Geometry-Induced Bursting Dynamics in Gene Expression

B. Meyer,[†] O. Bénichou,[†] Y. Kafri,[‡] and R. Voituriez[†*]
[†]UMR 7600, Université Pierre et Marie Curie/CNRS, Paris, France; and [‡]Department of Physics, Technion, Haifa, Israel

ABSTRACT In prokaryotes and eukaryotes, genes are transcribed stochastically according to various temporal patterns that range from simple first-order kinetics to marked bursts, resulting in temporal and cell-to-cell variations of mRNA and protein levels. Here, we consider the effect of the transport of regulatory molecules on the noise in gene expression by taking into account explicitly the dynamics of a finite number of transcription factors confined in the cell. We calculate analytically time-dependent correlation functions of mRNA levels for a wide range of transport mechanisms and find that in the limit of small-transcription-factor copy number, the results differ significantly from standard approaches, which ignore confinement. It is shown how such dynamical quantities, which can now be obtained experimentally, can be used to identify the underlying mechanisms of transcription. Of particular importance, it is demonstrated that the geometry of transcription-factor trajectories in the cellular environment plays a key role in transcription kinetics, and can intrinsically generate the observed various transcription patterns ranging from simple first-order kinetics to bursts.

INTRODUCTION

Gene transcription is initiated by a series of intrinsically stochastic events of binding and unbinding of specific factors to their target sequence. These events, as well as others, such as cell division, result in fluctuations in time of messenger RNA (mRNA) and protein levels that play a key role in phenotypic variability, adaptation, and developmental transitions (1,2). Observations made possible over the last decade by new imaging techniques (3,4) show that mRNAs are produced in transcriptional patterns that depend strongly on both the gene and the cell type. For example, the transcription of constitutive genes in yeast was shown to occur in single uncorrelated events, i.e., to be well captured by simple first-order kinetics (5,6). In contrast to this simple scenario, in many cases it is found that the mRNA production occurs in bursts, which suggest more complicated kinetic schemes. In particular, transcriptional bursts seem extremely common in genes of higher eukaryotic cells (7–10). In prokaryotes, most of the data suggest that transcription is described well by first-order kinetics (3,11), although transcriptional bursts have also been observed in these cells (12).

Along with such increasing experimental evidence of variability in transcriptional patterns, it has been recognized theoretically that bursts give rise to a source of stochasticity in mRNA and protein levels that is much larger than that predicted by first-order stochastic chemical kinetics (13–16). These theoretical studies generally focus on the variance and distribution of mRNA and protein levels in the stationary state and reveal that such time-independent observables only poorly constrain transcriptional models (5,16). The microscopic origin of bursting is consequently largely undetermined, even though possible mechanisms, such as chromatin remodeling, have been proposed (2).

In this article, we discuss a prototypical model in which the limiting step in mRNA synthesis is given by the binding to the gene locus of a specific transcription factor (TF). Activation is then controlled by the distribution of the first-passage time of the TF to the target locus. It has been suggested that in many cases this transport step is a limiting factor of transcription (6), but so far, with the exception of the numerical results of van Zon et al. (17), theoretical approaches to quantify its impact on noise have been mostly restricted to classical Brownian motion (17–19) and have considered the case of a homogeneous concentration of transcription factors in infinite space.

The distribution of the first-passage time of a particle to a target in a confined volume was obtained recently for a wide range of transport processes (20) and was shown to strongly depend on the confinement and to involve a nontrivial range of timescales dependent on the geometry of the trajectories. Making use of these new tools, we quantify analytically the effect of the transport of regulatory molecules on noise in gene expression and find that in the regime of small copy numbers of TFs, the results differ significantly from those obtained by assuming a homogeneous concentration in infinite space (see Fig. 2 *c*). More precisely, our findings are as follows. 1), We quantify analytically the contribution of the transport step to the noise and show that, in contrast to the time-independent observables that have been the focus of most studies so far, time-dependent correlation functions, which can now be obtained experimentally (6,21), make it possible to identify microscopic mechanisms of transcription. 2), We distinguish between uncorrelated mRNA production events, which stem from an underlying Poisson process, and correlated bursts, which originate from nontrivial kinetics; we show that these two mechanisms are characterized by markedly different behaviors of correlation

Submitted February 10, 2012, and accepted for publication March 26, 2012.

*Correspondence: voiturie@lptmc.jussieu.fr

Editor: Edda Klipp.

0006-3495/12/05/2186/6 $2.00

doi: 10.1016/j.bpj.2012.03.060

functions. 3), Of most importance, we show that within the framework of the model, the geometry of the TF trajectories in the cellular environment can intrinsically generate the various observed transcription patterns, ranging from simple first-order kinetics to marked correlated bursts.

MODEL

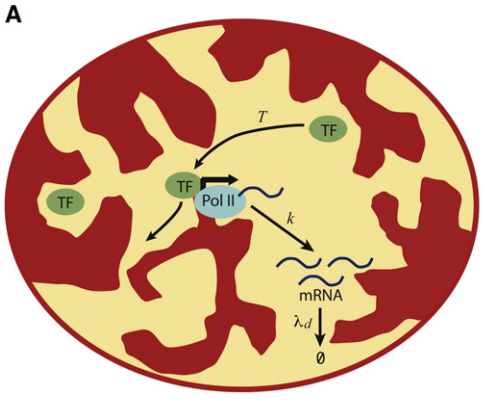
We consider the basic reaction scheme in which a gene, G, when activated by a TF, produces mRNA molecules, M, at rate k (see Fig. 1). Note that strictly speaking, we consider pre-mRNA molecules and do not aim at describing the maturation process. The TFs are assumed to diffuse in a nucleus or cell of volume V, and we consider that their copy number, n, is regulated over long timescales and remains constant. The gene G is activated only when a TF is localized at the gene locus. The time, T, between successive production events of M is then a random variable (hereafter, time variables are discrete and expressed in the unit of a small elementary time step, τ_0) whose distribution, $f_n(T)$, is governed by successive arrivals of TFs to the locus. The degradation process of a given mRNA molecule is generically described by the probability distribution, $h(t)$, of its lifetime; in practice, we will assume a single-step reaction and write $h(t) = \lambda_d e^{-\lambda_d t}$. The case of aging, where $h(t)$ is not exponential (16), is discussed in the Supporting Material. Since the typical lifetime of mRNAs is smaller than a cell cycle, cell division can be ignored. We denote by τ the average residence time/visit at the locus, which is assumed to be much smaller than the mean waiting time between activation events, $\langle T\rangle_n$ (12); the number m of transcripts produced per visit is then Poisson-distributed with mean $\langle m\rangle = k\tau$. The case of larger binding times at the gene locus, which seems relevant to higher eukaryotes, can be treated in a similar manner and leads to increased noise levels. We focus on the random variable, $n_M(t)$, which gives the number of mRNA molecules in the cell at time t.

To analyze quantitatively the fluctuations of $n_M(t)$ in the stationary state, we go beyond the standard tools, which yield the variance, σ_M^2, of $n_M(t)$ (14,16), and focus on dynamical properties, which can be quantified by the autocorrelation function $R_M(t) = \langle n_M(t_0)\, n_M(t_0+t)\rangle - \langle n_M(t_0)\rangle^2$, where $\langle n_M(t_0)\rangle$ denotes the mean number of mRNAs in the stationary state. Denoting the Fourier transform of a given function, g, by $\widehat{g}(\omega) = \int_{-\infty}^{\infty} e^{i\omega t} g(t)dt$, the autocorrelation function $R_M(t)$ can be derived analytically using queuing-statistics tools (22,23) and yields in Fourier space (see Supporting Material for details)

$$\widehat{R}_M(\omega) = \frac{\langle m\rangle}{\langle T\rangle_n}\left(1+\frac{\langle m\rangle}{2}\right)\frac{2}{\lambda_d^2+\omega^2} + \frac{\langle m\rangle^2}{\langle T\rangle_n} \times \left(\mathrm{Re}\left(\frac{\widehat{f}_n(\omega)}{1-\widehat{f}_n(\omega)}\right) - \frac{\pi}{\langle T\rangle_n}\delta(\omega)\right)\frac{2}{\lambda_d^2+\omega^2}. \quad (1)$$

This equation elucidates the dynamics of fluctuations of the mRNA copy number, and in particular readily yields the usual static measure of noise, $\sigma_M^2 = \frac{1}{2\pi}\int_{-\infty}^{\infty}\widehat{R}_M(\omega)d\omega$. It has a clear interpretation: the first term accounts for random uncorrelated mRNA production events, whereas the second term accounts for memory effects generated by the waiting-time distribution, $f_n(T)$. In particular, it can be checked that the second term vanishes only if $f_n(T)$ is a single exponential. In that simple case, the autocorrelation function $R_M(t)$ is a single exponential and reads

$$R_M(t) = \frac{\langle m\rangle}{\lambda_d\langle T\rangle_n}\left(1+\frac{\langle m\rangle}{2}\right)e^{-\lambda_d t}. \quad (2)$$



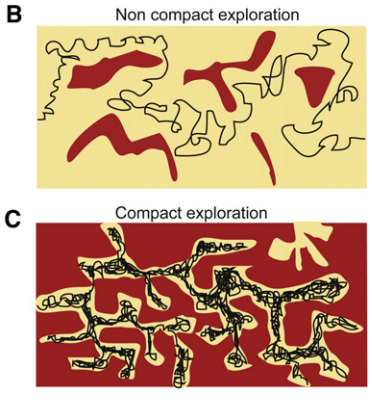


FIGURE 1 (*A*) Effect of the transport step on the noise in mRNA levels. A gene, G, is activated only when a TF is localized at the gene locus and produces mRNA molecules, M, at rate k. The TFs, of copy number n, are assumed to diffuse in a nucleus or cell of volume V. The space accessible to TFs is depicted in yellow and the space made inaccessible due to crowding effects in red. T denotes the random waiting time between successive production events of M, whose distribution, $f_n(T)$, is governed by successive returns of TFs to the locus. The mRNA molecules are degraded at rate λ_d. We will consider here both cases of eukaryotes (for which degradation generally occurs outside the nucleus) and prokaryotes. (*B*) Schematic TF trajectory when crowding effects are weak: the typical number of fast returns of a TF to a given site is small and the exploration of the cellular or nuclear environment is called noncompact. The production of mRNA is then shown to follow classical first-order kinetics. (*C*) Schematic TF trajectory when crowding effects are important, as in the case of a fractal organization of DNA: the typical number of fast returns of a TF to a given site is very large and the exploration is called compact. We show that the geometry of compact trajectories induces large correlated bursts of mRNA.

This is reminiscent of the fact that the dynamics is in this case Markovian, as is generally assumed in the literature, with the notable exception of the article by Pedraza and Paulsson (16). One can then distinguish two cases. In the first case, each visit of a TF yields a small number of mRNAs, so that $\langle m \rangle \sim 1$, as observed in Zenklusen et al. (5), leading to first-order kinetics with no bursts. The second case $\langle m \rangle \gg 1$ corresponds to a standard mechanism for producing bursts (2): each visit of a transcription factor yields a large number of mRNA molecules. Note that the bursts, whose amplitude is directly controlled by $\langle m \rangle$, are then uncorrelated, since the waiting time distribution, $f_n(T)$, is exponential. This manifests itself in $R_M(t)$ decaying in time as a single exponential.

RESULTS AND DISCUSSION

It is important to note that the assumption that the waiting-time distribution is exponential ignores the geometry of TF trajectories. Very recently, it was shown that depending on the geometry of the accessible space, the waiting-time distribution, $f_1(T)$, of a single diffusing particle to a given target site can significantly depart from a single exponential (20). Furthermore, this quantity is very different from the infinite volume limit implicitly used in works by Tkacik and colleagues (18,19) in the case of Brownian diffusion. As we now show, this results in markedly different dynamics for $n_M(t)$.

The so-called survival probability, $S_n(T)$, for n particles, i.e., the probability that the gene, activated at $t = 0$, remains unactivated until $t = T$, is readily obtained from the explicit expression of $f_1(T)$ derived in Bénichou et al. (20). One finds that $S_n(T) = \overline{S}^{n-1} S_1$, where $\overline{S}$ is the survival probability of a TF averaged over all possible starting positions, which was derived in Bénichou et al. (24). The desired waiting-time distribution, $f_n(T)$, for n particles is then eventually given by $f_n = -dS_n/dT$. At this point, we stress that the derivation of Eq. 1 for the autocorrelation function relies on a renewal property (25). Renewal holds exactly in the case of a single TF but is approximate in the case of n TFs, as it does not account for events where several TFs are present simultaneously in the vicinity of the locus. This approximation is, however, accurate in the regime of a small TF copy number, as is checked in our numerical simulations (see Figs. 2 and 3).

As discussed in Bénichou et al. (20), on general grounds, the range of timescales involved in $f_1(T)$, and therefore in $f_n(T)$, crucially depends on the compact or noncompact nature of the transport process. This property is controlled by the geometry of the trajectories and therefore by the geometry of the accessible space. In dilute 3-dimensional environments, crowding effects are weak and the average number of visits of a diffusing particle to a given site is finite: exploration is then weakly redundant and said to be noncompact (26) (see Fig. 1). When crowding effects are important, as exemplified by diffusion in fractal environments, trajectories can become highly redundant due to multiple dead ends. The average number of visits to a given site can then be infinite, and exploration is called compact (26) (see Fig. 1). For concreteness, we will consider first the case where the accessible space has fractal properties and discuss more general cases later on. The transport process can then be parameterized by its walk dimension, d_w, defined by the scaling with time of the mean-squared displacement, $\langle r^2 \rangle \propto t^{2/d_w}$ (strictly, for diffusion on a lattice, t should be taken as the number of steps), and the fractal dimension, d_f, of the medium defined by the scaling of the accessible volume with its linear dimension, $V \propto r^{d_f}$. Note that diffusion in a fractal medium is generally subdiffusive (i.e., $d_w \geq 2$). The compact case then corresponds to $d_w \geq d_f$, and the noncompact case to $d_f > d_w$ (26).

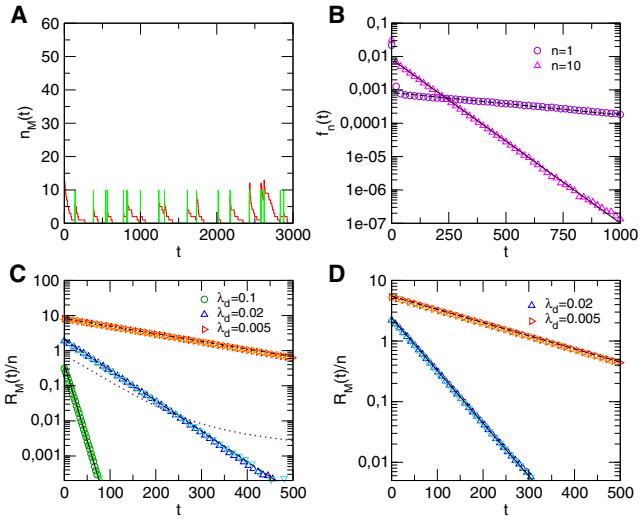


FIGURE 2 Fluctuations in mRNA level in the case of diffusion in a dilute environment, leading to noncompact exploration. The nucleus volume is $V = 10^3\ \mu m^3$, and we take $D = 1\ \mu m^2/s$, and $\langle m \rangle = 5$. (*A*) Typical evolution of the mRNA copy number, $n_M(t)$ (*red line*), obtained by numerical simulations of the model for $n = 10$, $\lambda_d = 0.02\ \text{min}^{-1}$. Here, the nucleus is modeled as a 3-dimensional cubic lattice and $\langle T \rangle = 648$ min. Green bars mark the activation of the gene. (*B*) Waiting-time distribution, $f_n(T)$, between activation events for $n = 1$ (*violet circles*) and $n = 10$ (*magenta triangles*), compared to the theoretical prediction, derived in the Supporting Material (*plain* and *dashed lines*, respectively). (*C*) Normalized autocorrelation function $R_M(t)/n$ for diffusion on a 3-dimensional cubic lattice. Numerical simulations (*symbols*) for different values of λ_d and n are compared to the theoretical prediction (*plain* and *dashed lines*). $\lambda_d = 0.1\ \text{min}^{-1}$ ($n = 1$ (*green circles*) and $n = 10$ (*light green crosses*)); $\lambda_d = 0.02\ \text{min}^{-1}$ ($n = 1$ (*dark blue triangles*) and $n = 10$ (*light blue inverted triangles*)); $\lambda_d = 0.005\ \text{min}^{-1}$ ($n = 1$ (*red triangles*) and $n = 10$ (*orange triangles*)). The dotted line is obtained by assuming a homogeneous concentration $c_0 = 1/V$ of TFs in infinite space (case $\lambda_d = 0.02\ \text{min}^{-1}$). (*D*) Normalized autocorrelation function $R_M(t)/n$ for diffusion with dilute obstacles (here modeled by a 3-dimensional supercritical percolation cluster with $p = 0.8$). $\lambda_d = 0.02\ \text{min}^{-1}$ ($n = 1$ (*dark blue triangles*) and $n = 10$ (*light blue inverted triangles*)); $\lambda_d = 0.005\ \text{min}^{-1}$ ($n = 1$ (*red triangles*) and $n = 10$ (*orange triangles*)).

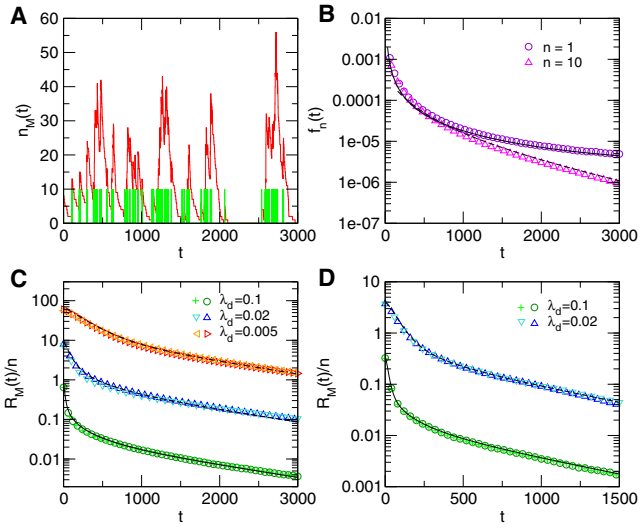


FIGURE 3 Fluctuations in mRNA levels in the case of diffusion with fractal crowding leading to compact exploration. (*A*) Typical evolution of the mRNA copy number, $n_M(t)$ (*red line*), obtained by numerical simulations of the model. The nucleus is modeled as a fractal environment (here, a 3-dimensional critical percolation cluster embedded in a 12^3 cube) embedded in a volume $V = 1.7\ 10^3\ \mu m^3$, such that $\langle T\rangle = 500$ min. Other parameter values are $\langle m\rangle = 5$, $n = 10$, $\lambda_d = 0.02\ \text{min}^{-1}$. Green bars mark the activation of the gene. (*B*) Waiting time distribution, $f_n(T)$, between activation events for $n = 1$ (*violet circles*) and $n = 10$ (*magenta triangles*), compared to the theoretical prediction derived in the Supporting Material (*plain* and *dashed lines*, respectively) for the same critical percolation cluster. (*C*) Normalized autocorrelation function $R_M(t)/n$ for diffusion on the same 3-dimensional critical percolation cluster. Numerical simulations (*symbols*) for different values of λ_d and n are compared to the theoretical prediction (*plain* and *dashed lines*, respectively). $\lambda_d = 0.1\ \text{min}^{-1}$ ($n = 1$ (*dark green circles*) and $n = 10$ (*light green crosses*)), $\lambda_d = 0.02\ \text{min}^{-1}$ ($n = 1$ (*dark blue triangles*) and $n = 10$ (*light blue inverted triangles*)), $\lambda_d = 0.005\ \text{min}^{-1}$ ($n = 1$ (*red triangles*) and $n = 10$ (*orange triangles*)). (*D*) Normalized autocorrelation function $R_M(t)/n$ for diffusion on a Sierpinski gasket (example of deterministic fractal) such that $\langle T\rangle = 729$ min. $\lambda_d = 0.1\ \text{min}^{-1}$ ($n = 1$ (*dark green circles*) and $n = 10$ (*green crosses*)), $\lambda_d = 0.02\ \text{min}^{-1}$ ($n = 1$ (*dark blue triangles*) and $n = 10$ (*light blue inverted triangles*)).

In the case of proteins in cellular and nuclear media, the existence of both compact and noncompact exploration seems to depend on protein types and their interactions with DNA. In particular, a fractal organization of the chromatin implying compact transport has been put forward by Bancaud et al. (27) and Lieberman-Aiden et al. (28). We give details of each scenario below and discuss the impact of each on the dynamics of fluctuations of the mRNA number.

Noncompact exploration

In the case of noncompact exploration, e.g., for regular diffusion in a 3-dimensional environment with weak crowding effects, the distribution of return times to the locus for n TFs can be written explicitly in the limit of large volume:

$$f_n(T) \;=\; (1-\Pi)\delta(T) + \Pi\frac{n}{\overline{\langle T\rangle}}e^{-nT/\overline{\langle T\rangle}}, \qquad (3)$$

where $\overline{\langle T\rangle} \sim V$ is the mean first-passage time for a single TF to the target averaged over all possible starting positions, the Dirac delta function $\delta(T)$ accounts for trajectories returning to the gene locus within timescales much shorter than $\overline{\langle T\rangle}$, and $1-\Pi$ is the probability of such fast returns (see Supporting Material for explicit expressions of all quantities). Using Eq. 1, the following expression of the autocorrelation function is obtained:

$$R_M(t) \simeq \frac{n\langle m\rangle}{\langle T\rangle_1\lambda_d}\left(1+\frac{\langle m\rangle}{2}+\frac{\langle m\rangle(1-\Pi)}{\Pi}\right)e^{-\lambda_d t}, \qquad (4)$$

where $\langle T\rangle_1$ denotes the mean waiting time for a single TF. The analysis of the autocorrelation function, which differs significantly from the results obtained by assuming a homogeneous concentration of TFs in infinite space (see Fig. 2 *c*) therefore shows that the dynamics of mRNA is in this case very similar to the case where $f_n(T)$ is a single exponential, discussed above: it is governed only by the degradation rate, λ_d, of mRNA and the production events of mRNAs, of typical size $\langle m\rangle$, are fully uncorrelated (see Fig. 2).

This result has important consequences. 1), First-order kinetics with a Poisson distribution of the mRNA number is recovered in the regime of small $\langle m\rangle$, and the noise in the mRNA level is then quantified by $\sigma_M^2/\langle n_M\rangle^2 = R_M(0)/\langle n_M\rangle^2 \simeq \langle T\rangle_1\lambda_d/(n\langle m\rangle) \simeq 1/\langle n_M\rangle$. 2), The quantity $\langle T\rangle_1 \sim V$ can be very large, which indicates that the transport step is a significant source of noise, especially in the regime of small TF copy number. For example, taking biologically relevant values for eukaryotes (see Fig. 2) indeed yields $\langle T\rangle_1$ of the order of hours and noise levels >1. 3), The scaling of the noise as the inverse of the mean mRNA number, here with a prefactor of order 1 (unless $\langle m\rangle$ is large), is consistent with experimental findings across various genes and even organisms (29–31). 4), Last, we note that the autocorrelation function depends linearly on the TF copy number, n, and that the dependence on $\langle m\rangle$ appears only in the amplitude of $R_M(t)$ but not in its time dependence.

Compact exploration

In the case of compact exploration, as exemplified by diffusion in fractal environments, the distribution of return times to the gene locus can be asymptotically written in the limit of large volume:

$$f_n(T) \;=\; (1-\Pi)\delta(T) + \Pi\sum_{i=1}^{\infty}\alpha_i e^{-\beta_i nT/\overline{\langle T\rangle}}, \qquad (5)$$

where $\overline{\langle T\rangle} \sim V^{d_w/d_f}$ is again the mean first-passage time for a single TF to the target averaged over all possible starting positions (see Supporting Material for explicit expressions of Π and the constants α_i, β_i). In this case the autocorrelation function can be written in Fourier space using Eq. 1.

Due to the wide spectrum of timescales involved in $f_n(T)$, fluctuations are clearly non-Poissonian in this case and display large correlated bursts even for $\langle m \rangle$ small (see Fig. 3). The major difference from the noncompact case is that in the compact case the probability Π of long waiting times is extremely small (see Supporting Material) so that repeated fast returns to the gene locus are very likely and eventually collectively contribute to large bursts of mRNA. This amplification mechanism has a clear signature in the autocorrelation function, which is in marked contrast to the Poissonian case: it shows two different timescales, one of order $1/\lambda_d$, characterizing the burst duration, and one controlled by $\overline{\langle T \rangle} \sim V^{d_w/d_f}$, stemming from the waiting times between successive bursts. More precisely, in the long time regime the autocorrelation function decays as a series of exponentials whose characteristic times are all of the order of magnitude of V^{d_w/d_f} (see Supporting Material). This shows finally that in the case of compact exploration, the very geometry of TF trajectories has a striking impact on mRNA dynamics and is at the origin of a mechanism generating large correlated bursts, which constitutes a central result of this study.

Several comments are in order. 1), The correlated bursts that we describe for compact transport are very different by nature from bursts that result only from the production of a large amount of mRNA at each visit of a TF. The latter case, which corresponds to $\langle m \rangle \gg 1$, leads to uncorrelated bursts that are characterized by an autocorrelation function with a single-exponential decay, as shown by Eq. 2. 2), The resulting noise in the mRNA level can be obtained and reads $\sigma_M^2/\langle n_M \rangle^2 \sim \langle T \rangle_1/(n\,\tau_0) \sim (\lambda_d\,\tau_0)^{-1}/\langle n_M \rangle$, where $\langle T \rangle_1 \sim V$ can be very large. Taking biologically relevant values for eukaryotes (see Fig. 3) yields $\langle T \rangle_1$ of the order of hours and typical noise levels >10, which is much larger than in the noncompact case. 3), The scaling of the noise as the inverse of the mean mRNA number is recovered and is similar to the noncompact case. However, the prefactor $(\lambda_d\,\tau_0)^{-1}$ can be large in the compact case and depends more strongly on the gene, as seems to be the case for eukaryotes (10). 4), The dependence of the autocorrelation function on n and $\langle m \rangle$ is similar to that in the noncompact case: it affects only the amplitude of the autocorrelation function but not its time dependence. In particular, it is linear in n, as illustrated in Fig. 3 (see Supporting Material for a derivation of this property).

The analysis above clearly illustrates the importance of the geometry of TF trajectories in the dynamics of mRNA levels. This result is in fact robust and remains unchanged for a broad class of transport processes. More precisely, the discussion above makes it clear that the mechanism of burst amplification relies on the occurrence with high probability of short timescales in the waiting-time distribution, $f_n(T)$, so that the hypothesis of a compact-scale invariant process made above can be relaxed and in fact only needs to hold over a given length or timescale, which only needs to be long enough to allow for multiple returns. In particular, our analysis shows that burst amplification can occur in environments that are fractal only over a finite range (as in the examples of Fig. 3), as is observed for the nuclear medium (27). This mechanism also applies to TFs whose trajectories, because of nonspecific interactions with DNA, are best described by a mechanism usually termed facilitated diffusion (32–34), that combines phases of compact sliding on the DNA and phases of noncompact bulk excursions (see, e.g., (35–39) for experimental evidence). The waiting-time distribution, $f_n(T)$, for facilitated diffusion can be obtained analytically according to the methods of others (34,40–42) and proves to be well approximated in the limit of long DNA length (compared to the average DNA length explored during a sliding phase) by the functional form described in Eq. 5. Π in this case denotes the weight of the long timescales, which is extremely small, since the compact property of sliding favors successive fast rebinding events to the gene locus (43). The dynamics of mRNA synthesis is then very similar to the compact case above: correlated bursts of mRNA of amplitude much larger than $\langle m \rangle$ emerge due to the repeated fast returns of the TFs to the gene locus, and the mRNA dynamics can be characterized by the emergence of two timescales in the autocorrelation function. We note also that the very nature of the nonspecific TF/DNA interaction was shown to yield in some cases a marked timescale separation leading to a waiting-time distribution, $f_n(T)$, similar to that in Eq. 5 (41,44).

Finally, our analysis underlines the importance of the geometry of TF trajectories and reveals the role of correlations induced by the statistics of returns. One could expect that in the case of multiple target genes regulated by the same TFs, correlations of expression levels could arise in the compact case due to a similar mechanism. This would be at work only if genes are close to each other, i.e., colocalized, which seems consistent with the available experimental data (45).

CONCLUSION

To conclude, autocorrelation functions of mRNA levels, which can now be obtained experimentally (6,21), carry clear signatures of the underlying mechanisms of transcription, as opposed to the classical static measures of noise. The geometry of TF trajectories, namely, their compact or noncompact nature, can generically generate markedly different transcription patterns ranging from first-order kinetics to correlated bursts, as observed in vivo. The compact or noncompact nature of transport strongly depends on the geometry of the cellular or nuclear environment as probed by TFs, and therefore on the nonspecific affinity for TFs with DNA, which we suggest could be used as a regulating parameter. Overall, our results highlight the importance of determining the geometry of in vivo TF trajectories, which was initiated recently (27).

SUPPORTING MATERIAL

Derivation of the correlation function of the mRNA number is available at http://www.biophysj.org/biophysj/supplemental/S0006-3495(12)00401-8.

REFERENCES

1. Elowitz, M. B., A. J. Levine, …, P. S. Swain. 2002. Stochastic gene expression in a single cell. *Science.* 297:1183–1186.
2. Raj, A., and A. van Oudenaarden. 2008. Nature, nurture, or chance: stochastic gene expression and its consequences. *Cell.* 135:216–226.
3. Cai, L., N. Friedman, and X. S. Xie. 2006. Stochastic protein expression in individual cells at the single molecule level. *Nature.* 440:358–362.
4. Li, G.-W., and X. S. Xie. 2011. Central dogma at the single-molecule level in living cells. *Nature.* 475:308–315.
5. Zenklusen, D., D. R. Larson, and R. H. Singer. 2008. Single-RNA counting reveals alternative modes of gene expression in yeast. *Nat. Struct. Mol. Biol.* 15:1263–1271.
6. Larson, D. R., D. Zenklusen, …, R. H. Singer. 2011. Real-time observation of transcription initiation and elongation on an endogenous yeast gene. *Science.* 332:475–478.
7. Becskei, A., B. B. Kaufmann, and A. van Oudenaarden. 2005. Contributions of low molecule number and chromosomal positioning to stochastic gene expression. *Nat. Genet.* 37:937–944.
8. Blake, W. J., M. KAErn, …, J. J. Collins. 2003. Noise in eukaryotic gene expression. *Nature.* 422:633–637.
9. Raj, A., C. S. Peskin, …, S. Tyagi. 2006. Stochastic mRNA synthesis in mammalian cells. *PLoS Biol.* 4:e309.
10. Suter, D. M., N. Molina, …, F. Naef. 2011. Mammalian genes are transcribed with widely different bursting kinetics. *Science.* 332:472–474.
11. Ozbudak, E. M., M. Thattai, …, A. van Oudenaarden. 2002. Regulation of noise in the expression of a single gene. *Nat. Genet.* 31:69–73.
12. Golding, I., J. Paulsson, …, E. C. Cox. 2005. Real-time kinetics of gene activity in individual bacteria. *Cell.* 123:1025–1036.
13. Paulsson, J. 2004. Summing up the noise in gene networks. *Nature.* 427:415–418.
14. Paulsson, J. 2005. Models of stochastic gene expression. *Phys. Life Rev.* 2:157–175.
15. Friedman, N., L. Cai, and X. S. Xie. 2006. Linking stochastic dynamics to population distribution: an analytical framework of gene expression. *Phys. Rev. Lett.* 97:168302.
16. Pedraza, J. M., and J. Paulsson. 2008. Effects of molecular memory and bursting on fluctuations in gene expression. *Science.* 319:339–343.
17. van Zon, J. S., M. J. Morelli, …, P. R. ten Wolde. 2006. Diffusion of transcription factors can drastically enhance the noise in gene expression. *Biophys. J.* 91:4350–4367.
18. Tkacik, G., T. Gregor, and W. Bialek. 2008. The role of input noise in transcriptional regulation. *PLoS ONE.* 3:e2774.
19. Tkacik, G., and W. Bialek. 2009. Diffusion, dimensionality, and noise in transcriptional regulation. *Phys. Rev. E.* 79:051901.
20. Bénichou, O., C. Chevalier, …, R. Voituriez. 2010. Geometry-controlled kinetics. *Nat. Chem.* 2:472–477.
21. Dunlop, M. J., R. S. Cox, 3rd, …, M. B. Elowitz. 2008. Regulatory activity revealed by dynamic correlations in gene expression noise. *Nat. Genet.* 40:1493–1498.
22. Eliazar, I. 2008. Spectral analysis of random source-medium-sink flows. *Europhys. Lett.* 82:30005.
23. Jia, T., and R. V. Kulkarni. 2011. Intrinsic noise in stochastic models of gene expression with molecular memory and bursting. *Phys. Rev. Lett.* 106:058102.
24. Bénichou, O., C. Chevalier, …, R. Voituriez. 2011. Facilitated diffusion of proteins on chromatin. *Phys. Rev. Lett.* 106:038102.
25. Redner, S. 2001. A guide to First-Passage Processes. Cambridge University Press, Cambridge, UK.
26. de Gennes, P. G. 1982. Kinetics of diffusion-controlled processes in dense polymer systems. i. nonentangled regimes. *J. Chem. Phys.* 76:3316–3321.
27. Bancaud, A., S. Huet, …, J. Ellenberg. 2009. Molecular crowding affects diffusion and binding of nuclear proteins in heterochromatin and reveals the fractal organization of chromatin. *EMBO J.* 28:3785–3798.
28. Lieberman-Aiden, E., N. L. van Berkum, …, J. Dekker. 2009. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science.* 326:289–293.
29. Bar-Even, A., J. Paulsson, …, N. Barkai. 2006. Noise in protein expression scales with natural protein abundance. *Nat. Genet.* 38:636–643.
30. Taniguchi, Y., P. J. Choi, …, X. S. Xie. 2010. Quantifying *E. coli* proteome and transcriptome with single-molecule sensitivity in single cells. *Science.* 329:533–538.
31. So, L.-H., A. Ghosh, …, I. Golding. 2011. General properties of transcriptional time series in *Escherichia coli*. *Nat. Genet.* 43:554–560.
32. Berg, O. G., R. B. Winter, and P. H. von Hippel. 1981. Diffusion-driven mechanisms of protein translocation on nucleic acids. 1. Models and theory. *Biochemistry.* 20:6929–6948.
33. Slutsky, M., and L. A. Mirny. 2004. Kinetics of protein-DNA interaction: facilitated target location in sequence-dependent potential. *Biophys. J.* 87:4021–4035.
34. Coppey, M., O. Bénichou, …, M. Moreau. 2004. Kinetics of target site localization of a protein on DNA: a stochastic approach. *Biophys. J.* 87:1640–1649.
35. Elf, J., G.-W. Li, and X. S. Xie. 2007. Probing transcription factor dynamics at the single-molecule level in a living cell. *Science.* 316:1191–1194.
36. Bonnet, I., A. Biebricher, …, P. Desbiolles. 2008. Sliding and jumping of single EcoRV restriction enzymes on non-cognate DNA. *Nucleic Acids Res.* 36:4118–4127.
37. Loverdo, C., O. Bénichou, …, P. Desbiolles. 2009. Quantifying hopping and jumping in facilitated diffusion of DNA-binding proteins. *Phys. Rev. Lett.* 102:188101–188104.
38. van den Broek, B., M. A. Lomholt, …, G. J. Wuite. 2008. How DNA coiling enhances target localization by proteins. *Proc. Natl. Acad. Sci. USA.* 105:15738–15742.
39. Lomholt, M. A., B. van den Broek, …, R. Metzler. 2009. Facilitated diffusion with DNA coiling. *Proc. Natl. Acad. Sci. USA.* 106:8204–8208.
40. Bénichou, O., C. Loverdo, …, R. Voituriez. 2008. Optimizing intermittent reaction paths. *Phys. Chem. Chem. Phys.* 10:7059–7072.
41. Bénichou, O., Y. Kafri, …, R. Voituriez. 2009. Searching fast for a target on DNA without falling to traps. *Phys. Rev. Lett.* 103:138102–138104.
42. Bénichou, O., C. Loverdo, …, R. Voituriez. 2011. Intermittent search strategies. *Rev. Mod. Phys.* 83:81–129.
43. Kolesov, G., Z. Wunderlich, …, L. A. Mirny. 2007. How gene order is influenced by the biophysics of transcription regulation. *Proc. Natl. Acad. Sci. USA.* 104:13948–13953.
44. Sheinman, M., O. Bénichou, …, R. Voituriez. 2012. Classes of fast and specific search mechanisms for proteins on DNA. *Rep. Prog. Phys.* 75:026601.
45. Fraser, P., and W. Bickmore. 2007. Nuclear organization of the genome and the potential for gene regulation. *Nature.* 447:413–417.