

Diffusion of Transcription Factors Can Drastically Enhance the Noise in Gene Expression

Jeroen S. van Zon,* Marco J. Morelli,[†] Sorin Tănase-Nicola,[†] and Pieter Rein ten Wolde[†]

*Division of Physics and Astronomy, Vrije Universiteit, Amsterdam, The Netherlands; and [†]FOM Institute for Atomic and Molecular Physics, Amsterdam, The Netherlands

ABSTRACT We study by Green's Function Reaction Dynamics the effect of the diffusive motion of repressor molecules on the noise in mRNA and protein levels for a gene that is under the control of a repressor. We find that spatial fluctuations due to diffusion can drastically enhance the noise in gene expression. After dissociation from the operator, a repressor can rapidly rebind to the DNA. Our results show that the rebinding trajectories are so short that, on this timescale, the RNA polymerase (RNAP) cannot effectively compete with the repressor for binding to the promoter. As a result, a dissociated repressor molecule will on average rebind many times, before it eventually diffuses away. These rebindings thus lower the effective dissociation rate, and this increases the noise in gene expression. Another consequence of the timescale separation between repressor rebinding and RNAP association is that the effect of spatial fluctuations can be described by a well-stirred, zero-dimensional model by renormalizing the reaction rates for repressor-DNA (un) binding. Our results thus support the use of well-stirred, zero-dimensional models for describing noise in gene expression. We also show that for a fixed repressor strength, the noise due to diffusion can be minimized by increasing the number of repressors or by decreasing the rate of the open complex formation. Lastly, our results emphasize that power spectra are a highly useful tool for studying the propagation of noise through the different stages of gene expression.

INTRODUCTION

Cells process information from the outside and regulate their internal state by means of proteins and DNA that chemically and physically interact with one another. These biochemical networks are often highly stochastic because in living cells the reactants often occur in small numbers (1,2). This is particularly important in gene expression (3–10), where transcription factors are frequently present in copy numbers as low as tens of molecules per cell. Although it is generally believed that biochemical noise can be detrimental to cell function (8), it is increasingly becoming recognized that noise can also be beneficial to the organism (9). Understanding noise in gene expression is thus important for understanding cell function, and this observation has recently stimulated much theoretical and experimental work in this direction (8–10). However, the theoretical analyses usually employ the zero-dimensional chemical master equation (10–12). This approach takes into account the discrete character of the reactants and the probabilistic nature of chemical reactions. It does assume, however, that the cell is a “well-stirred” reactor, in which the particles are uniformly distributed in space at all times; the reaction rates only depend upon the global concentrations of the reactants and not upon the spatial positions of the reactant molecules. Yet, to react, reactants first have to move toward one another. They do so by diffusion, or, in the case of eukaryotes, by a combination of diffusion and active transport. Both processes are

stochastic in nature and this could contribute to the noise in the network. Here, we study by computer simulation the expression of a single gene that is under the control of a repressor R in a spatially resolved model. We find that at low repressor concentration, i.e., $[R] < 50$ nM, the noise in gene expression is dominated by the noise arising from the diffusive motion of the repressor molecules. Our results thus show that spatial fluctuations of the reactants can be an important source of noise in biochemical networks. Our analysis also reveals that the effects of diffusion can nevertheless be described by a well-stirred model, provided that the reaction rates of repressor-DNA (un)binding are properly renormalized.

The simulations show that in gene expression significant fluctuations occur on both short and long length- and timescales. As expected from earlier work (13–15), the fluctuations on long timescales are predominantly due to protein degradation; we assume that proteins are degraded by dilution, which means that the relaxation rate of this process is on the order of 1 h. Our results, however, also elucidate an important process on much shorter length- and timescales. It is associated with the competition between repressor and RNA polymerase (RNAP) for binding to the promoter. When a repressor molecule dissociates from the DNA, it can rebind very rapidly: in our model, which neglects one-dimensional diffusion along the DNA, it can rebind on a timescale of milliseconds, or less. This timescale is much shorter than that on which the RNAP binds to the promoter, which is on the order of 0.01–0.1 s. Hence, when a repressor molecule has just dissociated, the probability that a RNAP will bind before the repressor molecule rebinds, is very

Submitted March 29, 2006, and accepted for publication September 6, 2006.

Address reprint requests to Pieter Rein ten Wolde, E-mail: tenwolde@amolf.nl.

© 2006 by the Biophysical Society

0006-3495/06/12/4350/18 \$2.00

doi: 10.1529/biophysj.106.086157

small. This has two important consequences. The first is that a repressor molecule will on average rebind many times before it eventually diffuses away from the promoter and a RNAP molecule, or another repressor molecule, can bind to the promoter. This decreases the effective dissociation rate, which increases the noise in gene expression.

The second consequence of the rapidity of the rebindings is that noise propagation during gene expression can be described by a well-stirred, zero-dimensional model. In the commonly used zero-dimensional models, chemical reactions are exponentially distributed in time. In a spatially resolved model, the distribution of repressor-DNA association times deviates markedly from Poisson statistics. Although at long timescales the distribution is exponential, at short timescales it is algebraic, due to the diffusive nature of the rebinding trajectories. However, these repressor rebindings are so fast that they do not significantly affect the dynamics of RNAP-DNA association; the latter is only affected by the repressor-DNA (un)binding dynamics at longer timescales, which obey Poisson statistics. The reason that the effect of spatial fluctuations on noise in gene expression can be described by a zero-dimensional model is thus a separation of timescales. In fact, it is conceivable that in a more realistic model of gene expression, which includes one-dimensional sliding along the DNA, the timescale of repressor rebinding is not separated from that of the RNAP dynamics. Under these conditions, the effect of spatial fluctuations might be detected in the statistics of mRNA production. However, as we discuss in the Discussion and Outlook section, the noise strength (variance) of the mRNA level can probably still be described by a zero-dimensional model, because the timescale of the spatial fluctuations, even in those more refined models, is expected to be still shorter than the typical life time of an mRNA molecule.

Because fluctuations in gene expression span orders of magnitude in length- and timescales, the simulation technique should be sufficiently detailed to resolve the events at short length- and timescales, yet also efficient enough to access the long length- and timescales. Recently, several simulation techniques have been developed for the stochastic modeling of reaction-diffusion systems (16,17). These techniques, however, do not satisfy both criteria: they either describe the system in a coarse-grained way, i.e., on the level of local concentrations rather than single particles (16,17) or are too slow to accurately model the dynamics on the long timescales (18). Our simulations have been made possible via the use of our recently developed Green's function reaction dynamics (GFRD) algorithm (19,20). GFRD is an event-driven algorithm that uses Green's functions to combine in one step the propagation of the particles in space with the reactions between them. The event-driven nature of the algorithm makes it particularly useful for problems, such as gene expression, in which the events are distributed over a wide range of length- and timescales: the algorithm takes small steps when the reactants are close to each other—such

as when a repressor molecule has just dissociated from the DNA—whereas it takes large jumps in time and space when the molecules are far apart from each other, like when the repressor molecule has eventually diffused away from the promoter. The event-driven nature of GFRD makes it orders of magnitude more efficient than brute-force particle-based algorithms (20) and this has allowed us to simulate gene expression on the relevant biological timescales of hours.

Several publications (21–29) have discussed the effect of fluctuations in the binding of transcription factors to their site on the DNA (called operator) on the noise in gene expression. Most of these models are relatively simple, ignoring, for instance, production of mRNA (23–26,29). Moreover all these studies, with the exception of (24,28), ignore the role of the spatial fluctuations of the transcription factors. Our aim is to study gene expression in a biologically meaningful model. We have therefore constructed a rather detailed model, although we will also use minimal models that can be studied analytically to interpret the simulation results. The full model, which is described in the next section, contains the diffusive motion of repressor molecules, open complex formation, promoter clearance, transcription elongation, and translation (30).

In the section “Simulation results: dynamics and noise”, we discuss the simulation results for both the noise in mRNA and protein level. The results reveal that for $[R] < 50$ nM, the noise in the spatially resolved model can be more than five times larger than the noise in the well-stirred model. We also show that a cell could minimize the effect of spatial fluctuations, either by tuning the open complex formation rate or by changing the number of repressors and their affinity for the binding site on the DNA. In the section “Simulations results: operator binding”, we elucidate the origin of the enhanced noise in the spatially resolved model. In the subsequent section, we show that in the model employed here the effect of spatial fluctuations can be quantitatively described by a well-stirred model in which the reaction rates for repressor binding and unbinding are appropriately renormalized; however, as alluded to above, and as we will discuss in more detail in the last section, we expect that in a more refined model the effect of diffusion will be more complex, impeding such a simplified description. In the section “Power spectra”, we discuss how the operator state fluctuations propagate through the different stages of gene expression using power spectra for the operator state, elongation complex, mRNA, and protein. The results show that these power spectra are highly useful for unraveling the dynamics of gene expression. We hope that this stimulates experimentalists to measure power spectra of not only mRNA and protein levels (31), but also of the dynamics of transcription initiation and elongation using, e.g., magnetic tweezers (32). As we argue in the last section, such experiments should make it possible to determine the importance of spatial fluctuations for the dynamics of gene expression.

MODEL

Diffusive motion of repressors

We explicitly simulate the diffusive motion of the repressor molecules in space. However, since the experiments of Riggs et al. (33) and the theoretical work of Berg et al. (34), it is well known that proteins could find their target sites via a combination of one-dimensional (1D) sliding along the DNA and three-dimensional (3D) diffusion through the cytoplasm—“hopping” or “jumping” from one site on the DNA to another. This mechanism could speed up the search process and make it faster than the rate at which particles find their target by free 3D diffusion; this rate is given by $k = 4\pi\sigma D_3[R]$, where σ is the cross section, which is on the order of a protein diameter or DNA diameter, D_3 is the diffusion constant of the protein in the cytoplasm, and $[R]$ is the concentration of the (repressor) protein. However, although it is clear that the mechanism of 3D diffusion and 1D sliding could potentially speed up the search process, whether this mechanism in living cells indeed drastically reduces the search time is still under debate (35). In this context, it is instructive to discuss the two main results of recent studies on this topic (35–40). The first is that the mean search time τ is given by (40)

$$\tau \sim \frac{L}{\lambda} \left[\frac{\lambda^2}{D_1} + \frac{r^2}{D_3} \right], \quad (1)$$

where L is the total length of the DNA, λ is the average distance over which the protein slides along the DNA before it dissociates, D_1 is the diffusion constant for sliding, r is the typical mesh size in the nucleoid (the characteristic distance between two segments on the DNA (40)), and D_3 is the diffusion constant in the cytoplasm. This formula has a clear interpretation (40): λ^2/D_1 is the sliding time, r^2/D_3 is the time spent on 3D diffusion, the sum of these terms is thus the time to perform one round of sliding and diffusion, and L/λ is the total number of rounds needed to find the target. The other principal result is that the search time is minimized when the sliding distance λ is

$$\lambda = \sqrt{\frac{D_1}{D_3}} r. \quad (2)$$

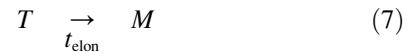
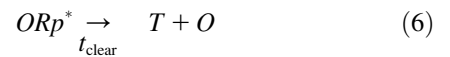
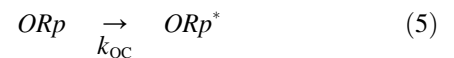
Under these conditions, a protein spends equal amounts of time on 3D diffusion and 1D sliding (a protein is thus half the time bound to the DNA). Equation 2 is a useful result because it shows that the average sliding distance λ depends upon the ratio of diffusion constants and on the typical mesh size in the nucleoid. If we now assume that D_1 and D_3 are equal (which is not obvious given that proteins bind relatively strongly to DNA— D_1 could thus very well be much smaller than D_3) and if we take the mesh size to be given by $r \sim \sqrt{v/L}$ (40), where $v \approx 1 \mu\text{m}^3$ is the volume of an *Escherichia coli* cell and $L \approx 10^3 \mu\text{m}$, we find that $\lambda \approx 10 \text{ nm}$ (30 bp). This corresponds to the typical diameter of a

protein or DNA double helix and is thus not very large. Interestingly, recent experiments seem to confirm this: experiments from Halford et al. on restriction enzymes (*EcoRV* and *BbcCI*) with a series of DNA substrates with two target sites and varying lengths of DNA between the two sites, suggest that under the in vivo conditions, sliding is indeed limited to relatively short distances, i.e., to distances less than 50 bp ($\approx 16 \text{ nm}$) (41,42).

Now, it should be realized that on lengthscales beyond the sliding length, the motion is essentially 3D diffusion: the sliding/hopping mechanism corresponds to 3D diffusion with a jump distance given by the sliding distance (36). Moreover, since the sliding distance is only on the order of a particle diameter, as discussed above, we have therefore decided to model the motion of the repressor molecules as 3D diffusion. But it should be remembered that on lengthscales shorter than 10–30 nm, this approach is not correct. As we discuss in the “Discussion and Outlook” section, this might have significant implications for the importance of spatial fluctuations for the noise in gene expression.

Transcription and translation

Most repressors bind to a site that (partially) overlaps with the core promoter—the binding site of the RNA polymerase (RNAP). When a repressor molecule is bound to its operator site, it prevents RNAP from binding to the promoter, thereby switching off gene expression. Only in the absence of a repressor on the operator site can RNAP bind to the promoter and initiate transcription and translation, ultimately resulting in the production of a protein. We model this by the following reaction network:



Equations 3 and 4 describe the competition between the binding of the repressor R and the RNAP molecules Rp to the promoter (O is the operator site). In our simulation we fix

the binding site O in the center of a container with volume $V = 1 \mu\text{m}^3$, comparable to the volume of a single *E. coli* cell. We simulate both the operator site O and the repressor molecules as spherical particles with diameter $\sigma = 10 \text{ nm}$. The operator site O is surrounded by N_R repressor molecules that move by free 3D diffusion (see previous section) with an effective diffusion constant $D = 1 \mu\text{m}^2\text{s}^{-1}$, as has been reported for proteins of a similar size (43). The intrinsic forward rate $k_{\text{IR}} = 6 \times 10^9 \text{M}^{-1}\text{s}^{-1}$ for the repressor particles R at contact is estimated from the Maxwell-Boltzmann distribution (19). The backward rate k_{bR} depends on the interaction between the DNA binding site of the repressor and the operator site on the DNA and varies greatly between different operons, with stronger repressors having a lower k_{bR} . In our simulations, we vary k_{bR} between 1 and 0.01 s^{-1} , as discussed in more detail below. The concentration of RNAP is much higher than that of the repressor (44). Because of this, we treat the RNAP as distributed homogeneously within the cell and we do not take diffusion of RNAP into account explicitly. Instead, RNAP associates with the promoter with a diffusion-limited rate $k_{\text{IRp}} = 4\pi\sigma D[Rp]$. In our simulations, the concentration of free RNAP is $[Rp] = 0.5 \mu\text{M}$ (44), leading to a forward rate $k_{\text{IRp}} = 38 \text{ s}^{-1}$. Finally, the backward rate $k_{\text{bRp}} = 0.5$ is determined such that $K_{\text{eq}} = 4\pi\sigma D/k_{\text{bRp}} = 1.4 \times 10^9 \text{M}^{-1}$ (45).

Transcription initiation is described by Eqs. 5 and 6. Before productive synthesis of RNA occurs, first the RNAP in the RNAP-promoter complex ORp unwinds approximately one turn of the promoter DNA to form the open complex ORp^* . The open complex formation rate k_{OC} has been measured to be on the order of $0.3\text{--}3 \text{ s}^{-1}$ (32). We approximate open complex formation as an irreversible reaction. Some experiments find this step to be weakly reversible (32). However, adding a backward reaction to the model did not change the dynamics of the system in a qualitative way, as long as the backward rate is smaller than k_{OC} , which is in agreement with experimental results. After open complex formation, RNAP must first escape the promoter region before another RNAP or repressor can bind. Because elongation occurs at a rate of $50\text{--}100$ nucleotides per second and between 30 and 60 nucleotides must be cleared by RNAP before the promoter is accessible, a waiting time of $t_{\text{clear}} = 1 \text{ s}$ is required before another binding can occur. Since promoter clearance consists of many individual elongation events that obey Poisson statistics individually, we model the step as one with a fixed time delay t_{clear} , not as a Poisson process with rate $1/t_{\text{clear}}$.

Equations 7–11 describe the dynamics of mRNA and protein numbers. After clearing the promoter region, RNAP starts elongation of the transcript T . As for clearance, the elongation step is modeled as a process with a fixed time delay $t_{\text{elon}} = 30 \text{ s}$, corresponding to an elongation rate of $50\text{--}100$ nucleotides per second and a 1500 bp gene. When an mRNA M is formed, it can degrade with a rate k_{dm} . Here, the mRNA degradation rate is determined by fixing the average

mRNA concentration in the unrepresed state, as described below. Furthermore, an mRNA molecule can form an mRNA-ribosome complex M_{ribo} and start translation. We assume that $b = 5$ proteins are produced on average from a single mRNA molecule (7), so that the start of translation occurs at a rate $k_{\text{ribo}} = bk_{\text{dm}}$. After a fixed time delay, $t_{\text{trans}} = 30 \text{ s}$, a protein P is produced. The mRNA is available for ribosome binding immediately after the start of translation. Due to the delay in protein production, M can start to be degraded, while the mRNA-ribosome complex M_{ribo} is still present; M thus represents the mRNA leader region rather than the entire mRNA molecule. Finally, the protein P degrades at a rate k_{dp} , which is determined by the requirement that the average protein concentration in the unrepresed state has a desired value, as we describe now.

We vary the free parameters in the reaction network described in Eqs. 3–11, N_R , k_{bR} , k_{dm} , k_{dp} , in the following way: first, we choose the concentration of mRNA and protein in the absence of repressor molecules. In this case, tuning of the concentrations is most straightforward by adjustment of the mRNA and protein decay rates k_{dm} and k_{dp} . For the above reaction network one can show that the average mRNA number N_M and protein number N_P is given by

$$N_M = \frac{K_4 K_1 V}{K_2 N_R + V(1 + K_1(1 + K_3))}, \quad (12)$$

$$N_P = K_5 N_M, \quad (13)$$

where $K_1 = k_{\text{IRp}}/(k_{\text{bRp}} + k_{\text{OC}})$, $K_2 = k_{\text{IR}}/k_{\text{bR}}$, $K_3 = k_{\text{OC}}t_{\text{clear}}$, $K_4 = k_{\text{OC}}/k_{\text{dm}}$, and $K_5 = k_{\text{ribo}}/k_{\text{dp}}$ are equilibrium constants, V is the volume of the cell, and N_R is the total number of repressors. The unrepresed state corresponds to $N_R = 0$. In our simulations, we fix the mRNA and protein numbers in the unrepresed state at $N_M = 50$ and $N_P = 2 \times 10^5$. The mRNA and protein decay rates then follow straightforwardly from Eqs. 12 and 13: the mRNA degradation rate is $k_{\text{dm}} = 0.019 \text{ s}^{-1}$ (46) and the protein degradation rate is $k_{\text{dp}} = 2.4 \times 10^{-4} \text{ s}^{-1}$; the latter corresponds to protein degradation by dilution with a cell cycle time of $\sim 1 \text{ h}$.

Next, we determine by what factor these concentrations should decrease in the repressed state. This can be done by changing the number of repressors, N_R , and the repressor backward rate, k_{bR} . We define the repression level f as the transcription initiation rate in the absence of repressors, divided by the initiation rate in the repressed state (47). For a repression level f , the concentration of mRNA and proteins in the repressed state is a fraction $1/f$ of the concentration in the unrepresed state and it follows that

$$\frac{N_R}{k_{\text{bR}}} = (f - 1) \frac{V(1 + K_1(1 + K_3))}{k_{\text{IR}}}. \quad (14)$$

Thus, a fixed repression level f does not specify a unique combination of N_R and k_{bR} : increasing the number of repressors twofold, while also increasing the repressor

backward rate by the same factor, gives the same repression level. This means that the cell can control mRNA and protein levels in the repressed state either by having a large number of repressors that stay on the DNA for a short time or by having a small number of repressors, possibly even one, that stay on the DNA for a long time. Even though it is conceivable that the latter is preferable for economic reasons, there is no difference between the two extremes in terms of the average gene expression. In our simulations, we vary N_R and k_{bR} , but use a fixed repression level $f = 100$. Consequently, in the repressed state, on average $N_M = 0.5$ and $N_P = 200$.

Lastly, we would like to emphasize that, while all reaction rates were, as much as possible, taken from experiments, it should be realized that the measured rates might not be very precise. However, we believe that this does not affect the main conclusions of our work.

SIMULATION TECHNIQUE

We simulate the above reaction network using GFRD (19, 20). GFRD is an event-driven algorithm, which combines in one step the propagation of the particles in space with the reactions between them. The main idea is to determine at each iteration of the simulation, a maximum time step, such that only single particles or pairs of particles have to be considered. For these particles, the Smoluchowski equation (48) can be solved exactly using Green's functions. For each single particle, the Green's function is just the Gaussian distribution function $p_1(\mathbf{r}, t|\mathbf{r}_0, t_0)$, which yields the probability that, given that the particle is at point \mathbf{r}_0 at time t_0 , it is at position \mathbf{r} at a later time t . For each pair of particles, two Green's functions are obtained: one for their center-of-mass, and one for their interparticle vector \mathbf{r} ; the latter, $p_2(\mathbf{r}, t|\mathbf{r}_0, t_0)$, yields the probability that the interparticle vector \mathbf{r}_0 at time t_0 becomes \mathbf{r} at a later time t . Importantly, the interparticle Green's function does not only take into account the diffusion of the particles, but also the reactions between them. This makes it possible to derive for each pair of particles the propensity function $q(t|r_0)$, which yields the probability that the pair will react for the first time at time t , given that the particles were separated by a distance r_0 initially. The propensity functions, then, can be used to set up an event-driven algorithm, quite analogous to kinetic Monte Carlo algorithms for zero-dimensional master equations, such as the Gillespie algorithm (49). The event-driven nature allows GFRD to make large jumps in time and space when the particles are far apart from each other, making it up to five orders of magnitude more efficient than brute-force Brownian dynamics. For details of the algorithm, in particular on how the Green's functions and the propensity functions are derived, we refer to Refs. (19,20).

As discussed above, only the operator site O and the repressor particles R are simulated in space. All other reactions are assumed to occur homogeneously within the cell and are simulated according to the well-stirred model (49) or

with fixed time delays for reaction steps involving elongation. A few modifications with respect to the algorithm described in van Zon and ten Wolde (19,20) are implemented to improve simulation speed. First, we neglect excluded volume interactions between repressor particles mutually because the concentration of repressor is very low. This means that the only potential reaction pairs we consider are operator-repressor pairs. Secondly, we use periodic boundary conditions instead of a reflecting boundary, which leads to a larger average time step. Because the operator site O is both small compared to the volume of the cell and is far removed from the cell boundary, this has no effect on the dynamics of the system. Finally, because the repressor backward rate k_{bR} is rather small, the operator site can be occupied by a repressor for a time long compared to the average simulation time step. If the repressor is bound to the operator site longer than a time $L^2/6D$, where L is the length of the sides of our container, the other repressor molecules diffuse on average from one side of the box to the other. Consequently, when the repressor eventually dissociates from the operator site, the other repressor molecules have lost all memory of their positions at the time of repressor binding. Here, when a repressor will dissociate after a time longer than $L^2/6D$, we do not propagate the other repressors with GFRD, but we only update the master equation and fixed delay reactions. We update the positions of the free repressors at the moment that the operator site becomes accessible again, by assigning each free repressor molecule a random position in the container; the dissociated repressor is put at contact with the operator site. We see no noticeable difference between this scheme and results obtained by the full GFRD algorithm described previously (19,20).

To obtain accurate statistics, especially upon notoriously difficult quantities such as noise and power spectra, very long simulations were performed. A total number of 24 simulations were performed, one for each combination of parameter values (N_R , k_{OC}). A single simulation took on average 24 h of CPU time on a Pentium IV 3.0 GHz processor.

SIMULATION RESULTS: DYNAMICS AND NOISE

To study the effect of spatial fluctuations on the repression of genes, we simulate the reaction network described in Eqs. 3–11 both by GFRD—thus explicitly taking into account the diffusive motion of the repressor particles—and according to the well-stirred model, where the repressor particles are assumed to be homogeneously distributed in space and the dynamics depends only on the concentration of repressor. In Fig. 1 we show the behavior of mRNA and protein numbers for a system with open complex formation rate $k_{OC} = 30 \text{ s}^{-1}$ and with varying numbers of repressors N_R . We keep the repression factor fixed at $f = 100$ so that with increasing N_R the repressor backward rate k_{bR} is also increased, i.e., repressor particles are bound to the DNA for a shorter time.

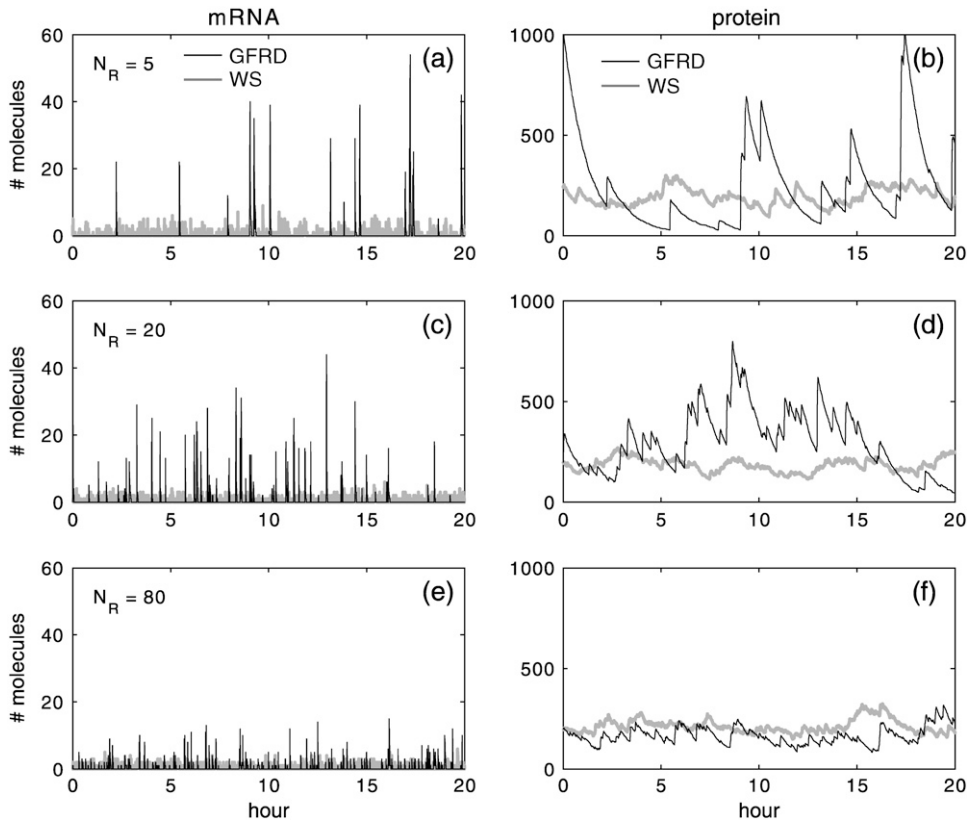


FIGURE 1 Dynamics of mRNA and protein numbers in the repressed state for different number of repressors N_R . The number of mRNA and protein molecules is shown for simulations with GFRD (black line) and according to the zero-dimensional master equation, the well-stirred model (WS, gray line). In the GFRD simulation, diffusion of repressor particles is explicitly included. (a, b) $N_R = 5$. (c, d) $N_R = 20$. (e, f) $N_R = 80$. In general, there is a dramatic difference in dynamics due to the spatial fluctuations of the repressor molecules. This difference becomes more pronounced as the number of repressors decreases. However, we find that in all cases $\langle N_{\text{mRNA}} \rangle = 0.5$ and $\langle N_P \rangle = 200$, on average. The rate of open complex formation $k_{\text{OC}} = 30 \text{ s}^{-1}$ and, when N_R varies, the repression strength is kept constant at $f = 100$ by changing k_{bR} (see Eq. 14).

It is clear from Fig. 1 that there is a dramatic difference between the behavior of mRNA and protein numbers between the GFRD simulation and the well-stirred model. When spatial fluctuations of the repressor molecules are included, mRNA is no longer produced in a continuous fashion, but instead in sharp, discontinuous bursts during which the mRNA level can reach levels comparing to those of the unrepresed state. These bursts in mRNA production consequently lead to peaks in protein number. As the protein decay rate is much lower than that of mRNA, these peaks are followed by periods of exponential decay over the course of hours. Due to these fluctuations, protein numbers often reach levels of $\sim 5\text{--}10\%$ of the protein levels in the unrepresed state. In contrast, in the absence of repressor diffusion, the fluctuations around the average protein number are much lower. For both cases, however, the average behavior is identical: even though the dynamics is very different, we always find that on average $\langle N_{\text{mRNA}} \rangle = 0.5$ and $\langle N_P \rangle = 200$. Also, in all cases the fluctuations in mRNA number are larger than those in protein number. This means that the translation step functions as a low-pass filter to the repressor signal.

When we increase the number of repressors N_R and change k_{bR} in such a way that the repression level f remains constant, we find that both for GFRD and the well-stirred model the fluctuations in mRNA and protein number decrease. In the absence of spatial fluctuations this effect is minor, but for GFRD this decrease is sharp: for a large

number of repressors, the burst in mRNA become both weaker and more frequent. This in turn leads to smaller peaks and shorter periods of exponential decay in protein numbers. In fact, as N_R is increased both approaches converge to the same behavior. At around $N_R \approx 100$, the dynamics of the protein number is similar for the well-stirred model and the spatially resolved model. The same happens for mRNA number when $N_R \approx 500$.

In Fig. 2, we quantify the noise in mRNA and protein number, defined as standard deviation divided by the mean, while we change the number of repressors, N_R . As we keep the amount of repression fixed at $f = 100$, we simultaneously vary the backward rate k_{bR} according to Eq. 14. When all parameters are the same, the noise for the GFRD simulation, including the diffusive motion of the repressors, is always larger than the noise for the well-stirred model, where the diffusive motion is ignored. In both cases, the noise decreases when the number of repressors is increased and the repressor backward rate becomes larger. This is consistent with the mRNA and protein tracks shown in Fig. 1. We also investigated the effect of changing the open complex formation rate k_{OC} . In nature, this rate can be tuned by changing the basepair composition of the promoter region on the DNA. When we change k_{OC} , we change the mRNA decay rate k_{dm} so that the average mRNA and protein concentrations remain unchanged (see section “Transcription and translation”). We find that when k_{OC} is lowered, the fluctuations in mRNA and protein levels are sharply reduced.

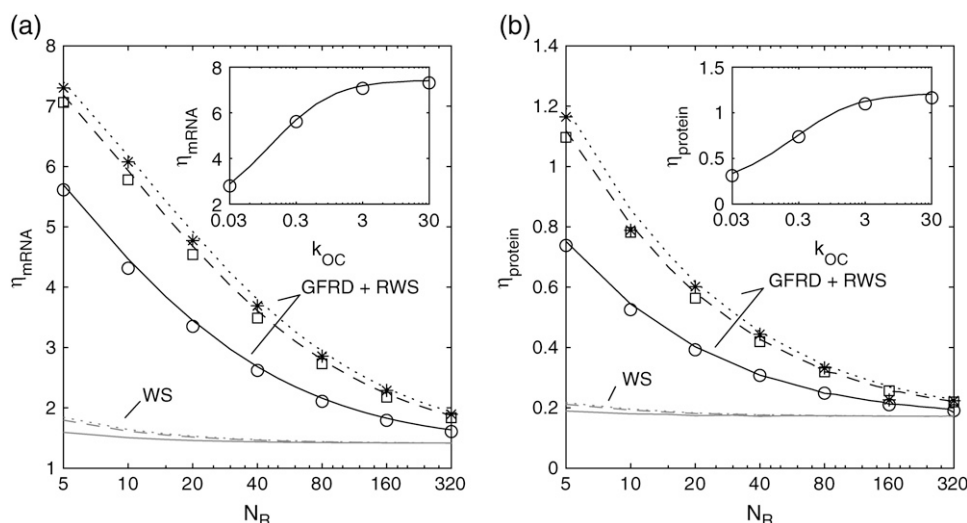


FIGURE 2 Noise in (a) mRNA number and (b) protein number as a function of the number of repressors N_R and for constant repression factor $f = 100$. Data obtained by GFRD simulation is shown for $k_{OC} = 0.3$ (○), 3 (□), and 30 (*) s^{-1} . Noise levels for the well-stirred model (WS) are shown as gray lines and those for the well-stirred model with reaction rates renormalized according to Eqs. 16 and 17 (the renormalized well-stirred (RWS) model) are shown as black lines, both for $k_{OC} = 0.3$ (solid lines), 3 (dashed lines), and 30 (dotted lines) s^{-1} . Only when the reaction rates are properly renormalized does the noise in the well-stirred model agree well with the noise in the GFRD simulations, which include the effect of diffusion. (Insets) Noise levels as a function of k_{OC} for $N_R = 5$. Symbols indicate results for GFRD and lines are results for the chemical master equation with renormalized reaction rates (RWS model).

When k_{OC} is much larger than the RNAP backward rate $k_{bRP} = 0.5 s^{-1}$, almost every RNAP binding to the promoter DNA will result in transcription of an mRNA. For k_{OC} smaller than k_{bRP} , RNAP binding will lead to transcription only infrequently. As a consequence, the operator filters out part of the fluctuations in RNAP binding due to the diffusive motion of the repressor particles, leading to the decrease in noise observed in Fig. 2. This shows that the open complex formation rate plays a considerable role in controlling noise in gene expression.

SIMULATIONS RESULTS: OPERATOR BINDING

To understand how the diffusive motion of repressor molecules leads to increased fluctuations in mRNA and protein

numbers, it is useful to look in some detail at the dynamics of repressor-DNA binding. In Fig. 3 A, we show the *OR* bias for both GFRD and the well-stirred model. The *OR* bias is a moving time average over *OR*(*t*) with a 50-s time window and should be interpreted as the fraction of time the operator site was bound by repressor particles over the last 50 s. The results we show here are for $N_R = 5$ repressors and a repression factor $f = 2$. At this repression factor, k_{bR} is such that the repressor molecules are bound to the operator only 50% of the time, making it easier to visualize the operator dynamics than in the case of $f = 100$ as used above.

The *OR* bias for the well-stirred model fluctuates around the average value $\langle OR \rangle = 0.5$, indicating that on the time-scale of 50 s several binding and unbinding events occur, in agreement with $k_{bR} = 1.26 s^{-1}$ for $f = 2$. On the other hand,

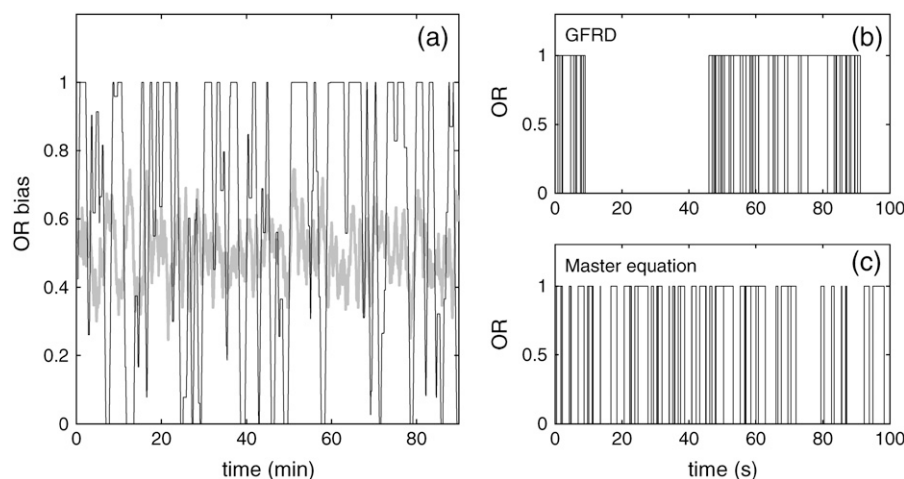


FIGURE 3 Dynamics of repressor binding for a repression factor of $f = 2$ and $N_R = 5$. (a) The *OR*-bias for GFRD (black line) and the well-stirred model (gray line). The *OR*-bias is defined as the fraction of time a repressor is bound to the operator site in the last 50 s. When the diffusive motion of repressor molecules is included (black line), the *OR*-bias switches between periods where repressors are continuously bound to or absent from the DNA for long times. (b, c) Time trace of the occupancy of the operator site by repressor molecules. When $OR = 1$ a repressor is bound to the operator site and $OR = 0$ indicates either a free operator site or one with RNAP bound. For the GFRD simulations, an initial binding is followed by several rapid rebindings, whereas for the well-stirred model binding and rebinding is much more unstructured. Note that here, for reasons of clarity, $f = 2$ instead of $f = 100$ as used in the text and Figs. 1 and 2.

when including spatial fluctuations, the *OR* bias switches between periods in which repressors are bound to the DNA continuously and periods in which the repressors are virtually absent, both on timescales much longer than the 50-s time window. How is it possible that repressors are bound to the operator site for times much longer than the timescale set by the dissociation rate from the DNA? The answer to that question can be found in Fig. 3, *B* and *C*, where a time trace is shown of the operator occupancy by the repressor for both GFRD and the well-stirred model. The time trace for the simulation of the well-stirred model in Fig. 3 *C* shows a familiar picture: binding and dissociation of the repressor from the operator occurs irregularly, the time between events given by Poisson distributions. The time trace for GFRD in Fig. 3 *B* looks rather different. Here, in general a dissociation event is followed by a rebinding very rapidly. Only occasionally does a dissociation result in the operator being unbound by repressors for a longer time. When this happens, repressors stay away from the operator for a time much longer than the typical time separating binding events in Fig. 3 *C*. These series of rapid rebindings followed by periods of prolonged absence from the operator result in the aberrant *OR* bias shown in Fig. 3 *A*.

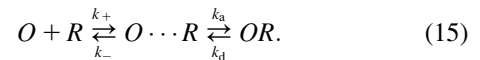
The occurrence of rapid rebindings is intimately related to the nature of diffusion. When diffusion and the positions of the reactants are ignored all dynamics is based only on the average concentration of the reactants. As a consequence, when in this approach a repressor dissociates from the operator site, the probability of rebinding depends only on the concentration of repressor in the cell. On the level of actual positions of the reactants, this amounts to placing the repressor at a random position in the container. The situation is very different for the GFRD approach, where the positions of the reactants are taken into account. After a dissociation from the operator site, the repressor particle is placed at contact with the operator site. Because of the close proximity of the repressor to its binding site, it has a high probability of rapidly rebinding to, and only a small probability of diffusing away from, the binding site. At the same time, when the repressor eventually diffuses away from the operator site, the probability that the same, or more likely, another repressor diffuses to and binds the operator site is much smaller than the probability of binding in the well-stirred model, as will be shown quantitatively in the section “Two-step kinetic scheme”. This results in the behavior observed in Fig. 3 *B*.

It can now be understood that the bursts in mRNA production correspond to the prolonged absence of repressor from the operator site compared to the well-stirred model. Especially for low repressor concentrations, these periods of absence can be long enough that the concentration of mRNA reaches values comparable to those in the unrepresed state for brief periods of time. When a repressor binds to the operator site, due to the rapid rebindings it will remain bound effectively for a time much longer than the mRNA lifetime, leading to long periods where mRNA is absent in the cell.

This shows that under these conditions spatial fluctuations and not stochastic chemical kinetics are the dominant contribution to the noise in mRNA and protein numbers in the repressed state.

TWO-STEP KINETIC SCHEME

In the current model, the average repressor concentration profile is uniform. It is therefore natural to investigate to what extent the effect of diffusion on the repressor dynamics can be described by a zero-dimensional, well-stirred, model, via the following two-step kinetic scheme (50,51):



The first step in Eq. 15 describes the diffusion of repressor to the operator site resulting in the encounter complex $O \cdots R$, with the rates k_+ and k_- depending on the diffusion coefficient D and the size of the particles. The next step describes the subsequent binding of repressor to the DNA. In this case the rates are related to the microscopic rates defined in Eq. 3. When the encounter complex is assumed to be in steady state, the two-step kinetic scheme can be mapped onto the reaction described in Eq. 3, but with effective rate constants $k'_{fr} = k_+k_a/(k_- + k_a)$ and $k'_{br} = k_-k_d/(k_- + k_a)$ (50). The two-step kinetic scheme should yield the same average concentrations as the scheme in Eq. 3, so that the equilibrium constant $K = k_a/k_d = k'_{fr}/k'_{br} = k_{fr}/k_{br}$, where k_{fr} and k_{br} are the reaction rates defined in Eq. 3.

It is possible to express the effective rate constants k'_{fr} and k'_{br} in terms of the microscopic rate constants k_{fr} and k_{br} . For the setup used here, where a single operator O is surrounded by a homogeneous distribution of repressor R , the rate k_+ follows from the solution of the steady-state diffusion equation with a reactive boundary condition with rate $k = k_a$ at contact (48,51) and is given by the diffusion-limited reaction rate $k_D = 4\pi\sigma D$. The rates k_- and k_a depend on the exact definition of the encounter complex $O \cdots R$. It is natural to identify the rate k_d with the intrinsic dissociation rate k_{br} , thus $k_d = k_{br}$. From these expressions for k_+ and k_d and the requirement that the equilibrium constant should remain unchanged, one finds that $k_a/k_- = k_{fr}/k_D$. Using this result one obtains $k'_{fr} = k_D k_{fr}/(k_D + k_{fr})$ and $k'_{br} = k_D k_{br}/(k_D + k_{fr})$.

These renormalized rate constants have a clear interpretation. For the effective forward rate it follows, for instance, that: $1/k'_{fr} = 1/k_D + 1/k_{fr}$; that is, on average, the time required for repressor binding is given by the time needed to diffuse toward the operator plus the time for a reaction to occur when the repressor is in contact with the operator site (51). The effective backward rate has a similar interpretation. The probability that after dissociation the repressor diffuses away from the operator site and never returns is given by $S_{irr}(t \rightarrow \infty | \sigma)$, where $S_{irr}(t, r_0)$ is the irreversible survival probability for two reacting particles (52). Using that $S_{irr}(t \rightarrow \infty | \sigma) = k_D/(k_{fr} + k_D)$, the expression for k'_{br} can

be written as $k'_{\text{bR}} = k_{\text{bR}} S_{\text{IT}}(t \rightarrow \infty | \sigma)$: that is, the effective dissociation rate is the microscopic dissociation rate multiplied by the probability that after dissociation the repressor escapes from the operator site (51).

For diffusion-limited reactions, such as the reaction considered here, we have that $k_{\text{fR}} \gg k_{\text{D}}$. Now, the renormalized rate constants reduce to:

$$k'_{\text{fR}} = k_{\text{D}}, \quad (16)$$

$$k'_{\text{bR}} = k_{\text{D}} k_{\text{bR}} / k_{\text{fR}}. \quad (17)$$

In Fig. 2, we compare the noise profiles for the GFRD algorithm with those obtained by a simulation of the well-stirred model, where instead of the microscopic rates k_{fR} and k_{bR} we use the renormalized rates from Eqs. 16 and 17. Surprisingly, we find complete agreement. One of the main reasons why this is unexpected, is that for the master equation the time between events is Poisson-distributed, whereas after a dissociation the time to the next rebinding is distributed according to a power-law distribution when diffusion is taken into account (52). The reason that this power-law behavior of rebinding times is not of influence on the noise profile, is that the timescale of rapid rebinding is much shorter than any of the other relevant timescales in the network. Specifically, rebinding times are so short that the probability that a RNAP will bind before a rebinding is negligible. As a consequence, the transcription network is not at all influenced by the brief period the operator site is accessible before rebinding: for the transcription machinery the series of consecutive rebindings, albeit distributed algebraically in time individually, is perceived as a single event. And on much longer timescales, when a repressor diffuses in from the bulk toward the operator site, the distribution of arrival times is expected to be Poissonian, because on these timescales the repressors are distributed homogeneously in the bulk. This is succinctly summarized in Fig. 4, which shows the distribution of association times. It is seen that at short timescales, the association events are algebraically distributed in time—these arise from the rapid rebindings—whereas at long timescales, they are distributed exponentially in time. For comparison, we also show the distribution of the repressor-DNA association times in the well-stirred model, with appropriately renormalized rate constants for repressor (un)binding (Eqs. 16 and 17). As expected, the number of association events is much smaller at short timescales, but follows the same distribution as that of the spatially resolved model at long timescales. As described quantitatively in the section “Noise propagation”, the rate constant for the exponential relaxation is given not only by the diffusion-limited rate of repressor-DNA association, but also by the RNAP promoter occupancy.

It is possible to reinterpret the effective rate constants in Eqs. 16 and 17 in the language of rapid rebindings. The probability p that a rebind will occur after a dissociation from the DNA is given by $p = 1 - S_{\infty}$, where $S_{\text{t}} = S_{\text{IT}}(t, r_0 = \sigma)$.

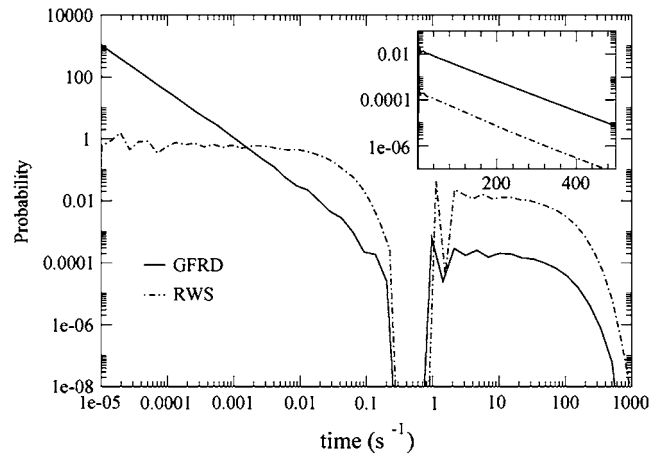


FIGURE 4 Distribution of repressor-DNA association times for the spatially resolved model (*GFRD*, solid black line) and for the well-stirred, zero-dimensional model in which the rate constants for repressor-DNA (un)binding are given by the intrinsic (un)binding rates divided by the number of repressor rebindings in the absence of RNAP (see Eqs. 18 and 19)—the renormalized well-stirred model (*RWS*, dot-dashed line). It is seen that for the spatially resolved model, at short times the distribution follows a power law, while at long times it is exponential. Moreover, at long times, the distribution of the spatially resolved model (*GFRD*) follows that of the renormalized well-stirred model (*RWS*). To a good approximation, the relaxation rate at long times in both the spatially resolved model and in the renormalized well-stirred model is given by $k'_{\text{fR}}[R_{\text{T}}][O]'$, where k'_{fR} is the effective repressor association rate (see Eq. 18), $[R_{\text{T}}]$ is the total repressor concentration, and $[O]'$ is the probability that the promoter is not occupied by RNAP, given that it is not occupied by repressor (see Eq. 28). The dip in the distribution at $t \approx 0.1\text{--}1.0 \text{ s}^{-1}$ is due to the competition with the RNAP for binding to the promoter. The model parameters are $f = 100$, $k_{\text{OC}} = 30 \text{ s}^{-1}$, $N_{\text{R}} = 5$.

The probability that n consecutive rebindings occur before the repressor diffuses away from the operator site is then given by $p_n = (1 - S_{\infty})^n S_{\infty}$. From this follows that the average number of rebindings is $N_{\text{RB}} = (1 - S_{\infty})/S_{\infty}$. Using again that $S_{\infty} = k_{\text{D}}/(k_{\text{fR}} + k_{\text{D}})$, we find that $N_{\text{RB}} = k_{\text{fR}}/k_{\text{D}}$. Combining this with Eqs. 16 and 17, we get:

$$k'_{\text{fR}} = k_{\text{fR}}/N_{\text{RB}}, \quad (18)$$

$$k'_{\text{bR}} = k_{\text{bR}}/N_{\text{RB}}. \quad (19)$$

In words, after an initial binding the repressor spends N_{RB} times longer on the DNA than expected on the basis of the microscopic backward rate, as it rebinds on average N_{RB} times. Because the average occupancy should not change, the forward rate should be renormalized in the same way. In conclusion, in this model the effects of diffusion can be properly described by a well-stirred model when the reaction rates are renormalized by the average number of rebindings.

POWER SPECTRA

In this section, we study how the noise due to the stochastic dynamics of the repressor molecules propagates through the different steps of gene expression for both the spatially

resolved model and the well-stirred model. This analysis will also provide further insight into why the well-stirred model with renormalized rate constants for the (un)binding of the repressor molecules works so well.

In biochemical networks, the noise in the output signal depends upon the noise in the biochemical reactions that constitute the network, the so-called intrinsic noise, and on the noise in the input signal, called extrinsic noise (6,14,53–56). In our case, the output signal is the protein concentration, whereas the input signal is provided by the repressor concentration. The intrinsic noise arises from the biochemical reactions that constitute the transcription and translation steps. Moreover, we consider the noise in the protein concentration that is due to the (un)binding of the RNAP to (from) the DNA to be part of the intrinsic noise. The extrinsic noise is provided by the fluctuations in the binding of the repressor to the operator, i.e., in the state *OR*. Because the total repressor concentration, $[R_T] = [R] + [OR]$, is constant, the extrinsic noise is also given by the fluctuations in the concentration of unbound repressor.

The noise properties of biochemical networks are most clearly elucidated via the power spectra of the time traces of the copy numbers of the components. Recently, we have shown that if the fluctuations in the input signal are uncorrelated with the noise in the biochemical reactions that constitute the processing network, the power spectrum of the output signal is given (56)

$$S_P(\omega) = S_{in}(\omega) + g(\omega)S_{ex}(\omega). \quad (20)$$

Here, $S_P(\omega)$ is the power spectrum of the output signal, the protein concentration. The spectrum $S_{in}(\omega)$ denotes the intrinsic noise of the processing network; it is defined as the noise in the output signal in the absence of noise in the input signal. Here, the intrinsic noise is due to the biochemical reactions of transcription and translation. The spectrum $S_{ex}(\omega)$ is the power spectrum of the input signal, which, in this case, is given by the noise in the concentration of unbound repressor: $S_{ex}(\omega) = S_R(\omega)$; because the total repressor concentration is constant, this power spectrum is also directly related to that of the repressor-bound state of the operator, $S_{OR}(\omega)$. The function $g(\omega)$ is a transfer function, which indicates how fluctuations in the input signal are transmitted toward the output signal. If the extrinsic noise is uncorrelated with the intrinsic noise, then $g(\omega)$ is an intrinsic quantity that only depends upon properties of the processing network, and not upon properties of the incoming signal (56). However, for the network studied here, the noise in the input signal is not uncorrelated with the intrinsic noise (56). As we have shown recently, this means that Eq. 20 is not strictly valid (56); the extrinsic contribution to the power spectrum of the output signal can no longer be factorized into a function that only depends upon intrinsic properties of the network, $g(\omega)$, and one that only depends upon the input signal, $S_{ex}(\omega)$. This relation is nevertheless highly instructive. Indeed, Eq. 20

could be interpreted as a heuristic definition of the transfer function $g(\omega)$.

The diffusive motion of the repressor molecules impede an analytical evaluation of the power spectrum for the extrinsic noise. Moreover, whereas power spectra can be calculated analytically for linear reaction networks (57), the delays in transcription resulting from promoter clearance and elongation, preclude the derivation of an analytical expression for the power spectrum of the intrinsic noise. We have, therefore, obtained the power spectra $S_P(\omega)$, $S_{ex}(\omega)$, and $S_{in}(\omega)$, directly from the time traces of the copy numbers. The power spectrum of a component *X* is given by $S_X(\omega) = \langle |\tilde{X}(\omega)|^2 \rangle$, where $\tilde{X}(\omega)$ is the Fourier transform of the concentration $X(t)$ of component *X*. Conventional FFT algorithms are not convenient for computing the power spectra, because our signals vary over a wide range of timescales. We therefore adopted a novel and efficient procedure, which is described in the Appendix. This procedure should prove useful for computing the power spectra of time traces of copy numbers of species in biochemical networks, as obtained by kinetic Monte Carlo simulations.

As indicated above, the intrinsic noise, $S_{in}(\omega)$, is defined as the noise in the output signal in the absence of fluctuations in the input signal. To determine the intrinsic contribution to the noise in the protein concentration, we discarded the (un)binding reaction of the repressor to the DNA (Eq. 3), while rescaling the rate k_{bRP} for the dissociation reaction of the RNAP from the DNA (Eq. 4) in such a way that the average concentration of the protein *P* remains unchanged. This eliminates the extrinsic noise arising from the repressor dynamics, thereby allowing us to obtain the intrinsic noise of the reactions in Eqs. 4–11. The rescaled backward rate k_{bRP}^* is given by

$$k_{bRP}^* = k_{bRP}(1 + K_2 N_R/V) + k_{OC} K_2/V \quad \text{where} \\ K_2 = k_{fR}/k_{bR}. \quad (21)$$

For the interpretation of the power spectra of the mRNA and protein concentration, as discussed below, it is instructive to recall the power spectrum of a linear birth-and-death process,



with rate constants k and μ . For the interpretation of the spectra of repressor binding to the DNA, it is useful to recall the spectrum of a two-state model,



with rate constants k_1 and k_2 . For both models, the power spectrum is a Lorentzian function of the form (58):

$$S(\omega) = \frac{2\sigma^2\mu}{\mu^2 + \omega^2}. \quad (24)$$

For the birth-and-death process, the variance in the concentration of *A*, σ^2 , is k/μ , whereas for the two-state system,

the variance σ^2 in the occupancy n is $n(1 - n)$; the decay rate in the two-state model is $\mu = k_1 + k_2$. The corner frequency μ (in both models) yields the timescale on which fluctuations relax back to steady state. We also note that the noise strength σ^2 is given by the integral of the power spectrum $S(\omega)$: $\sigma^2 = 1/(2\pi) \int_{-\infty}^{\infty} d\omega S(\omega)$. The noise strength is thus dominated by those frequencies at which the power spectrum is largest.

In the next subsection, we discuss the effect of spatial fluctuations on the noise in gene expression and explain why a well-stirred model with renormalized rate constants for repressor (un)binding can capture its effect. In the subsequent section, we discuss how the noise is propagated through the different stages of gene expression.

Spatial fluctuations

In Fig. 5, we show the power spectra for the input and output signals, for both the spatially resolved model and the well-stirred model with renormalized rate constants for repressor (un)binding (see previous section). We recall that the output signal is the protein concentration, whereas the input signal is the concentration of unbound repressor (the extrinsic noise). Fig. 5 also shows the power spectrum of the intrinsic noise. This is the noise in the protein concentration (the output signal), when the noise in the input signal resulting from the repressor dynamics has been eliminated by the procedure outlined above. The power spectra have been obtained in a

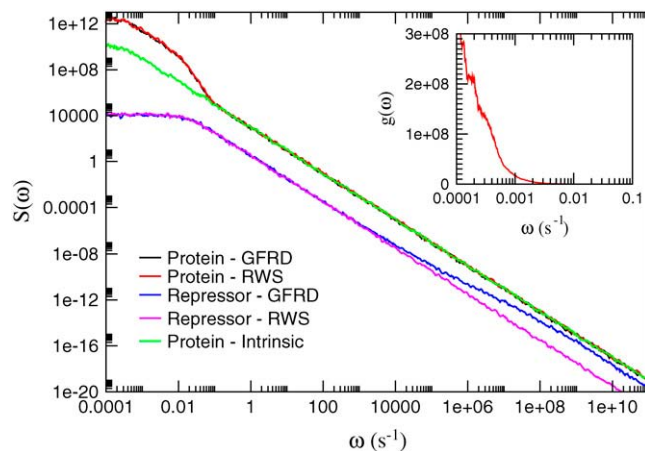


FIGURE 5 Power spectra of the repressor and protein concentrations for $f = 100$, $k_{OC} = 30 \text{ s}^{-1}$, $N_R = 5$. Data are shown both for the renormalized well-stirred model (RWS) with reaction rates renormalized according to Eqs. 16–17, and for GFRD, taking into account the spatial fluctuations of the repressor molecules explicitly. Also shown is the power spectrum of the intrinsic noise, which is the power spectrum of the protein concentration in the absence of fluctuations in the repressor concentration (extrinsic noise). For large ω , the repressor spectrum (extrinsic noise) differs between the well-stirred and the spatially resolved model. However, this difference does not appear in the power spectra of the protein concentration. The inset shows the frequency-dependent gain $g(\omega)$ (see Eq. 20).

parameter regime where the diffusing repressors have a large effect on the noise: $k_{OC} = 30 \text{ s}^{-1}$, $N_R = 5$ (see Fig. 2).

Fig. 5 shows that the power spectrum of the protein concentration in the spatially resolved model is identical to that in the well-stirred model for the entire range of frequencies observed. This confirms the observation in the section “Two-step kinetic scheme” that the effect of the spatial fluctuations of the repressor molecules on the noise in the protein concentration can be described by a well-stirred model in which the reaction rates for repressor (un)binding to the DNA are properly renormalized.

Fig. 5 also elucidates the reason why a well-stirred model with properly renormalized rate constants for repressor (un)binding can successfully describe the effect of the diffusive motion of the repressor molecules on the noise in gene expression. It is seen that the repressor spectrum for the renormalized well-stirred model is accurately described by a Lorentzian function with a corner frequency $\mu = 0.02 \text{ s}^{-1}$ (see also Eq. 24), as expected for the dynamics of repressor (un)binding dynamics (see next section). The repressor spectrum of the spatially resolved model fully overlaps with that of the well-stirred model up to a frequency of $\omega \approx 10^6 \text{ s}^{-1}$, but for higher frequencies it shows a clear deviation from the ω^{-2} behavior. This deviation is caused by the diffusive motion of the repressor molecules. Indeed, the deviation occurs at frequencies comparable to the inverse of the typical timescale for rapid rebindings ($\sim \mu\text{s}$). However, this difference between the spectrum of the repressor dynamics in the spatially resolved model and that in the well-stirred model does not lead to a difference in the noise strength of the protein concentrations of the two respective models (see Fig. 2), for two reasons: 1), the difference only occurs at high frequencies, i.e., in a frequency regime where the fluctuations only marginally contribute to the noise strength (the difference in area under the curves of the repressor power spectra for the two models is $< 5\%$); 2), the repressor fluctuations in this frequency range are filtered out by the processing network of transcription and translation; as a result of this, the effect of the small difference in area under the curves of the repressor power spectra for the two models is reduced even further. The filtering properties of the processing network are illustrated in the inset of Fig. 5, which shows the transfer function $g(\omega)$ as obtained from $g(\omega) = (S_P(\omega) - S_{in}(\omega))/S_{ex}(\omega)$ (see Eq. 20). Clearly, the transfer function rapidly decreases as the frequency increases. This shows that the processing network of transcription and translation acts as a low-pass filter, rejecting the high frequency noise in the repressor dynamics that originates from the rapid rebindings.

The only effect of the repressor rebindings on the noise in gene expression is thus that it lowers the effective dissociation rate (and association rate), as explained in the previous section. As compared to the well-stirred model with the unrenormalized rate constants for repressor (un)binding, this decreases the corner frequency μ in the repressor power

spectrum (see Fig. 6), but increases the power at low frequencies—recall that for a two-state model, which relaxes monoexponentially, the power spectrum at zero frequency is $S(\omega = 0) = 2\sigma^2/\mu$, which thus increases as the relaxation rate $\mu = k_1 + k_2$ decreases as a result of the slower binding and unbinding of repressor (see Eq. 24). The higher power in the repressor spectrum at low frequencies for the spatially resolved model and for the well-stirred model with the renormalized rate constants, as compared to that for the well-stirred model with the unrenormalized rate constants, is not filtered by the processing network of transcription and translation and thus manifests itself in the power spectrum of the protein concentration. Spatial fluctuations of gene regulatory proteins thus increase the noise in gene expression by increasing the power of the input signal at low frequencies.

Noise propagation

In Fig. 7 we show how fluctuations in the input signal, arising from the dynamics of repressor binding and unbinding, are propagated through the different stages of gene expression. In Fig. 7 *a* we illustrate how the noise in the repressor concentration (the extrinsic noise) is transferred to the level of transcription. The figure shows, for both the spatially resolved model and for the well-stirred model with renormalized rate constants for repressor (un)binding, the power spectrum of the repressor concentration and the spectrum of the concentration of the elongation complex, defined as $[ORp^*] + [T]$. It is clear from Fig. 7 *a* that already at the level of the elongation complex, the high-frequency

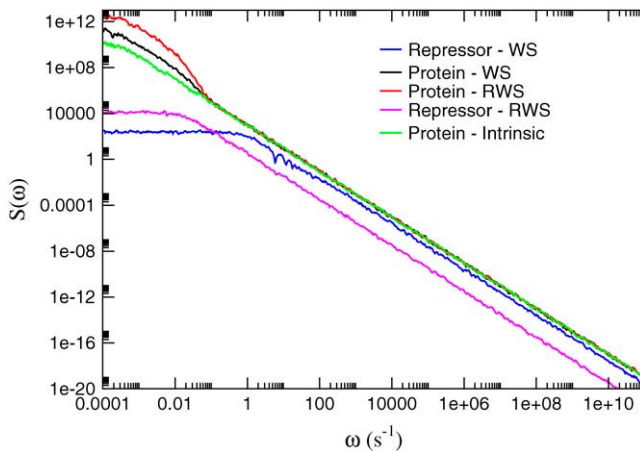
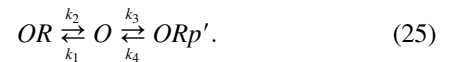


FIGURE 6 The power spectra for the well-stirred model with unrenormalized rate constants (WS) and for the well-stirred model with renormalized rate constants for repressor (un)binding to (from) the DNA (RWS). The intrinsic noise of gene expression is the same for both models. The extrinsic noise arising from the repressor dynamics is, however, markedly different. The repressor spectrum for the well-stirred model with renormalized rate constants has lower corner frequency but, more importantly, also a higher power at low frequencies. The increased power at low frequencies is not filtered by the processing network and increases the noise in gene expression. For parameter values, see Fig. 5.

noise due to the rapid rebinding is filtered. Transcription can thus already be described by a well-stirred model with properly renormalized rate constants for repressor (un)binding to (from) the DNA.

The power spectrum of the elongation complex exhibits two corner frequencies, one around $\omega_+ \approx 40 \text{ s}^{-1}$ and another one at $\omega_- \approx 0.02 \text{ s}^{-1}$. These two corner frequencies arise from the competition between repressor and RNAP for binding to the promoter. To elucidate this, we have plotted in the inset the power spectrum for RNAP bound to the promoter, thus the power spectrum for $[ORp] + [ORp^*]$. It is seen that this power spectrum has the same two corner frequencies as that of the elongation complex, showing that their dynamics is dominated by the same processes—repressor binding and RNAP binding to the promoter. These two corner frequencies can be estimated analytically by considering the reactions in Eqs. 3–6 as a three-state system, in which repressor and RNAP compete for binding to the promoter:



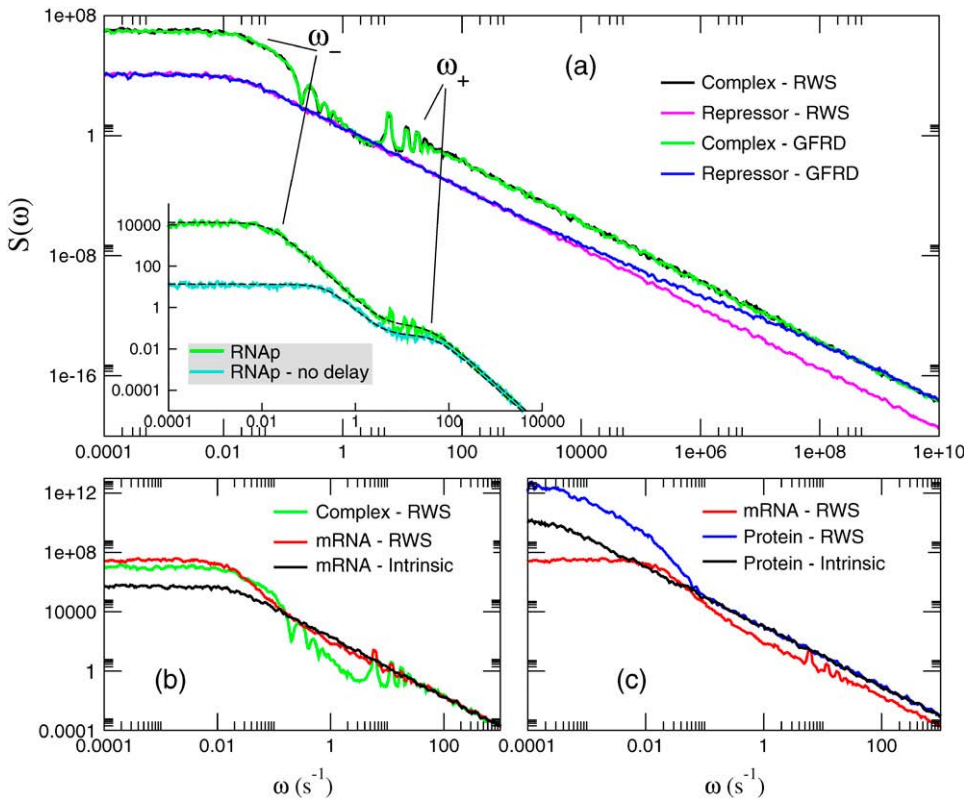
Here, $ORp' = ORp + ORp^*$, where ORp denotes the RNAP bound to the promoter in the closed complex and ORp^* denotes RNAP bound to the promoter in the open complex. The rate constant k_1 denotes the rate at which a repressor binds to the promoter; it is given by $k_1 = k'_{\text{IR}}[R_T]$, where k'_{IR} is the renormalized association rate (see Eq. 16). The rate constant k_2 denotes the renormalized rate for repressor unbinding, $k_2 = k'_{\text{bR}}$ (see Eq. 17); $k_3 = k_{\text{IRp}}$ denotes the rate at which RNAP binds to the promoter. The rate constant k_4 is the rate at which the RNAP leaves the promoter. Since the promoter can become accessible for the binding of another RNAP or repressor by either the dissociation of RNAP from the closed complex or by forming the open complex and then clearing the promoter, this rate is given by $k_4 = k_{\text{bRp}} + (k_{\text{OC}}^{-1} + t_{\text{clear}})^{-1}$. If promoter clearance would be neglected, then, indeed, $k_4 = k_{\text{bRp}} + k_{\text{OC}}$.

The power spectrum of the RNAP dynamics in Eq. 25 can be calculated analytically and is given by a sum of two Lorentzians:

$$S_{ORp'}(\omega) = \frac{A\omega_-}{\omega_-^2 + \omega^2} + \frac{B\omega_+}{\omega_+^2 + \omega^2}, \quad (26)$$

where A and B are coefficients. The corner frequencies ω_- and ω_+ are given by $\omega_{\pm} = (k \pm \sqrt{k^2 - 4h})/2$, where $k = \sum_i k_i$ and $h = k_1k_4 + k_2(k_3 + k_4)$. The dynamics of repressor binding and unbinding is much slower than that of RNAP binding and unbinding, meaning that $k_1, k_2 \ll k_3, k_4$. This allows us to approximate the corner frequencies as $\omega_+ = k_3 + k_4$ and $\omega_- = k_2 + k_1k_4/(k_3 + k_4)$. This yields the following expressions for the corner frequencies:

$$\omega_+ = k_{\text{IRp}} + k_{\text{bRp}} + (k_{\text{OC}}^{-1} + t_{\text{clear}})^{-1} \quad (27)$$



high frequencies, the mRNA dynamics is well described by a linear birth-and-death process. (c) Spectra of mRNA and protein. The slow protein dynamics filters out all the peaks resulting from the delays in the RNAP dynamics. The only difference between the full spectrum of the output signal and that of the intrinsic noise is an increased noise at low frequencies, due to the repressor dynamics. For parameter values, see Fig. 5.

$$\omega_- = k'_{br} + k'_{fr}[R_T][O]'. \quad (28)$$

Here, $[O]' \equiv k_4/(k_3 + k_4)$ is the conditional probability that the promoter is not occupied by the RNAP, given that it is not occupied by repressor; it is given by the occupancy of the promoter by RNAP in the absence of any repressor molecules in the system. We can now see that the highest corner frequency, ω_+ , describes the fast dynamics of RNAP binding to, and clearing from, the promoter and that the other corner frequency, ω_- , represents the slow dynamics of repressor (un)binding to the DNA in the presence of the fast RNAP bindings to the promoter; the lower corner frequency, ω_- , is also the corner frequency in the repressor spectrum of the renormalized well-stirred model (see Figs. 5 and 6). In Fig. 7a we plot the power spectrum $S_{ORp^*}(\omega)$ as predicted by the three-state model (Eq. 26; with fitted coefficients A and B) on top of the power spectrum obtained from the simulations and find excellent agreement. We also show the power spectra when we neglect the delay due to promoter clearance. As expected, in the absence of the delay due to promoter clearance, the lower corner frequency, ω_- , and, to a smaller extent, the higher corner frequency, ω_+ , are shifted to higher frequencies.

The power spectrum of the elongation complex in Fig. 7a contains information that is not easily observed in the time

domain and could as a result be helpful in the interpretation of the results. It is seen that there are two series of peaks. Those are associated with the two processes with fixed time delays. The first process is the promoter clearance, which takes a fixed time t_{clear} . Indeed, the first peak in the corresponding series of peaks in the power spectrum of the elongation complex, is at $\omega \approx 2\pi/(t_{clear}) = 6.3 \text{ s}^{-1}$; the other peaks in the series are the higher harmonics that naturally arise for processes with fixed time delays. The second process is the transcript elongation process. After the elongation complex has been formed, it takes a fixed time $t_{clear} + t_{elon}$ before the full transcript is formed and the RNAP dissociates from the DNA; the first valley of the corresponding series of peaks/valleys is, indeed, at $\omega \approx 2\pi/(t_{clear} + t_{elon}) = 0.2 \text{ s}^{-1}$. While the frequency $2\pi/t_{clear}$ yields, to a good approximation, the rate at which the elongation complex signal increases, the frequency $2\pi/(t_{clear} + t_{elon})$ corresponds to the frequency at which the elongation complex signal decreases; this explains why the shapes of the respective series of peaks and valleys are reciprocal. Lastly, the reason that both peaks and valleys are broadened is that the delay in the formation of the elongation complex is not fully deterministic: the duration of the delay is not only determined by the promoter clearance time, which, indeed, is fixed, but also by the time it takes for another RNAP to bind the DNA

and then form the open complex—in the absence of repressor, the average frequency at which an elongation complex is formed is given by $2\pi/(k_{\text{fRP}}^{-1} + k_{\text{OC}}^{-1} + t_{\text{clear}})$ (see also Eqs. 4–6). Both RNAP binding and open complex formation are modeled as Poisson processes, and this leads to a distribution of delay times for the formation of the elongation complex.

For completeness, in Fig. 7, *b* and *c*, we examine how the noise in the dynamics of the elongation complex propagates to the level of mRNA and protein dynamics. In Fig. 7 *b*, we compare the full power spectrum of the mRNA concentration with that of the elongation complex—the input signal (extrinsic noise) for the mRNA signal—and that of the intrinsic noise of the mRNA signal; to compute the intrinsic noise, we have modeled the mRNA dynamics as a birth-and-death process (see Eq. 22) with a production rate as given by the average production rate for the full system in Eqs. 3–11. As expected, for higher frequencies ($\omega > 0.1 \text{ s}^{-1}$), the full spectrum of mRNA overlaps almost fully with that of the intrinsic noise, although some traces of the input signal (the elongation complex) are still apparent in this high frequency regime; these are the peaks at $\omega \approx 6.3 \text{ s}^{-1}$ corresponding to promoter clearance. At lower frequencies ($\omega < 0.1 \text{ s}^{-1}$), the noise in the mRNA signal is dominated by the extrinsic noise, which is the noise in the elongation complex (the input signal). Indeed, both the spectrum of the elongation complex and that of mRNA have a corner frequency at ω_- , which, as discussed above, arises from the slow repressor (un)binding to the DNA in the presence of the fast DNA-(un)binding kinetics of RNAP.

Fig. 7 *c* shows how the noise in the mRNA concentration is propagated to that in the protein concentration. Again, at higher frequencies, the spectrum of the protein concentration coincides with that of the intrinsic noise of protein synthesis, which, as above for mRNA, has been computed by modeling protein production as a birth-and-death process; note also that the remnants of operator clearance (the peaks in the spectrum at $\omega \approx 6.3 \text{ s}^{-1}$) have been filtered by the slow protein dynamics. Only for frequencies smaller than $\omega \approx 0.1 \text{ s}^{-1}$, does the extrinsic noise—the noise in the mRNA concentration—strongly contribute to the noise in the protein concentration. A careful inspection of the protein spectrum shows that it has a “corner” at ω_- , which arises from the repressor DNA-(un)binding dynamics (the extrinsic noise), and one, albeit much less visible, at $\omega \approx k_{\text{dp}} = 2 \times 10^{-4} \text{ s}^{-1}$, which is due to the intrinsic dynamics of protein degradation.

DISCUSSION AND OUTLOOK

Our analysis reveals that at high frequencies both mRNA and protein synthesis are well described by a linear birth-and-death model. In this frequency regime, the effect of spatial fluctuations, originating from the rapid repressor rebindings, is completely filtered by the slow dynamics of transcription and translation. These rebindings do, however, decrease the

effective rate at which the repressor molecules associate with, and dissociate from, the promoter. This increases the intensity of the extrinsic (repressor) noise in the low frequency regime. Moreover, the low-frequency fluctuations in the repressor binding do propagate through the different stages of gene expression. In particular, they lead to sharp bursts in the production of mRNA and protein. These bursts increase the noise intensity at the lower frequencies in the noise spectrum of mRNA and protein. And since the noise strength σ^2 is dominated by fluctuations in the low-frequency regime, spatial fluctuations ultimately strongly increase the noise in mRNA and protein concentration.

Recently, experiments have been performed in which the synthesis of individual mRNA transcripts (59) and individual protein molecules (60) could be detected. The systems in these studies were very similar to that studied here: a gene under the control of a (Lac) repressor. These studies unambiguously demonstrated that mRNA production (59) and protein synthesis can occur in bursts (60). Of particular interest is the pulsatile transcription, which has been observed in experiments by Golding et al. (59) and in our simulations, but not in the experiments of Yu et al. (60). We therefore address the question of whether our analysis on transcription initiation in the section “Noise propagation” can reconcile these observations. Transcription occurs in bursts if: a), the operator is mostly in the repressed state, meaning that the repression strength f must be large; and b), when the operator is in the derepressed state, more than one transcript is formed; this means that transcription initiation must be sufficiently fast as compared to repression-DNA association (see also Eq. 25). In our simulations and in the Lac system studied by Yu et al. (60), the repression strength is indeed large, $f \approx 100$. With a typical in vivo repressor concentration of $[R_T] \approx 20 \text{ nM}$ ($N_R = 10$), the average repressor-DNA association rate, in the presence of RNAP, is $k'_{\text{fR}}[R_T][O]' \approx 0.1 \text{ s}^{-1}$ (see the reaction scheme in Eq. 25). The rate of open complex formation of the lac promoter has been measured to be on the order of 0.1 s^{-1} (32). Hence, in the Lac system approximately one mRNA molecule is produced per gene expression event. This is consistent with the observations of Yu et al. (60,61). The observed burst-like protein production in these experiments is indeed due to the fact that more than one protein is formed from one mRNA transcript (7,60,62). The repression strength, open complex formation rate, and repressor-DNA (un)binding rates for the system studied by Golding et al. are not known in similar detail (59), but, clearly, the observed pulsatile production of mRNA must mean that the repressor-DNA association rate is sufficiently slow as compared to the open complex formation rate.

The spatial fluctuations due to diffusion of the repressor molecules could have significant implications for the functioning of gene regulatory networks. Under some conditions, it might be crucial that the protein number is not only low on average, but remains low at all times. For instance, if the protein itself functions as a transcription factor, it might by

accident induce the expression of another gene, when, due to a fluctuation, its concentration crosses a particular activation threshold. Thus, not all combinations of repressor copy number N_R and repressor backward rate k_{bR} that obey Eq. 14 and thus have the same average repression strength, are necessarily equivalent in terms of function when diffusion is taken into account. If the fluctuations in the repressed state need to be small, then the cell could increase the number of repressors and decrease the binding affinity to the operator site, such that the repressor molecules stay bound to the DNA only briefly. Alternatively, the cell could minimize the effect of fluctuations by reducing the rate at which the open complex is formed by RNAP—our analysis shows that the process of open complex formation can act as a strong low-pass filter.

The rapid rebindings observed in our simulations are a general phenomenon. We now address the question of when the effect of spatial fluctuations due to diffusion can be described by a well-stirred model in which the association and dissociation rates are renormalized. In the current problem, the rebinding time for a dissociated repressor is exceedingly short. As a consequence, the probability that a RNAP binds to the promoter during this time, is vanishingly small. This is precisely the reason that the effective dissociation rate is simply the bare dissociation rate divided by the number of rebindings (see Eq. 18); the effective association rate is renormalized accordingly, because the equilibrium constant should remain unchanged (see Eq. 19). The success of the renormalized well-stirred model is thus a result of the strong separation of timescales—the timescale of repressor rebinding is well separated from that of RNAP binding. In fact, because of this strong separation of timescales, one could argue that the states in which a repressor has just dissociated from the operator should not be counted as unrepressed states, but rather as states that belong to the ensemble of microscopic states that together form the mesoscopic repressed state.

The separation of timescales also makes it possible to account for the effect of spatial fluctuations by renormalizing the association and dissociation rates in other cases. For instance, we have simulated a system in which repression occurs in a cooperative manner. In this system, the repressor backward rate is smaller when two repressors are bound to the operator than when a single repressor is bound. However, when one of the two repressors dissociates, its rebinding time is so short that the probability for the other repressor to dissociate in the mean time, is negligible for reasonable values of cooperativity. As a result, the effect of spatial fluctuations can be described by a well-stirred model with properly renormalized reaction rates. We have also studied a system in which the expression of a gene is not under the control of a repressor, but rather under the control of an activator. In this system, too, diffusion of the transcription factors leads to an enhancement of noise in gene expression through a similar mechanism.

Do these observations imply that the effect of spatial fluctuations can always be described by a well-stirred model? In the system studied here, the ligand (repressor) molecules bind to a single site. We expect that the effect of spatial fluctuations becomes more intricate when the number of binding sites for a particular ligand increases—the binding of the ligand to the different sites will then exhibit correlations. This could be important when the ligand binds to receptors that occur in dense clusters, as in bacterial chemotaxis (63,64) and in the immune response (65). In gene regulatory networks this effect could also be significant. Recently, we have shown that in *E. coli*, pairs of coregulated genes—genes that are controlled by a common transcription factor—tend to lie exceedingly close to each other on the genome (66): their promoter regions are often separated by a distance shorter than a few hundred basepairs. It is conceivable that spatial fluctuations of the transcription factors introduce correlations between the noise in the expression of these pairs of coregulated genes. This study also revealed that pairs of genes that regulate each other often lie close together, again suggesting that the diffusive motion of transcription factors could be important for the functioning of gene regulatory networks (66).

Even in the case of a single gene, the effect of spatial fluctuations is expected to be more complicated than that reported here. First and foremost, in this study we have assumed that the repressor, RNAP, and ribosome molecules diffuse freely through the cytoplasm. This is likely to be a gross assumption. In fact, it has recently been observed in *Bacillus subtilis* that RNAP resides principally inside the nucleoid whereas ribosomes are localized almost exclusively outside the nucleoid (67), suggesting that transcription and translation occur in separate spatial domains. Moreover, we have modeled the operator as a spherical site. However, as mentioned in the section “Diffusive motion of repressors”, transcription factors are believed to find their operator site via a combination of free 3D diffusion and 1D sliding along the DNA. Although on lengthscales longer than the sliding distance this process is indeed essentially 3D diffusion, on length- and timescales shorter than the sliding distance and sliding time, respectively, the dynamics is more complicated. We expect that sliding could have two important effects. First, it will increase the number of rebindings—the probability that in 1D a random walker returns to the origin is one, whereas in 3D there is a finite probability that it will escape and never return. Secondly, sliding is expected to also increase the duration of the rebindings, especially when diffusion along the DNA is much slower than diffusion in the cytoplasm. It is thus likely that with sliding, the nonexponential relaxation of the operator state, arising from the rebindings, shifts to lower frequencies (see Fig. 5). Indeed, it is conceivable that with sliding, a dissociated repressor molecule can compete with RNAP for binding to the promoter. Under these conditions, the effect of spatial fluctuations might be detected experimentally in the statistics of the

synthesis of the individual mRNA molecules, which could be useful for unraveling the mechanism and dynamics of transcription initiation. Importantly, we nevertheless expect that even under these conditions, the mRNA noise strength (variance) can be described by a zero-dimensional model because the life time of the mRNA molecules, setting the timescale for time averaging, is probably still longer than the duration of the rebinding trajectories. However, the effective rate constant for repressor-DNA dissociation might no longer simply be given by the bare dissociation rate divided by the number of rebindings in the absence of RNAP. Indeed, it will depend upon the spatial fluctuations of the repressor molecules and their interplay with the RNAP-DNA association dynamics in a nontrivial manner, and deriving it would probably require a spatially resolved model. We leave this for future work.

Finally, we address the question of whether spatial fluctuations, and, more in particular, the rebindings, could be studied experimentally. Interestingly, recent biochemical data on the restriction enzyme *EcoRV* suggests that after an initial dissociation, 10–100 rebindings occur before the enzyme escapes into the bulk solution (41,42), in good agreement with the average number of rebindings calculated in the section “Two-step kinetic scheme”. However, in our gene expression model, the rebinding times are so short that it would seem difficult to probe the repressor rebindings directly in an experiment. In fact, reaction rates measured biochemically will probably already be corrected for according to Eqs. 16 and 17. Sliding along the DNA, however, may extend the rebinding times to accessible experimental time-scales. Moreover, recent experiments suggest that the motion of proteins in the nucleoid might be subdiffusive, which would increase the importance of the rebindings (68). Recently, magnetic tweezer experiments on a mechanically stretched, supercoiled, single DNA have made it possible to study the kinetics of the open complex formation and promoter clearance (32). Performing these experiments on a promoter that is under the control of a repressor seems a promising approach for studying the effect of spatial fluctuations due to the diffusive motion of transcription factors on the dynamics of gene expression.

APPENDIX: COMPUTING POWER SPECTRA

The power spectrum of the time trace of the copy number $X(t)$ of a species X can be efficiently computed by exploiting the fact that in between the times t_k the signal $X(t)$ is constant. The Fourier Transform $S_X(\omega)$ of $X(t)$ is

$$\tilde{X}(\omega) = \int X(t)e^{-i\omega t} dt = \sum_k \int_{t_{k-1}}^{t_k} X_k e^{-i\omega t} dt. \quad (29)$$

As $X(t)$ is constant within every interval $\{t_{k-1}, t_k\}$, the integration can easily be performed:

$$\tilde{X}(\omega) = \sum_k X_k \frac{1}{-i\omega} (e^{-i\omega t_k} - e^{-i\omega t_{k-1}}). \quad (30)$$

Shifting up by one the index j in the second part of the sum, we obtain:

$$\tilde{X}(\omega) = \frac{1}{i\omega} \sum_k (X_{k+1} - X_k) (e^{-i\omega t_k}). \quad (31)$$

The real and imaginary parts of the Fourier transform are thus:

$$\Re[\tilde{X}(\omega)] = \frac{1}{\omega} \sum_k \delta_k (\sin \omega t_k) \quad (32)$$

$$\Im[\tilde{X}(\omega)] = \frac{1}{\omega} \sum_k \delta_k (\cos \omega t_k), \quad (33)$$

where we have defined $\delta_k = X_{k+1} - X_k$. The power spectrum $S_X(\omega) = \Re[\tilde{X}(\omega)]^2 + \Im[\tilde{X}(\omega)]^2$ is thus given by

$$S_X(\omega) = \left(\frac{1}{\omega} \sum_k \delta_k \cos(\omega t_k) \right)^2 + \left(\frac{1}{\omega} \sum_k \delta_k \sin(\omega t_k) \right)^2. \quad (34)$$

The Fourier transforms were computed at 10,000 logarithmically spaced angular frequencies starting from $\omega_{\min} = 10 \times 2\pi/T$, where T is the total length of the signal. Power spectra obtained according to Eq. 34 were filtered with a box average over 20 neighboring points.

We thank Johan Paulsson, Mans Ehrenberg, and Frank Poelwijk for useful discussions and a critical reading of the manuscript.

The work is part of the research program of the “Stichting voor Fundamenteel Onderzoek der Materie (FOM)”, which is financially supported by the “Nederlandse organisatie voor Wetenschappelijk Onderzoek (NWO)”.

REFERENCES

1. Delbruck, M. 1945. The burst size distribution in the growth of bacterial viruses (bacteriophages). *J. Bacteriol.* 50:131–135.
2. Novick, A., and M. Weiner. 1957. Enzyme induction as an all-or-none phenomenon. *Proc. Natl. Acad. Sci. USA.* 43:553–566.
3. McAdams, H. H., and A. Arkin. 1997. Stochastic mechanisms in gene expression. *Proc. Natl. Acad. Sci. USA.* 94:814–819.
4. Rigney, D. R., and W. C. Schieve. 1977. Stochastic model of linear, continuous protein synthesis in bacterial populations. *J. Theor. Biol.* 69:761–766.
5. Berg, O. G. 1978. A model for the statistical fluctuations of protein numbers in a microbial population. *J. Theor. Biol.* 71:587–603.
6. Elowitz, M. B., A. J. Levine, E. D. Siggia, and P. S. Swain. 2002. Stochastic gene expression in a single cell. *Science.* 297:1183–1186.
7. Ozbudak, E. M., M. Thattai, I. Kurtser, A. D. Grossman, and A. van Oudenaarden. 2002. Regulation of noise in the expression of a single gene. *Nat. Genet.* 31:69–73.
8. Rao, C. V., D. M. Wolf, and A. P. Arkin. 2002. Control, exploitation and tolerance of intracellular noise. *Nature.* 420:231–237.
9. Kaern, M., T. C. Elston, W. J. Blake, and J. J. Collins. 2005. Stochasticity in gene expression: from theories to phenotypes. *Nat. Rev. Genet.* 6: 451–464.
10. Paulsson, J. 2005. Models of stochastic gene expression. *Physics of Life Reviews.* 2:157–175.
11. van Kampen, N. G. 2001. *Stochastic Processes in Physics and Chemistry*. Elsevier, Amsterdam, The Netherlands.
12. Gillespie, D. T. 1976. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *J. Comp. Phys.* 22:403–434.
13. Swain, P. S., M. B. Elowitz, and E. D. Siggia. 2002. Intrinsic and extrinsic contributions to stochasticity in gene expression. *Proc. Natl. Acad. Sci. USA.* 99:12795–12800.

14. Paulsson, J. 2004. Summing up the noise in gene networks. *Nature*. 427:415–418.
15. Rosenfeld, N., J. W. Young, U. Alon, P. S. Swain, and M. Elowitz. 2005. Gene regulation at the single-cell level. *Science*. 307:1962–1965.
16. Elf, J., and M. Ehrenberg. 2004. Spontaneous separation of bi-stable biochemical systems into spatial domains of opposite phases. *Sys. Biol.* 2:230–236.
17. Ander, M., P. Beltrao, B. D. Ventura, J. Ferkinghoff-Borg, M. Foglierini, A. Kaplan, C. Lemerle, I. Tomas-Oliveira, and L. Serrano. 2004. Smartcell, a framework to simulate cellular processes that combines stochastic approximation with diffusion and localisation: analysis of simple networks. *Sys. Biol.* 1:129–138.
18. Andrews, S. S., and D. Bray. 2004. Stochastic simulation of chemical reactions with spatial resolution and single molecule detail. *Phys. Biol.* 1:137–151.
19. van Zon, J. S., and P. R. ten Wolde. 2005. Simulating biochemical networks at the particle level and in time and space: Green's function reaction dynamics. *Phys. Rev. Lett.* 94:128103.
20. van Zon, J. S., and P. R. ten Wolde. 2005. Green's-function reaction dynamics: a particle-based approach for simulating biochemical networks in time and space. *J. Chem. Phys.* 123:234910.
21. Ko, M. S. H. 1991. A stochastic model for gene induction. *J. Theor. Biol.* 153:181–194.
22. Peccoud, J., and B. Ycart. 1995. Markovian modeling of gene-product synthesis. *Theor. Popul. Biol.* 48:222–234.
23. Kepler, T. B., and T. C. Elston. 2001. Stochasticity in transcriptional regulation: origins, consequences, and mathematical representations. *Biophys. J.* 81:3116–3136.
24. Metzler, R. 2001. The future is noisy: the role of spatial fluctuations in genetic switching. *Phys. Rev. Lett.* 87:068103.
25. Karmakar, R., and I. Bose. 2004. Graded and binary responses in stochastic gene expression. *Phys. Biol.* 1:197–204.
26. Pirone, J. R., and T. C. Elston. 2004. Fluctuations in transcription factor binding can explain the graded and binary responses observed in inducible gene expression. *J. Theor. Biol.* 226:111–121.
27. Simpson, M. L., C. D. Cox, and G. S. Saylor. 2004. Frequency domain chemical Langevin analysis of stochasticity in gene transcriptional regulation. *J. Theor. Biol.* 229:383–394.
28. Bialek, W., and S. Setayeshgar. 2005. Physical limits to biochemical signaling. *Proc. Natl. Acad. Sci. USA*. 102:10040–10045.
29. Hornos, J. E. M., D. Schultz, G. C. P. Innocentini, A. M. W. J. Wang, J. N. Onuchic, and P. G. Wolynes. 2005. Self-regulating gene: an exact solution. *Phys. Rev. E*. 72:051907.
30. Record, M., W. S. Reznikoff, M. L. Craig, K. L. McQuade, and P. J. Schlax. 1996. Escherichia coli RNA polymerase (σ_{70}), promoters, and the kinetics of the steps of transcription initiation. In *Escherichia coli and Salmonella*, 2nd Ed. F. C. Neidhardt, R. Curtiss, III, J. L. Ingraham, E. C. C. Lin, K. B. Low, B. Magasanik, W. S. Reznikoff, M. Riley, M. Schaechter, and H. E. Umbarger, editors. ASM Press, Washington, DC. 792 – 821.
31. Austin, D. W., M. S. Allen, J. M. Coolom, R. D. Dar, J. R. Wilgus, S. G. Saylor, N. F. Sarmatova, C. D. Cox, and M. L. Simpson. 2006. Gene network shaping of inherent noise spectra. *Nature*. 439:608–611.
32. Revyakin, A., R. H. Ebright, and T. R. Strick. 2004. Promoter unwinding and promoter clearance by RNA polymerase: detection by single-molecule DNA nanomanipulation. *Proc. Natl. Acad. Sci. USA*. 101:4776–4780.
33. Riggs, A. D., S. Bourgeois, and M. Cohn. 1970. The lac repressor-operator interaction. III. Kinetic studies. *J. Mol. Biol.* 53:401 – 417.
34. Berg, O. G., R. B. Winter, and P. H. von Hippel. 1981. Diffusion-driven mechanisms of protein translocation on nucleic acids. I. Models and theory. *Biochemistry*. 20:6929–6948.
35. Halford, S. E., and J. F. Marko. 2004. How do site-specific dna-binding proteins find their targets? *Nucleic Acids Res.* 32:3040–3052.
36. Gerland, U., J. Moroz, and T. Hwa. 2002. Physical constraints and functional characteristics of transcription factor-DNA interaction. *Proc. Natl. Acad. Sci. USA*. 99:12015–12020.
37. Coppey, M., O. Benichou, R. Voituriez, and M. Moreau. 2004. Kinetics of target site localization of a protein on DNA: a stochastic approach. *Biophys. J.* 87:1640–1649.
38. Slutsky, M., and L. A. Mirny. 2004. Kinetics of protein-dna interaction: facilitated target location in sequence-dependent potential. *Biophys. J.* 87:4021–4035.
39. Klenin, K. V., H. Merlitz, J. Langowski, and C.-X. Wu. 2006. Facilitated diffusion of DNA binding proteins. *Phys. Rev. Lett.* 96:018104.
40. Hu, T., A. Y. Grosberg, and B. I. Shklovskii. 2006. How proteins search for their specific sites on DNA: the role of DNA conformation. *Biophys. J.* 90:2731–2744.
41. Stanford, N. P., M. D. Szczelkun, J. F. Marko, and S. E. Halford. 2000. One- and three-dimensional pathways for proteins to reach specific DNA sites. *EMBO J.* 19:6546–6557.
42. Gowers, D. M., G. G. Wilson, and S. E. Halford. 2005. Measurement of the contributions of 1D and 3D pathways to the translocation of a protein along DNA. *Proc. Natl. Acad. Sci. USA*. 102:15883–15888.
43. Elowitz, M. B., M. G. Surette, P.-E. Wolf, J. B. Stock, and S. Leibler. 1999. Protein mobility in the cytoplasm of Escherichia coli. *J. Bacteriol.* 181:197–203.
44. Bremer, H., P. Dennis, and M. Ehrenberg. 2003. Free RNA polymerase and modeling global transcription in Escherichia coli. *Biochimie*. 84:597–609.
45. McClure, M. R. 1983. A biochemical analysis of the effect of RNA polymerase concentration on the in vivo control of RNA chain initiation frequency. In *Biochemistry of Metabolic Processes*. D. L. F. Lennon, F. W. Stratman, and R. N. Zahltlen, editors. Elsevier, New York. 207–217.
46. Kushner, S. 1996. mRNA decay. In *Escherichia coli and Salmonella*, 2nd Ed. F. C. Neidhardt, R. Curtiss, III, J. L. Ingraham, E. C. C. Lin, K. B. Low, B. Magasanik, W. S. Reznikoff, M. Riley, M. Schaechter, and H. E. Umbarger, editors. ASM Press, Washington, DC. 849–860.
47. Vilar, J., and S. Leibler. 2003. DNA looping and physical constraints on transcription regulation. *J. Mol. Biol.* 331:981–989.
48. Smoluchowski, M. 1917. Versuch einer mathematischen theorie der koagulationskinetik kolloider losungen. *Z. Phys. Chem.* 92:129–168. [in German].
49. Gillespie, D. T. 1977. Exact stochastic simulation of coupled reactions. *J. Phys. Chem.* 81:2340–2361.
50. Eigen, M. 1974. Diffusion Control in Biochemical Reactions. Plenum Press, New York.
51. Shoup, D., and A. Szabo. 1982. Role of diffusion in ligand binding to macromolecules and cell-bound receptors. *Biophys. J.* 40:33–39.
52. Kim, H., and K. J. Shin. 1999. Exact solution of the reversible diffusion-influenced reaction for an isolated pair in three dimensions. *Phys. Rev. Lett.* 82:1578–1581.
53. Pedraza, J. M., and A. van Oudenaarden. 2005. Noise propagation in gene networks. *Science*. 307:1965–1969.
54. Detwiler, P. B., S. Ramanathan, A. Sengupta, and B. I. Shraimann. 2000. Engineering aspects of enzymatic signal transduction: photoreceptors in the retina. *Biophys. J.* 79:2801–2817.
55. Shibata, T., and K. Fujimoto. 2005. Noisy signal amplification in ultrasensitive signal transduction. *Proc. Natl. Acad. Sci. USA*. 102:331–336.
56. Tanase-Nicola, S., P. B. Warren, and P. R. ten Wolde. 2006. Signal detection, modularity, and the correlation between extrinsic and intrinsic noise in biochemical networks. *Phys. Rev. Lett.* 97:068102.
57. Warren, P. B., S. Tanase-Nicola, and P. R. ten Wolde. 2006. Exact results for noise power spectra in linear biochemical reaction networks. *J. Chem. Phys.* 125:144904.
58. van Kampen, N. G. 1992. Stochastic Processes in Physics and Chemistry. North-Holland, Amsterdam, The Netherlands.

59. Golding, I., J. Paulsson, S. M. Zawilski, and E. C. Cox. 2005. Real-time kinetics of gene activity in individual bacteria. *Cell*. 123:1025–1030.
60. Yu, J., J. Xiaojia, K. Lao, and X. Xie. 2006. Probing gene expression in live cells, one protein molecule at a time. *Science*. 311:1600–1603.
61. Cai, L., N. Friedman, and X. Xie. 2006. Stochastic protein expression in individual cells at the single molecule level. *Nature*. 440:358–362.
62. Thattai, M., and A. van Oudenaarden. 2001. Intrinsic noise in gene regulatory networks. *Proc. Natl. Acad. Sci. USA*. 98:8614–8619.
63. Bray, D., M. D. Levin, and C. J. Morton-Firth. 1998. Receptor clustering as a cellular mechanism to control sensitivity. *Nature*. 393:85–88.
64. Andrews, S. S. 2005. Serial rebinding of ligands to clustered receptors as exemplified by bacterial chemotaxis. *Phys. Biol.* 2:111–122.
65. Valitutti, S., S. Müller, M. Cella, E. Padovan, and A. Lanzavecchia. 1995. Serial triggering of many t-cell receptors by a few peptide-MHC complexes. *Nature*. 375:148–151.
66. Warren, P. B., and P. R. ten Wolde. 2004. Statistical analysis of the spatial distribution of operons in the transcriptional regulation network in *Escherichia coli*. *J. Mol. Biol.* 342:1379–1390.
67. Lewis, P. J., S. D. Thaker, and J. Errington. 2000. Compartmentalization of transcription and translation in *Bacillus subtilis*. *EMBO J.* 19:710–718.
68. Golding, I., and E. C. Cox. 2006. Physical nature of bacterial cytoplasm. *Phys. Rev. Lett.* 96:098102.