_{on}(s)+k_{off}) \) (see Sec. 1 in the Appendix).

To investigate the phenomenon of post-transcriptional bursting, we first solve numerically the system with constant mRNA amount (see Sec. 2 in the Appendix). We use the values of \( k_{on} \) and \( k_{off} \) experimentally determined for the interaction between the sRNA SgrS and the mRNA ptsG in bacteria, leading to an asymmetric on-off switching (slow activation and fast deactivation) [23]. When there is only one mRNA in the cell (and ten sRNAs), we find a protein dynamics with bursts, resulting in irregular periods of accumulation and relaxation [Fig. 1(b)]. In this scenario, bursts from heterogeneous translation occur at a frequency of \( k_{on}(s) \) and they have a typical amplitude of \( k_p/k_{off} \), although larger bursts can be observed due to the stochasticity of the system. Because the deactivation is produced in a few seconds, the RBS has to be strong so that bursts have a significant impact. This highlights the coupling between the sRNA-mRNA and ribosome-mRNA interactions. However, when there are many mRNAs in the cell, the fluctuations are less significant with respect to the mean value \( \langle y \rangle \) [Fig. 1(c)]. Indeed, bursts occurring at the single-molecule level are hidden at the cellular level in the same way as single-cell variability is hidden when a cell population as a whole is analyzed [8].

Simulations with a full kinetic model (see Sec. 3 in the Appendix) give similar results (Fig. 2) [26]. We indeed observe bursts in protein expression when the average number of mRNAs in the cell is \( \langle x \rangle < 10 \). Nevertheless, Kolmogorov-Smirnov tests reveal significant differences between the distributions coming from the stochastic differential equations and the full kinetic model for different transcription rates [Fig. 2(c)]. This reflects the approximation of the continuous model, which is instrumental to perform an analytical treatment.

We calculate analytically the amplitude of the stochastic fluctuations in expression under the mean-field approximation \( x = \langle x \rangle + \Delta x \) and \( y = \langle y \rangle + \Delta y \) [27]. The deterministic solution in the steady state is \( x = c k_m/\gamma_m \) and \( y = k_p p_{on}(x)/\gamma_p \). Applying the Wiener-Khinchin theorem, we obtain mathematical expressions for the noise in mRNA and protein amounts (\( \eta^2_x = \langle \Delta x^2 \rangle /x^2 \) and \( \eta^2_y = \langle \Delta y^2 \rangle /y^2 \).
POST-TRANSCRIPTIONAL BURSTING IN GENES …

PHYSICAL REVIEW E 97, 032401 (2018)

FIG. 2. (a) and (b) Dynamics of protein expression with the full kinetic model. On the right, the dynamics within a time window is presented. The transcription rate is (a) $k_\text{on} = 0.15 \text{ mol/min}$ and (b) $k_\text{on} = 1 \text{ mol/min}$. In both cases, $k_\text{off} = 0.01 \text{ mol}^{-1}\text{min}^{-1}$, $k_\text{eff} = 10 \text{ min}^{-1}$, $k_p = 1.5 \text{ mol/min}$ (transcription rate of sRNA), $\gamma_m = 0.15 \text{ min}^{-1}$, $k_p = 10 \text{ min}^{-1}$, and $\gamma_p = 0.02 \text{ min}^{-1}$ (characteristic of prokaryotes). Noise decreases by augmenting $k_\text{on}$ ($10 \times$) or reducing $k_\text{eff}$ ($0.1 \times$), while noise increases by reducing $k_p$ ($0.2 \times$) or augmenting $\gamma_p$ ($6 \times$).

FIG. 3. Noise in protein expression $\eta_p^2$ calculated analytically. The red line corresponds to the nominal parameter values of $k_\text{on} = 0.01 \text{ mol}^{-1}\text{min}^{-1}$, $k_\text{eff} = 10 \text{ min}^{-1}$, $c = 1$, $k_m = 1 \text{ mol/min}$, $\gamma_m = 0.15 \text{ min}^{-1}$, $k_p = 10 \text{ min}^{-1}$, and $\gamma_p = 0.02 \text{ min}^{-1}$ (characteristic of prokaryotes). Noise decreases by augmenting $k_\text{on}$ ($10 \times$) or reducing $k_\text{eff}$ ($0.1 \times$), while noise increases by reducing $k_p$ ($0.2 \times$) or augmenting $\gamma_p$ ($6 \times$).

While noise in the mRNA expression is simply Poissonian, noise in protein expression is contributed by three different sources. The first term in the equation accounts for the bursting process at the translation level, the second term for the transmission of noise from the mRNA and sRNA (having assumed the same degradation rate for these two molecules, $\gamma_m$), and the third term for the intrinsic noise in the processes of translation and protein degradation (Fig. 3). With the parametrization used here, the three terms have similar values. The bursting process increases by 1 the Fano factor (defined as $\langle y \rangle \eta_y^2$), representing a 37.5% of the noise.

These calculations provide noise-mean scaling laws, beyond the null model of a Poisson distribution, from which to confront experimental data, once technical variability has been corrected [28]. In this regard, the precise experimental determination of the Fano factor (greater than 1) of the effective noise term $\eta_y^2 = 2 \mp \eta_p^2$ could be a useful strategy to recognize scenarios of post-transcriptional bursting (if we assume $\gamma_p \ll \gamma_m$ and $\eta_m^2 = 0$). Note that the Fano factor of $\eta_p^2$ is already greater than 1 in the case of homogenous translation, because this process amplifies the noise in the mRNA amount [29].

B. Post-transcriptional vs transcriptional bursting

With the aim of performing a comparative analysis, we construct the system of stochastic differential equations that model the dynamics in the case of transcriptional on-off switching [as Eqs. (1)]. Now the telegraph process describes the transcription rate [$k_m \rightarrow k_m \zeta(t,c,s)$] and the translation rate is homogeneous [$k_p \zeta(t,x,s) \rightarrow k_p$]. In this regard, $s$ denotes the concentration of a TF that controls the promoter activity (i.e., the ability of the promoter to recruit RNA polymerases, leading to time-dependent transcription). Notably, this model might also be exploited to study the effect of genetic redundancy by playing with the parameter $c$ [30]. We simulate the two systems by ensuring the same time scale in all biochemical processes and the same value of $\langle y \rangle$. As expected, we find larger bursts of protein production in the case of transcriptional on-off switching [Figs. 4(a)–4(d)], as there is a process of signal amplification. We also obtain analytical expressions for the noise in mRNA and protein amounts (see Sec. 4 in the Appendix), already known [12].

A genetic implementation of a system having a high transcription rate and a low translation rate is widely accepted as a strategy to minimize the noise in protein expression, as a strong promoter recruits RNA polymerases constantly without bursting and a weak RBS does not amplify.
Subsequently, we analyze how information encoded on the signal molecule, an sRNA or a TF, is transmitted through these regulatory systems (by using mutual information as a metric, as previously done [30]). This is part of a longer information transmission chain, from the environment to the phenotype. Certainly, by increasing the number of signal molecules in the cell (s), the activation becomes faster and the dynamics more Poissonian. We find that the noise in protein expression is lower in the case of post-transcriptional on-off switching when the signal is weak (p_{on} < 0.1), but higher when the signal is strong [Fig. 4(e)]. We also find that only for small input dynamic ranges (s_{max}/s_{min}) post-transcriptional regulation outperforms transcriptional regulation in terms of information transfer [Fig. 4(f)]. This suggests that sRNAs may mainly work to fine-tune protein expression by sensing weak signals, while the gross expression level may be adjusted through transcriptional control.

### C. Predictability and tunability of the protein expression distribution

According to previous work, the distribution of protein expression can well be approached by a Gamma distribution [13,33]. This form,

\[ P(y) = \frac{y^{a-1}e^{-y/b}}{b^a\Gamma(a)} \]

(3)

can still be valid to describe a system with post-transcriptional bursting, provided we redefine its two characteristic parameters (shape and scale, denoted by a and b, respectively). Here b represents the total mean number of proteins produced per stochastic fluctuation (i.e., the Fano factor). Typically, k_{on}/k_{off} proteins are produced once the RBS is active and one mRNA and one protein are produced intrinsically. The mRNA molecule generates k_{off}/k_{on} proteins before it is degraded. In total, b = k_{on}/k_{off} + k_{off}/\gamma_m + 1 proteins are produced on average per stochastic fluctuation. In addition, a represents the total mean number of stochastic fluctuations during the protein lifetime, which is the total mean number of fluctuations per mRNA in that period (n) times the mean number of mRNAs (x) (a = n(x)). With r_m = 1/(k_{on}(s)), r_m = 1/\gamma_m, \tau_p = 1/(k_{on}p_{on}), and \tau_p = 1/\gamma_p, we have n = \tau_p(r_m + \tau_m + \tau_p). Notably, in the limit p_{on} \ll 1 and \gamma_p \ll \gamma_m, the Gamma distribution as constructed is compatible with Eqs. (1) and (2), as \ E[P(y)] = ab = \langle y \rangle and \ Var[P(y)] = ab^2 = \langle \Delta y^2 \rangle.

Because the bursting regime requires a low number of molecules, we then investigate two additional post-transcriptional regulations in bacteria that contribute to this end [Fig. 5(a)]. On the one hand, the mRNA can be quickly degraded if an antisense small RNA (asRNA) binds to it [22,23], a mechanism that can also occur in eukaryotes to fine-tune protein expression and noise [24]. To model this, we modify Eqs. (1) as \gamma_m \rightarrow \gamma_m + \mu_m, assuming a high constant amount of the asRNA and no binding interference with the sRNA. As a result of the additional regulation, the distribution is shifted to the left and becomes more skewed to the right [Fig. 5(b)]. This shape comes from the slowly decaying autocorrelation of protein expression due to \gamma_p \simeq 0 and it is similar to the distribution that would be obtained in the case of moderate transcriptional bursting. On the other hand, the...
protein sequence can contain a degradation tag [34]. In a regime of a few molecules due to a weak promoter (or the presence of an asRNA), the proteolytic machinery is not saturated and then such degradation can be considered to be of first order. We modify Eqs. (1) as $y_p \rightarrow y_p + \mu_p$ to model this updated scenario (typically, $\mu_p \gg y_p$). As a result, the distribution is again shifted to the left but with a radically different, monotonically decaying shape [Fig. 5(c)]. In both scenarios, the distributions coming from numerical simulations are in agreement with those obtained analytically (see also Fig. 6).

Overall, these results highlight how post-transcriptional regulations are useful to modulate not only the magnitude but also the shape of the intrinsic variability. This can end in different dynamic protein expression programs of stochastic nature that can be switched by regulatory RNAs according to particular environmental signals [25].

**D. Nonequilibrium thermodynamics of the sRNA-mRNA interaction**

Finally, to further understand the process of post-transcriptional bursting, we inspect the dependence of $k_{on}$ and $k_{off}$ on the energetics of the sRNA-mRNA interaction. This characterizes the nonequilibrium thermodynamics of the system, which, however, satisfies detailed balance $[p_{on} k_{off} = (1 - p_{on}) k_{on}(y)]$. Typically, the energy landscape is described by the free energy of activation ($\Delta G^\ddagger > 0$), the free energy required to get the two molecules together and the free energy of formation ($\Delta G < 0$, the net free-energy release upon interaction; see Fig. 7(a)). According to the transition-state theory [35], a kinetic constant is defined by the free-energy barrier faced with the reaction progress. Here the transition state corresponds to the sRNA-mRNA complex in which only the seed regions of each molecule are annealed [20]. In the classical scenario of equilibrium, the ratio between $k_{on}$ and $k_{off}$ yields the equilibrium constant $(K_{eq} = k_{on}/k_{off} \propto e^{-\beta \Delta G})$, which determines the protein production and only depends on the free energy of formation (not of activation). However, extensive experimental data have shown that only $\Delta G$ is not sufficient to explain the dynamic range in a scenario of riboregulation [36]. Instead, $\Delta G^\ddagger$ and $\Delta G$ are the two main determinants. Even if formation is very favorable, the reaction does not progress in the absence of a suitable seed interaction [37].

We hypothesize a nonequilibrium scenario in which the system continuously fluctuates between the two states and the dependence of the average protein production on the energetics is different [38]. On the one hand, the off to transition faces a free-energy barrier of $\Delta G^\ddagger$. Detailed experiments with nucleic acids evidence an exponential acceleration of $k_{on}$ with the seed region length, as well as a saturation for a sufficiently large seed [39]. In addition, the larger the seed region, the lower the free energy associated with the seed region annealing ($\Delta G^\ddagger < 0$). Hence, prior to the saturation limit, we can state that $\Delta G^\ddagger$ scales with $\Delta G^\ddagger$, leading to $e^{-\beta \Delta G^\ddagger} \propto e^{-\beta \Delta G^\ddagger}$. After that, the formation of the final intermolecular complex takes place, with low probability and slow kinetics due to the formation of intermolecular base pairs at the cost of breaking intramolecular ones [37]. Note that the seed region annealing contributes in part to the total complex formation ($\Delta G^\ddagger > \Delta G$). On the other hand, the on to off transition faces a free-energy barrier that depends on $\Delta G$, due to the process of recomposition of intramolecular base pairs by breaking intermolecular ones. This is assumed to not depend on $\Delta G^\ddagger$, as once the previous process has ended, most probably through different pathways, dissociation is immediate. Therefore, our coarse-grained energy model reads

$$k_{on} = k_{on}^0 e^{-\beta \Delta G^\ddagger}, \quad k_{off} = k_{off}^0 e^{-\beta \Delta G^\ddagger},$$

(4)

where $k_{on}^0$ and $k_{off}^0$ represent the asymptotic kinetic constants. The parameter $\beta$ is empirically estimated to be a value lower than the theoretical value given by the Boltzmann constant [40,41]. Importantly, both $\Delta G^\ddagger$ and $\Delta G$ can be directly calculated from the nucleotide sequences [42]. Figures 7(b) and 7(c) illustrate how the dynamics changes with such energetics.

This results in a computational tool $[(y) \propto k_{on}/k_{off} \propto e^{-\beta (\Delta G^\ddagger + \Delta G)}]$ to analyze the dynamics of natural and mutant systems [36] or to guide the design of synthetic sequences to implement different functionalities in the cell [20]. Of relevance is that the experimental characterization of
FIG. 6. Dynamics and distribution of protein expression in the case of (a) and (f) $k_{on} = 0.1 \text{ min}^{-1}$ and $k_{off} = 10 \text{ min}^{-1}$, (b) and (g) $k_{on} = 0.03 \text{ min}^{-1}$ and $k_{off} = 3 \text{ min}^{-1}$, (c) and (h) $k_{on} = 0.01 \text{ min}^{-1}$ and $k_{off} = 1 \text{ min}^{-1}$, (d) and (i) $k_{on} = 1 \text{ min}^{-1}$ and $k_{off} = 10 \text{ min}^{-1}$, and (e) and (j) $k_{on} = 0.1 \text{ min}^{-1}$ and $k_{off} = 1 \text{ min}^{-1}$ for (a)–(e) stable protein, $\gamma_p = 0.02 \text{ min}^{-1}$, and (f)–(j) unstable protein, $\gamma_p = 0.12 \text{ min}^{-1}$. In all cases, $k_m = 1 \text{ mol/min}$, $\gamma_m = 0.15 \text{ min}^{-1}$, and $k_p = 10 \text{ min}^{-1}$. This leads to (a)–(c) $p_{on} = 0.01$ and $\langle y \rangle = 33.33 \text{ mol}$, (d) and (e) $p_{on} = 0.1$ and $\langle y \rangle = 333.33 \text{ mol}$, (f)–(h) $p_{on} = 0.01$ and $\langle y \rangle = 5.56 \text{ mol}$, and (i) and (j) $p_{on} = 0.1$ and $\langle y \rangle = 55.56 \text{ mol}$. The analytical distribution for $y$ (Gamma) is shown in red.

III. CONCLUSION

We have presented a theoretical framework to study a distinctive mechanism at the post-transcriptional level for the generation of stochastic bursts in protein expression. This
In theory, post-transcriptional bursting arises when there are at least two clearly different translation modes (on and off). Moreover, the switching kinetics between the modes have to be highly asymmetric towards the off state ($k_{\text{on}} \ll k_{\text{off}}$) and the mean time in the on state sufficient to allow translation initiation ($k_{\text{p}} \gtrsim k_{\text{off}}$). In prokaryotes, a scenario of positive riboregulation can favor this [20], but not a scenario of negative riboregulation (see Sec. 5 in the Appendix), not in the case of repression of an active RBS [36] or degradation of the mRNA [22,23]. In both latter cases, translation occurs in the absence of the asRNA, which results in a kinetic asymmetry towards the active translation state (Fig. 8). In eukaryotes, by contrast, the translation initiation machinery is more complex, involving the participation of several cofactors, which can result in spontaneous on-off switchings with slower kinetics (e.g., periods in the on state of some minutes) [18]. Consequently, bursts can be much larger.

All in all, our results highlight a fundamental contribution of the post-transcriptional control mechanisms to the construction of dynamic gene expression programs, stressing the utility of the RNA layer for the cell [45]. More broadly, they highlight how the nonequilibrium thermodynamics of the functional macromolecular complexes in the cell [46], in particular those involved in translation (here sRNA-mRNA), can lead to different phenotypes with time. This quantitative and testable description of translation initiation dynamics and protein expression variability, explicitly considering time and energy, is parallel to the development of single-molecule experimental approaches able to monitor the activity of those complexes in vivo [18,19], a combination that ultimately leads to an understanding of biology from first principles.

**ACKNOWLEDGMENT**

This work was supported by the Spanish Ministry of Economy and Competitiveness under Grant No. BFU2015-66894-P (MINECO/FEDER).

**APPENDIX**

1. **Simplification of the kinetics for the telegraph process**

Because the on-off switching is very fast in comparison to other processes in the system ($k_{\text{off}} \gg \gamma_m \gg \gamma_p$), we approached a decay exponential by a Dirac delta $e^{-t/\tau} \simeq \delta(t/\tau)$. Here $1/\tau$ represents the mean switching rate. This simplified the analytical calculations to obtain Eqs. (2). Moreover, we used the first-order expansion $\zeta(t,x,s) \simeq \zeta(t,\langle x \rangle,\langle s \rangle) + \frac{\gamma_m}{k_{\text{on}}} (s) \Delta x + \langle x \rangle \Delta s$.

2. **Numerical simulations of the stochastic differential equations**

For numerical calculations of Eqs. (1), we followed $\int_0^{\Delta t} \xi_i(t) dt = z \mathcal{N}(0,1)$, and $\int_0^{\Delta t} \zeta(t,x,s) dt$ for $b = 1/k_{\text{off}}$, where $b$ is a random number binomially distributed, $b \sim \mathcal{B}(x,p_{\text{on}})$. This way, $x(t)$ and $y(t)$ were obtained.
3. Full kinetic model

The set of biochemical reactions that describe a system in which the sRNA activates the RBS of the mRNA is

\[
\begin{align*}
\emptyset + x & \xrightarrow{k_p} x^*, \\
x & \xrightarrow{\gamma_m} \emptyset, \\
\emptyset + y & \xrightarrow{k_{ps}} s, \\
s & \xrightarrow{\gamma_m} \emptyset, \\
x^* + s & \xrightarrow{k_{ps}} x^*, \\
 x + s & \xrightarrow{k_{m}} x + s, \\
x^* & \xrightarrow{k_{d}} x^* + y, \\
y & \xrightarrow{\gamma_m} \emptyset,
\end{align*}
\]

where \(x\) denotes the number of free mRNAs, \(s\) the number of free sRNAs, \(x^*\) the number of sRNA-mRNA complexes, and \(y\) the number of proteins. Note that we have assumed the same degradation rate for all RNA species. By solving this kinetic model with the Gillespie algorithm [26], we obtained the dynamics of protein expression.

4. Additional analytical results regarding bursting

In the case of transcriptional bursting, the noise is \(\eta_1^2 = \frac{k_p(1-p_p)}{k_p(t)} + \frac{1}{2} \gamma_m^2 + \frac{1}{2} \gamma_p^2\) and \(\eta_2^2 = \frac{\gamma_p}{\gamma_m} \gamma_m^2 + \frac{1}{2} \gamma_p^2\), in opposition to Eqs. (2). A Gamma distribution (3) explains well protein expression in different scenarios, with an appropriate definition of \(b\) and \(n\). In the case of homogenous transcription and translation, \(b = k_p/\gamma_m + 1\) and \(n = (1/\gamma_p)/(1/\gamma_m + 1/k_p)\).

In the case of transcriptional bursting, \(b = k_m k_p/(\gamma_m \gamma_{off}) + k_p/\gamma_m + 1\) and \(n = (1/\gamma_p)/(k_m/(\gamma_m \gamma_{off}) + 1/\gamma_m + 1/k_p)\).

5. Model of negative riboregulation

The stochastic differential equations that model a system in which the sRNA represses the RBS of the mRNA, without inducing degradation, are

\[
\begin{align*}
\frac{dx}{dt} &= c_{km} - \gamma_m x + \sqrt{ck_{km}} + \gamma_m x \xi_m(t), \\
\frac{dy}{dt} &= k_p[x - \xi(t,x,s)] - \gamma_p y \\
&\quad + \sqrt{k_p[x - \xi(t,x,s)] + \gamma_p y} \xi_p(t),
\end{align*}
\]

In this case, the active translation state corresponds to the mRNA alone (off, with respect to the sRNA binding) and the inactive state to the sRNA-mRNA complex (on, with respect to the sRNA binding).

[1] U. Alon, An Introduction to Systems Biology: Design Principles of Biological Circuits (Chapman & Hall/CRC, Boca Raton, 2006).

[2] A. Eldar and M. B. Elowitz, Functional roles for noise in genetic circuits, Nature (London) 467, 167 (2010).

[3] L. S. Tsimring, Noise in biology, Rep. Prog. Phys. 77, 26601 (2014).

[4] T. Kuhlman, Z. Zhang, M. H. Saier, Jr., and T. Hwa, Combinatorial transcriptional control of the lactose operon of Escherichia coli, Proc. Natl. Acad. Sci. USA 104, 6043 (2007).

[5] H. Maamar, A. Raj, and D. Dubnau, Noise in gene expression determines cell fate in Bacillus subtilis, Science 317, 526 (2007).

[6] E. Dekel and U. Alon, Optimality and evolutionary tuning of the expression level of a protein, Nature (London) 436, 588 (2005).

[7] C. Gonzalez, J. C. J. Ray, M. Manhart, R. M. Adams, D. Nevozhay, A. V. Morozov, and G. Balazsi, Stress-response balance drives the evolution of a network module and its host genome, Mol. Syst. Biol. 11, 827 (2015).

[8] M. B. Elowitz, A. J. Levine, E. D. Sigga, and P. S. Swain, Stochastic gene expression in a single cell, Science 297, 1183 (2002).

[9] A. Becskei, B. B. Kaufmann, and A. van Oudenaarden, Contributions of low molecule number and chromosomal positioning to stochastic gene expression, Nat. Genet. 37, 937 (2005).

[10] D. Huh and J. Paulsson, Non-genetic heterogeneity from stochastic partitioning at cell division, Nat. Genet. 43, 95 (2011).

[11] J. C. Locke, J. W. Young, M. Fontes, M. J. H. Jimenez, and M. B. Elowitz, Stochastic pulse regulation in bacterial stress response, Science 334, 366 (2011).

[12] A. Raj, C. S. Peskin, D. Tranchina, D. Y. Vargas, and S. Tyagi, Stochastic mRNA synthesis in mammalian cells, PLoS Biol. 4, e309 (2006).

[13] L. Cai, N. Friedman, and X. S. Xie, Stochastic protein expression in individual cells at the single molecule level, Nature (London) 440, 358 (2006).

[14] W. J. Blake, G. Balazsi, M. A. Kohanski, F. J. Isaacs, K. F. Murphy, Y. Kuang, C. R. Cantor, D. R. Walt, and J. J. Collins, Phenotypic consequences of promoter-mediated transcriptional noise, Mol. Cell 24, 853 (2006).

[15] J. M. Pedraza and J. Paulsson, Effects of molecular memory and bursting on fluctuations in gene expression, Science 319, 339 (2008).

[16] D. M. Suter, N. Molina, D. Gatfield, K. Schneider, U. Schibler, and F. Naef, Mammalian genes are transcribed with widely different bursting kinetics, Science 332, 472 (2011).

[17] R. D. Dar, B. S. Razooky, A. Singh, T. V. Trimeloni, J. M. McCollum, C. D. Cox, M. L. Simpson, and L. S. Weinberger, Transcriptional burst frequency and burst size are equally modulated across the human genome, Proc. Natl. Acad. Sci. USA 109, 17454 (2012).

[18] B. Wu, C. Eliscovich, Y. J. Yoon, and R. H. Singer, Translation dynamics of single mRNAs in live cells and neurons, Science 352, 1430 (2016).

[19] X. Yan, T. A. Hoek, R. D. Vale, and M. E. Tanenbaum, Dynamics of translation of single mRNA molecules in vivo, Cell 165, 976 (2016).

[20] G. Rodrigo, T. E. Landrain, and A. Jaramillo, De novo automated design of small RNA circuits for engineering synthetic riboregulation in living cells, Proc. Natl. Acad. Sci. USA 109, 15271 (2012).
[21] M. E. Tanenbaum, L. A. Gilbert, L. S. Qi, J. S. Weissman, and R. D. Vale, A protein-tagging system for signal amplification in gene expression and fluorescence imaging, Cell 159, 635 (2014).
[22] E. Levine, Z. Zhang, T. Kuhlman, and T. Hwa, Quantitative characteristics of gene regulation by small RNA, PLoS Biol. 5, e229 (2007).
[23] J. Fei, D. Singh, Q. Zhang, S. Park, D. Balasubramanian, I. Golding, C. K. Vanderpool, and T. Ha, Determination of in vivo target search kinetics of regulatory noncoding RNA, Science 347, 1371 (2015).
[24] J. M. Schmiedel, S. L. Klemm, Y. Zheng, A. Sahay, N. Blüthgen, D. S. Marks, and A. van Oudenaarden, MicroRNA control of protein expression noise, Science 348, 128 (2015).
[25] R. Arbel-Goren, A. Tal, B. Parasar, A. Dym, N. Costantino, J. Munoz-Garcia, and J. Stavans, Transcript degradation and noise of small RNA-controlled genes in a switch activated network in Escherichia coli, Nucleic Acids Res. 44, 6707 (2016).
[26] D. T. Gillespie, Stochastic simulation of chemical kinetics, Annu. Rev. Phys. Chem. 58, 35 (2007).
[27] T. D. Frank, A. Daffertshofer, and P. J. Beek, Multivariate Ornstein-Uhlenbeck processes with mean-field dependent coefficients: Application to postural sway, Phys. Rev. E 63, 011905 (2000).
[28] D. Grun, L. Kester, and A. van Oudenaarden, Validation of noise models for single-cell transcriptomics, Nat. Methods 11, 637 (2014).
[29] E. M. Ozbudak, M. Thattai, I. Kurtser, A. D. Grossman, and A. van Oudenaarden, Regulation of noise in the expression of a single gene, Nat. Genet. 31, 69 (2002).
[30] G. Rodrigo and J. F. Poyatos, Genetic redundancies enhance information transfer in noisy regulatory circuits, PLoS Comput. Biol. 12, e1005156 (2016).
[31] D. Nevozhay, R. M. Adams, K. F. Murphy, K. Josic, and G. Balazsi, Negative autoregulation linearizes the dose-response and suppresses the heterogeneity of gene expression, Proc. Natl. Acad. Sci. USA 106, 5123 (2009).
[32] A. Singh and J. P. Hespanha, Optimal feedback strength for noise suppression in autoregulatory gene networks, Biophys. J. 96, 4013 (2009).
[33] Y. Taniguchi, P. J. Choi, G.-W. Li, H. Chen, M. Babu, J. Hearn, A. Emili, and X. S. Xie, Quantifying E. coli proteome and transcriptome with single-molecule sensitivity in single cells, Science 329, 533 (2010).
[34] K. E. McGinness, T. A. Baker, and R. T. Sauer, Engineering controllable protein degradation, Mol. Cell 22, 701 (2006).
[35] P. Hanggi, P. Talkner, and M. Borkovec, Reaction-rate theory: Fifty years after Kramers, Rev. Mod. Phys. 62, 251 (1990).
[36] V. K. Mutalik, L. Qi, J. C. Guimaraes, J. B. Lucks, and A. P. Arkin, Rationally designed families of orthogonal RNA regulators of translation, Nat. Chem. Biol. 8, 447 (2012).
[37] N. Salim, R. Lamichhane, R. Zhao, T. Banerjee, J. Philip, D. Rueda, and A. L. Feig, Thermodynamic and kinetic analysis of an RNA kissing interaction and its resolution into an extended duplex, Biophys. J. 102, 1097 (2012).
[38] H. Feng, K. Zhang, and J. Wang, Non-equilibrium transition state rate theory, Chem. Sci. 5, 3761 (2014).
[39] N. Srinivas, T. E. Ouldridge, P. Sulc, J. M. Schaeffer, B. Yurke, A. A. Louis, J. P. K. Doye, and E. Winfree, On the biophysics and kinetics of toehold-mediated DNA strand displacement, Nucleic Acids Res. 41, 10641 (2013).
[40] Y. Hao, Z. J. Zhang, D. W. Erickson, M. Huang, Y. Huang, J. Li, T. Hwa, and H. Shi, Quantifying the sequence-function relation in gene silencing by bacterial small RNAs, Proc. Natl. Acad. Sci. USA 108, 12473 (2011).
[41] G. Rodrigo, E. Majer, S. Prakash, J. A. Daros, A. Jaramillo, and J. F. Poyatos, Exploring the dynamics and mutational landscape of riboregulation with a minimal synthetic circuit in living cells, Biophys. J. 109, 1070 (2015).
[42] D. H. Mathews, M. D. Disney, J. L. Childs, S. J. Schroeder, M. Zuker, and D. H. Turner, Incorporating chemical modification constraints into a dynamic programming algorithm for prediction of RNA secondary structure, Proc. Natl. Acad. Sci. USA 101, 7287 (2004).
[43] N. Kumar, T. Platini, and R. V. Kulkarni, Exact Distributions for Stochastic Gene Expression Models with Bursting and Feedback, Phys. Rev. Lett. 113, 268105 (2014).
[44] H. Ge, H. Qian, and X. S. Xie, Stochastic Phenotype Transition of a Single Cell in an Intermediate Region of Gene State Switching, Phys. Rev. Lett. 114, 078101 (2015).
[45] J. Chappell, M. K. Takahashi, S. Meyer, D. Loughrey, K. E. Watters, and J. B. Lucks, The centrality of RNA for engineering gene expression, Biotechnol. J. 8, 1379 (2013).
[46] A. Coulon, C. C. Chow, R. H. Singer, and R. D. Larson, Eukaryotic transcriptional dynamics: From single molecules to cell populations, Nat. Rev. Genet. 14, 572 (2013).