THE JOURNAL OF PHYSICAL CHEMISTRY LETTERS
A JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

pubs.acs.org/JPCL

Letter

How Transcription Factors Binding Stimulates Transcriptional Bursting

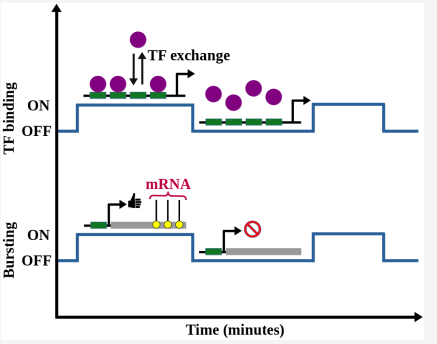
Anupam Mondal and Anatoly B. Kolomeisky*

Cite This: J. Phys. Chem. Lett. 2024, 15, 8781−8789

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

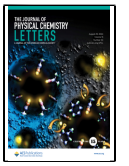
ABSTRACT: Transcription is a fundamental biological process of transferring genetic information which often occurs in stochastic bursts when periods of intense activity alternate with quiescent phases. Recent experiments identified strong correlations between the association of transcription factors (TFs) to gene promoters on DNA and transcriptional activity. However, the underlying molecular mechanisms of this phenomenon remain not well understood. Here, we present a theoretical framework that allowed us to investigate how binding dynamics of TF influences transcriptional bursting. Our minimal theoretical model incorporates the most relevant physical-chemical features, including TF exchange among multiple binding sites at gene promoters and TF association/dissociation dynamics. Using analytical calculations supported by Monte Carlo computer simulations, it is demonstrated that transcriptional bursting dynamics depends on the strength of TF binding and the number of binding sites. Stronger TF binding affinity prolongs burst duration but reduces variability, while an optimal number of binding sites maximizes transcriptional noise, facilitating cellular adaptation. Our theoretical method explains available experimental observations quantitatively, confirming the model's predictive accuracy. This study provides important insights into molecular mechanisms of gene expression and regulation, offering a new theoretical tool for understanding complex biological processes.



Gene transcription is a tightly regulated process that involves complex interactions between proteins and nucleic acids. It begins with protein transcription factors (TFs) binding to specific DNA segments in gene regulatory regions, such as promoters or distal enhancers.[1,2] These TFs then recruit other proteins such as coactivators, chromatin remodelers, and modifiers to prepare the system for gene activation.[3] This facilitates the assembly of the preinitiation complex, which includes transcription factors and RNA polymerase II (Pol II). Once assembled, Pol II initiates the transcription by escaping from the promoter and starting to synthesize mRNA by copying the gene.[4] Due to the random nature of molecular interactions and the low number of gene copies in the cellular nucleus, each step of transcription is inherently stochastic, leading to large fluctuations. Additionally, mRNA production is not continuous but occurs in bursts, characterized by periods of transcriptional activity during which a gene produces several mRNA molecules, followed by periods of quiescence without transcription.[5−10] This phenomenon is known as transcriptional bursting, and it has been observed in various species ranging from bacteria[11−13] to yeast[14,15] and to mammals.[16−18] While these observations suggest that transcriptional bursting is a universal phenomenon in gene expression, the molecular mechanisms underlying this process remain largely not well understood.

Multiple biochemical and biophysical factors, including sequences of regulatory elements,[10] enhancer-promoter interactions,[19,20] chromatin modifications,[21] nucleosomes turnover,[22] polymerases clustering,[23] TFs clustering[24] and DNA supercoiling,[25,26] have been shown to influence the transcriptional bursting. Among these factors, the binding kinetics of TFs have recently emerged as probably the most crucial element in regulating transcriptional activities. An increasing body of experimental data indicates that the bursting dynamics can be controlled by varying the stability of TF-DNA interactions.[22,27−33] For instance, recent single-molecule tracking measurements of transcription factor Gal4 binding kinetics with real-time mRNA synthesis observation in yeast have revealed that longer TF residence times are associated with increased burst duration.[22] This correlation suggests that transcription can continue as long as the TF remains bound to the DNA. Similarly, in mammalian cells, the glucocorticoid receptor (GR) has been studied with comparable high-resolution techniques, and it has been observed that stimuli increasing the amount of GR bound to chromatin enhance

Received: July 12, 2024
Revised: August 12, 2024
Accepted: August 14, 2024
Published: August 20, 2024



https://doi.org/10.1021/acs.jpclett.4c02050
J. Phys. Chem. Lett. 2024, 15, 8781−8789

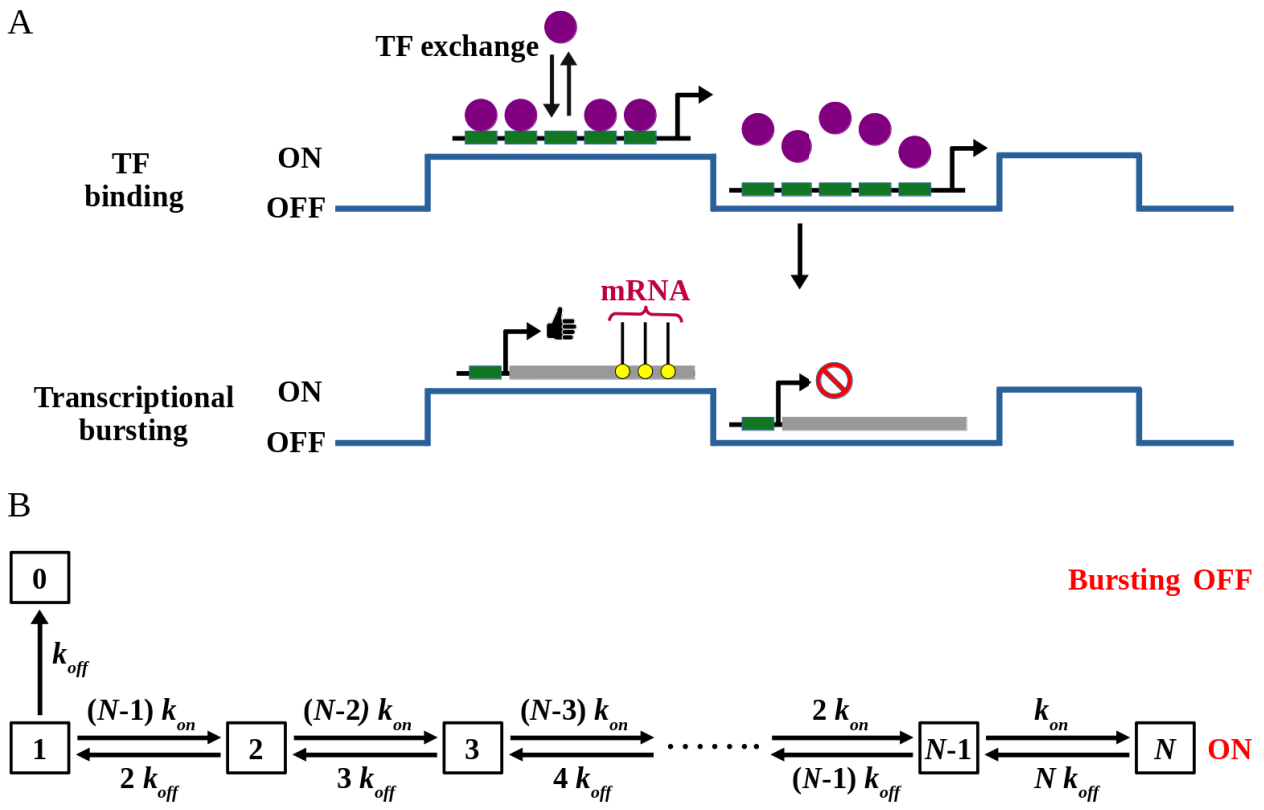


Figure 1. (A) Schematic representation of the theoretical model where TF exchange prolongs periods of high TF occupancy at the promoter, thereby increasing the burst duration. (B) Discrete-state stochastic model of the TF exchange dynamics. Each state corresponds to the number of TF-bound sites out of N total binding sites. The association/dissociation dynamics of TFs are governed by an association rate k_{on} per site and a dissociation rate k_{off} per site.

burst frequency, while longer GR residence times control burst duration.[31]

While the experiments identified the strong connection between TF binding and bursting dynamics, there is a notable discrepancy between the time scales of these two processes. TFs bind to DNA transiently, and the duration of such events is in the order of seconds,[34−36] while the time scale of transcriptional bursting typically ranges from minutes to hours.[16,37] This raises the question of how these transient TF−DNA interactions can influence prolonged periods of transcriptional activity. To bridge the discrepancy between these two time scales, three kinetic models have been proposed previously.[38−41] The first is the long-binding subpopulation model, which postulates that while most TFs bind DNA transiently, a subpopulation binds for extended periods such that dwell time directly determines burst duration. The second is the multistep activation model, which suggests that TF binding initiates a cascade of subsequent steps that continue to activate transcription even after the TF has dissociated. The third is the cooperative binding with exchange model, which proposes that DNA-bound TFs enhance the binding rate of subsequent TFs. Despite each TF having a short residence time, a continuous exchange of TFs maintains high TF occupancy during bursts.

A recent experimental study by Pomp et al.[42] thoroughly tested these proposals using the yeast TF Gal4 and its highly expressed target gene GAL10. By developing novel imaging techniques that simultaneously tracked Gal4 binding and transcriptional bursting at GAL10, it was shown that the dynamic exchange with cooperative binding of Gal4 at the GAL10 promoter predominantly drives prolonged transcriptional bursts. This continuous exchange allows each TF to bind transiently but be quickly replaced by another, thereby extending the period of high TF occupancy and bridging the time scale gap (see Figure 1A). While this study significantly enhances our understanding of the dynamic coupling between TF binding and transcription, the precise microscopic mechanisms that link these processes remain elusive.

In this letter, we propose a new theoretical framework for understanding the microscopic details of coupling between TF binding and transcriptional bursting. A discrete-state stochastic model that explicitly incorporates the continuous exchange of TFs to their binding sites and reversible associations/dissociations is introduced. It demonstrates that the active periods of transcriptional bursts only occur when TFs are bound to DNA. By solving explicitly this model using the method of first-passage probabilities, we obtain a comprehensive description of the dynamic coupling between TF binding and transcriptional activity. Our results indicate that increasing the number of TF binding sites leads to longer burst duration. However, the noise in the system exhibits a nonmonotonic behavior as the number of binding sites changes, suggesting an optimal number of sites for maximizing transcriptional noise. Physical-chemical arguments to describe these observations are presented. Our theoretical predictions also explain experimental observations from the Gal4-GAL10 system and are further supported by Monte Carlo computer simulations. The proposed theoretical method clarifies the regulatory potential of binding site multiplicity and TF affinity in fine-tuning of gene expression.

https://doi.org/10.1021/acs.jpclett.4c02050
J. Phys. Chem. Lett. 2024, 15, 8781−8789

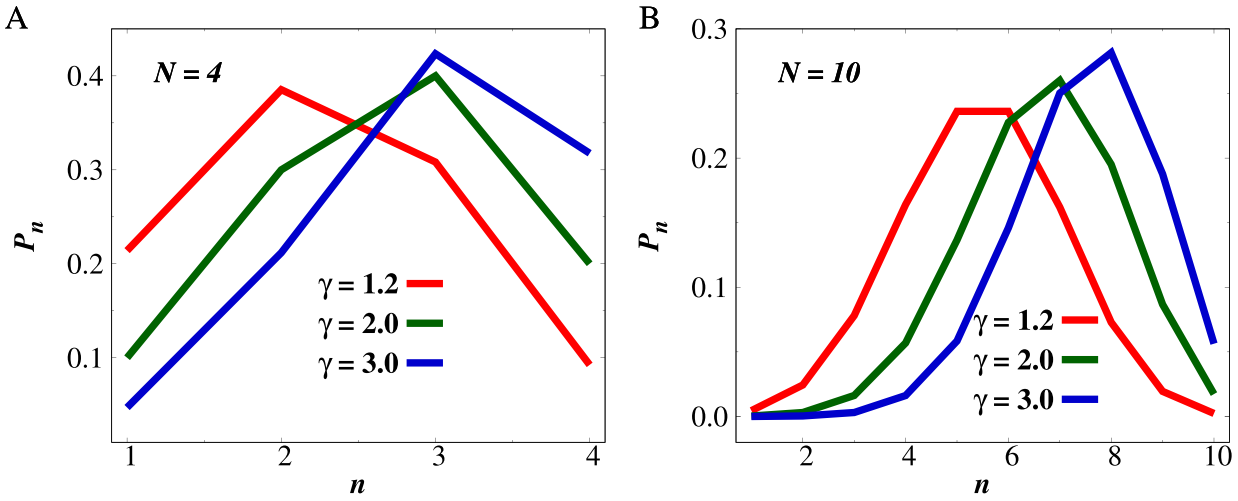


Figure 2. Stationary distributions of finding n TFs bound to the DNA promoter region for (A) $N = 4$ and (B) $N = 10$ binding sites during the active period of transcription for different values of the equilibrium binding constant γ.

Let us introduce a theoretical model to study the coupling between TFs binding and transcriptional activity as shown in Figure 1A. In this model, the continuous exchange and the association/dissociation of TFs to/from multiple binding sites determine the ON and OFF periods of bursting dynamics. Suppose there are N binding sites available for the association of TFs at the promoter region. When all TF binding sites are unoccupied, the corresponding state is denoted by 0, which also corresponds to the OFF period during burst (see Figure 1B), i.e., this is a transcriptional inactivity. However, the bursting takes place (in ON state) when the system is in states 1, 2, 3, ..., N, where each discrete state n corresponds to n sites occupied by TFs. This is the transcriptional active phase. It is assumed that initially, the beginning of transcriptional burst, (at $t = 0$) the system starts at the state 1 and TFs can sequentially bind to (or unbind from) their binding sites. The association rate constant of each TF is denoted by k_{on}, whereas the dissociation rate constant is k_{off}. This means that from the state n the overall rate of adding TFs is $(N - n)k_{on}$ because there are $(N - n)$ available sites, and the overall rate of dissociating TF is nk_{off} because there are n bound proteins. The corresponding discrete-state stochastic model is presented in Figure 1B. To couple the TF binding and transcriptional bursting, we assume that the bursting starts in state 1 and ends upon exiting again from the state 1 to the state 0, and the rate for this process is equal to k_{off}. One should also notice that our minimalist theoretical model is closely related to the well-known Langmuir adsorption model when in our approach DNA plays the role of the solid surface where molecules are adsorbed.

To quantify the transcription bursting dynamics, one can introduce a function $P_n(t)$ which is defined as the probability of having n bound TFs at time t. Forward master equations can describe the temporal evolution of these occupation probabilities

$$\frac{dP_1(t)}{dt} = 2k_{off}P_2(t) - [k_{off} + (N - 1)k_{on}]P_1(t) \quad (1)$$

for $n = 1$;

$$\frac{dP_n(t)}{dt} = (N - n + 1)k_{on}P_{n-1}(t) + (n + 1)k_{off}P_{n+1}(t) - [nk_{off} + (N - n)k_{on}]P_n(t) \quad (2)$$

for $1 < n < N$; and

$$\frac{dP_N(t)}{dt} = k_{on}P_{N-1}(t) - Nk_{off}P_N(t) \quad (3)$$

for $n = N$. In addition, there is a normalization condition, $\sum_{n=1}^{N} P_n(t)=1$. Assuming that the system quickly reaches steady-state conditions $[dP_n(t)/dt = 0]$, the stationary occupation probabilities P_n can be explicitly evaluated. This is because the number of states is finite and the system can be viewed as effectively in equilibrium. This leads to

$$P_n = \frac{N!}{n!(N - n)!} \cdot \frac{\gamma^n}{(1 + \gamma)^N - 1} \quad (4)$$

where $\gamma = \frac{k_{on}}{k_{off}}$ is the equilibrium binding constant, which can also be viewed as a measure of the affinity between TFs and the binding sites. The larger γ, the stronger the tendency for the TF to bind to an empty site. Also, in eq 4, $\frac{N!}{n!\,(N-n)!}$ is a binomial coefficient that reflects the combinatorial possibilities of arranging n TFs among total N possible sites. This means that all binding sites are independent of each other. Thus, in contrast to ref 42, we do not assume the cooperativity in the system when the presence of already bound TFs stimulate the association of new species. This is a minimal theoretical model that takes into account the most relevant physical-chemical processes in the system.

The results of our calculations for distributions of probabilities of different occupations of binding sites by TFs are presented in Figure 2. One can see that the stationary occupation probabilities P_n are always nonmonotonic reaching a maximum at some intermediate values of n that depend on the value of the equilibrium binding constant γ. As expected, increasing the affinity of TFs association to the DNA promoter region shifts the peak in the stationary distribution to larger values of n, while for weak affinities the peak shifts to smaller values of n. Since these probabilities determine all dynamic properties in the system, the results in Figure 2 indicate that

https://doi.org/10.1021/acs.jpclett.4c02050
J. Phys. Chem. Lett. 2024, 15, 8781−8789

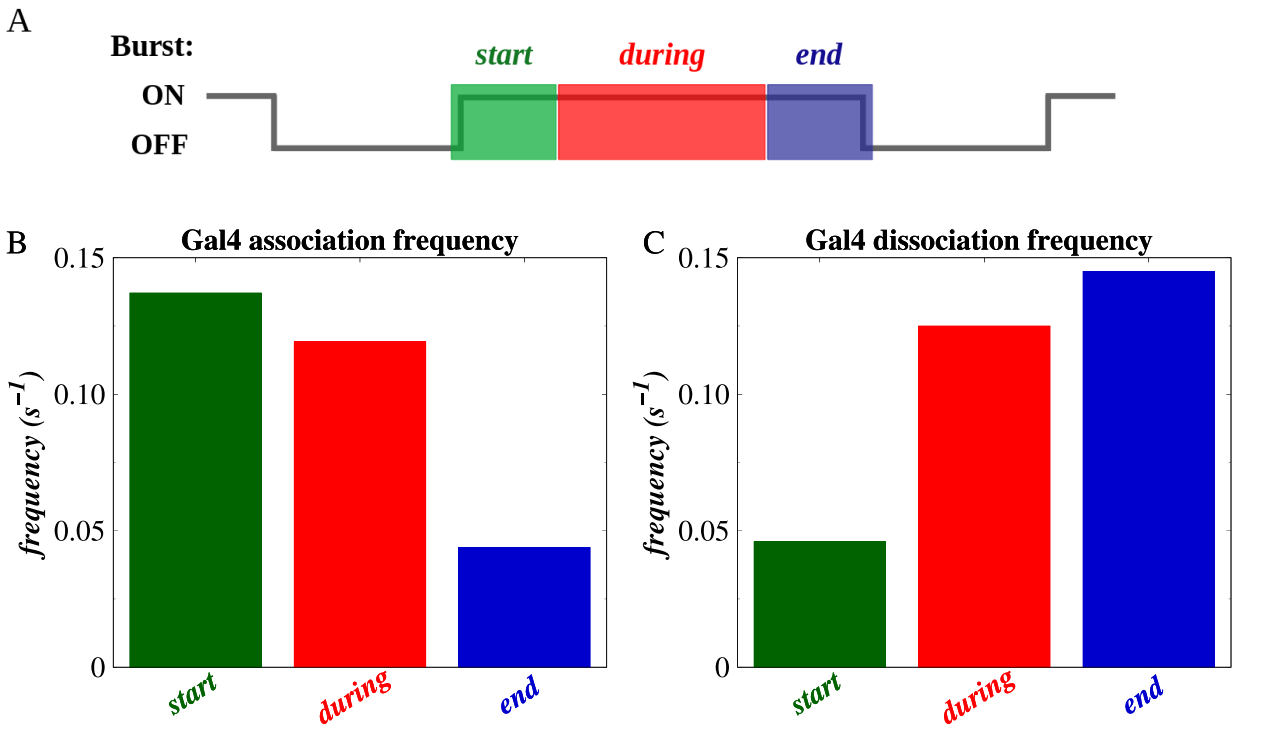


Figure 3. (A) Classification of the three periods: at the start, during, at the end of a burst. (B) Gal4 association frequency at GAL10 in these three periods. (C) Gal4 dissociation frequency at GAL10 in these three periods. The results are obtained from Monte Carlo computer simulations. The parameters used are $N = 4$, $k_{off} = 0.04\ s^{-1}$ and $\gamma = 1.2$.

transcription bursting strongly depends on the number of binding sites and the TF's affinity to bind them.

Our theoretical approach allows us to analyze quantitatively recent experiments that measured the binding of the TFs Gal4 and transcription of its target gene GAL10 in living yeast cells.[42] In these experiments, GAL10 promoter region that contains 4 Gal4 binding sites (i.e., $N = 4$) has been considered. The association of Gal4 TFs to these binding sites was monitored using novel 3D tracking imaging techniques, while simultaneously measurements of GAL10 transcription activity took place.[42] By quantifying the Gal4 intensity traces with a hidden Markov model, it was determined that TF binding quickly activates transcription and the burst ends as soon as all proteins dissociate from their binding sites. It was also observed that there are two binding populations with different residence times. One population is responsible for approximately two-thirds of the Gal4 binding events that lasted on average $\tau_{1,exp} \simeq 25 \pm 1$ s. In contrast, the second population corresponds to one-third of the binding events that lasted on average $\tau_{2,exp} \simeq 75 \pm 1$ s.[42] In addition, it was found that transcription burst duration on average is $T_{burst} \simeq 121 \pm 1$ s.

Our theoretical method can quantitatively explain these experimental observations. Since there are four binding sites in the experimentally studied system, we adopt the $N = 4$ model. Then, from eq 4 one can estimate the probabilities of different states with different numbers of bound TFs,

$$P_1 = \frac{1}{1 + 1.5\gamma + \gamma^2 + 0.25\gamma^3}, \quad P_2 = 1.5\gamma P_1,$$
$$P_3 = \gamma^2 P_1, \quad P_4 = 0.25\gamma^3 P_1 \tag{5}$$

Since the transcription burst ends when the last TF dissociates (transition 1 → 0 in Figure 1B), the average bursting time can be evaluated as

$$T_{burst} = \frac{1}{k_{off} P_1} = \frac{1 + 1.5\gamma + \gamma^2 + 0.25\gamma^3}{k_{off}} \tag{6}$$

This result can be easily explained. While the transcription activity starts when the first TF associates with the promoter region (state 1), the system might explore other states with a larger number of bound TFs (states 2, 3, and 4) before returning to the states 1 and exiting from the promoter region (transition 1 → 0 in Figure 1B). Due to the effective equilibrium between these discrete states, the average bursting time is longer, as reflected in the numerator by the terms that depend on γ in eq 6.

Fitting experimental data[42] using eqs 5 and (6), we estimate that $k_{off} \simeq 0.04\ s^{-1}$ and $\gamma \simeq 1.2$ for Gal4 TF binding. This leads to the following values of the occupational probabilities, $P_1 \simeq 0.214$, $P_2 \simeq 0.386$, $P_3 \simeq 0.308$, and $P_4 \simeq 0.092$. The existence of different states with variable numbers of bound TFs can explain the existence of different binding populations since these measurements might correspond to starting in different states of the system. Although, in principle, four different binding populations are expected, it is reasonable to assume that at given experimental conditions probably only the two fastest populations (starting from the state 1 or from the state 2) can be observed. The residence time before the dissociation starting from the state 1 is just $\tau_{1,theor} = 1/k_{off}$, yielding 25 s, which perfectly agrees with the experimentally observed time $\tau_{1,exp}$. To estimate the residence time starting from the state 2, we notice that these events correspond to the pathway 2 → 1 → 0, for which the corresponding flux can be written as

$$J_{2\to 1\to 0} = \frac{2k_{off}^2}{3k_{off} + 3k_{on}} = \frac{2k_{off}}{3(1 + \gamma)} \tag{7}$$

producing

https://doi.org/10.1021/acs.jpclett.4c02050
J. Phys. Chem. Lett. 2024, 15, 8781−8789

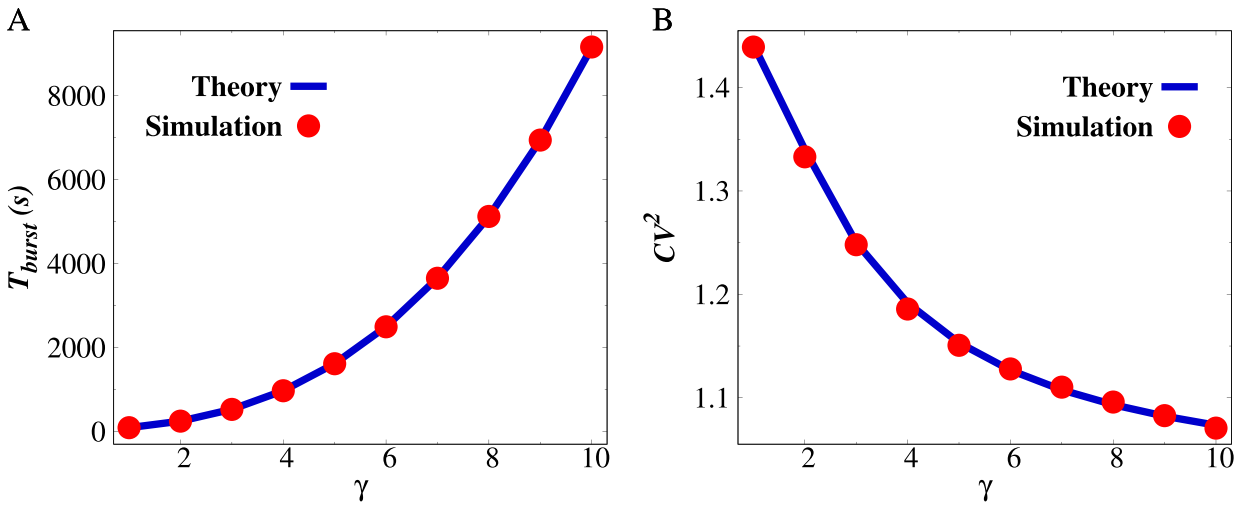


Figure 4. (A) Mean bursting time as a function of the equilibrium constant γ. (B) Temporal noise in transcriptional bursting as a function of γ. Analytical results are represented by blue lines, and red symbols are from Monte Carlo simulations. The following parameters are used in calculations: $N = 4$ and $k_{off} = 0.04\ s^{-1}$.

$$\tau_{2,theor} = \frac{1}{J_{2\to 1\to 0}} = \frac{3(1+\gamma)}{2k_{off}} \tag{8}$$

Substituting into this expression the experimentally fitted values of k_{off} and γ we obtain $\tau_{2,theor} \simeq 82$ s, which is only slightly larger than the experimentally measured time. Thus, our theoretical method provides a quantitative explanation of the existence of multiple binding populations.

Experiments[42] also reported that the frequency of association events was higher at the beginning of the transcriptional burst than at the end of this phase, while the frequency of dissociation events was higher at the end of the transcriptional burst than at the beginning of this phase. These observations can be easily explained using our theoretical picture. Since the start of the transcriptional burst can be associated with the state 1 (see Figure 1B) where the overall association rate is the highest $[(N-1)k_{on}]$, while the overall dissociation rate is the lowest (k_{off}). Later in the transcription activity phase, the situation is completely reversed: the system is more likely to be found in the states $n > 1$, where the dissociation rates are fast and association rates are slow.

However, to understand better the dependence of association and dissociation events at different stages of transcriptional bursting, we performed Monte Carlo computer simulations that mimic the experimental observations. In each simulation trajectory, we divided a long active period of bursting into three distinct segments: the start of a burst, during a burst, and the end of a burst (see Figure 3A). The first and last 30% of the binding events were designated as the start and end periods of a burst, respectively, while the remaining 40% of the events were considered the period during the burst. For each of these segments, the association and dissociation frequencies, defined as the number of TF association or dissociation events divided by the duration of each period, were evaluated from the computer simulations. The results of our computer simulations are presented in Figure 3. As one can see, the frequency of binding events decreases during the burst, while the frequency of unbinding events increases. This fully agrees with experimental observations. These findings show that our theoretical method captures main features of the processes that couple transcriptional activities with associations of TFs. In addition, it does not require for the TFs association process to be cooperative.

The advantage of our theoretical method is that it can be used to obtain more detailed molecular information on how TFs bindings stimulate transcriptional bursting. For this purpose, we introduce a method of first-passage probabilities that has been successfully explored before for clarifying dynamic aspects of bursting phenomena.[43,44] One can define $F_n(t)$ as the first-passage probability of transcriptional bursting ending for the first time by exiting from the state 1 at time t if the system started at $t = 0$ in the state n (see Figure 1B). Then the temporal evolution of these first-passage probabilities is governed by a set of backward master equations,[43]

$$\frac{dF_n(t)}{dt} = (N-n)k_{on}F_{n+1}(t) + nk_{off}F_{n-1}(t) - [(N-n)k_{on} + nk_{off}]F_n(t) \tag{9}$$

for $1 < n < N$, and

$$\frac{dF_1(t)}{dt} = (N-1)k_{on}F_2(t) + k_{off}F_{off}(t) - [(N-1)k_{on} + k_{off}]F_1(t) \tag{10}$$

$$\frac{dF_N(t)}{dt} = Nk_{off}F_{N-1}(t) - Nk_{off}F_N(t) \tag{11}$$

where $F_{off}(t)$ is the probability of being found in the OFF state immediately after exiting state 1. So, it is natural to assume that $F_{off}(t) = \delta(t)$. This physically means that if the system is in this state at $t = 0$, the process is immediately accomplished.

By modifying the backward master equations using Laplace transformations $\left(\tilde{F}_n(s) = \int_0^\infty e^{-st}F_n(t)dt\right)$, they can be exactly solved for all ranges of parameters. The detailed calculation is provided in the Supporting Information. From explicit expressions for first-passage probabilities, one can obtain all relevant dynamic properties of the system. We are specifically interested in two major quantities. The first one is the mean first-passage time (MFPT) or mean bursting time, which is defined as the average time it takes for the system to reach the

https://doi.org/10.1021/acs.jpclett.4c02050
J. Phys. Chem. Lett. 2024, 15, 8781−8789

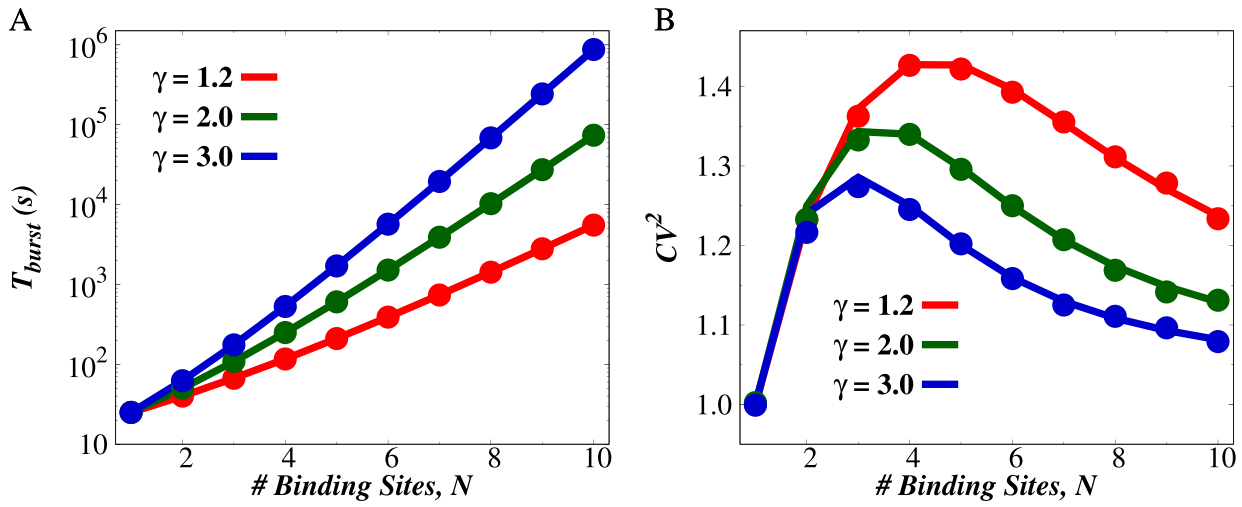


Figure 5. (A) Average bursting times as a function of the number of binding sites N for different equilibrium constants γ. (B) Temporal noise in transcriptional bursting as a function of N for different equilibrium constants γ. Analytical results are represented by solid lines, and symbols are from Monte Carlo simulations. Other parameter used for the calculations is $k_{off} = 0.04\ s^{-1}$.

OFF state (where no TFs are bound) from the initial ON state (where at least one TF is bound) for the first time. This quantifies the duration of a single transcriptional burst, which is the period of transcriptional activity. The second one is the coefficient of variation or noise in the system, which quantifies the extent of stochasticity in the bursting dynamics. These quantities are important as they can be measured in experiments and also provide crucial quantitative insights for clarifying the molecular mechanisms of bursting dynamics.

One can explicitly calculate the average burst times, as shown in the Supporting Information,

$$T_{burst} = \frac{1}{k_{off}}\left[1 + \sum_{n=2}^{N} \frac{(N-1)!}{n!(N-n)!}\gamma^{n-1}\right] \tag{12}$$

This expression has a clear physical meaning. If there would be only one binding site, then the overall time before the TF dissociation would be $1/k_{off}$ [see the first term in the brackets in eq 12]. However, for $N > 1$, there are multiple possibilities of TFs to be bound to the promoter region, and this is reflected by the second term in brackets in eq 12. Note that for $N = 4$ this result reduces to the estimate of bursting time that we already obtained in eq 6. Similarly, one can also estimate the coefficient of variation (CV) or temporal noise in the system, as given by

$$CV^2 = \frac{\langle T_{burst}^2\rangle - \langle T_{burst}\rangle^2}{\langle T_{burst}\rangle^2} \tag{13}$$

where $\langle T_{burst}^2\rangle$ is the mean-squared bursting time, and $\langle T_{burst}\rangle = T_{burst}$ is the mean bursting time: see the Supporting Information for more details.

Explicit expressions for dynamic properties of the system allow us to analyze in more details the molecular mechanisms that couple TFs exchange and transcriptional bursting. Figure 4 shows the results of our calculations for mean bursting times and the coefficient of variation as a function of the equilibrium constant γ. This quantity reflects the interaction between transcription factor and the corresponding binding site on DNA with larger γ corresponding to stronger interactions. As one can see (Figure 4A), increasing the attraction of TFs to DNA sites makes the average bursting times longer. This is an expected result since TFs can now spend more time on DNA, prolonging the transcriptional activities. At the same time, the noise in the transcription bursting times decreases with increasing γ (see Figure 4B). It is interesting to note that the Gal4 system corresponds to $\gamma \simeq 1.2$ when the bursting times are not too long and not too short either ($T_{burst} \simeq 121$ s) and the variations in times are significant ($CV^2 = 1.42$).

Our theoretical method can also be used to analyze the role of multiple binding sites on the dynamics of transcriptional bursting, as illustrated in Figure 5. As expected, the average bursting time increases for a larger number of binding sites N, and the effect is stronger for larger equilibrium constants γ: see Figure 5A. This is because it will take a much longer time for all TFs to dissociate to end the current cycle of transcriptional activities. Surprisingly, the dependence of the normalized variance on N is nonmonotonic (Figure 5B). It seems that noise is maximal for intermediate values of N. Interestingly, the experimental system Gal4 corresponds to the conditions with the maximal distribution in bursting times ($N = 4$). Another important observation is that increasing γ shifts the peak in the variance to smaller values of N.

Our theoretical method suggests that the number of binding sites and the strength of interactions between TFs and promoter regions are important parameters that influence the dynamics of transcription. Modifying these quantities might lead to significant variations in transcriptional bursting, indicating that they can be utilized for efficient genetic regulation. Applying this analysis to experimentally studied Gal4 system suggests that biological systems might be optimized concerning these properties in the following way. The affinity of TFs to their binding sites is tuned to achieve transcriptional bursting times that are not too short and not too long. Otherwise, there will be not enough time to reliably synthesize mRNA molecules if T_{burst} are too small. Very large continuous bursting times are also not good, probably because of the oversupply of mRNA molecules. Since the rate of translation processes is limited by the number of ribosomes where proteins are synthesized, adding more mRNA molecules might be energetically wasteful and unproductive.

Our analysis also suggests that the experimental transcription is taking place for largest noise, or widest distribution

https://doi.org/10.1021/acs.jpclett.4c02050
J. Phys. Chem. Lett. 2024, 15, 8781−8789

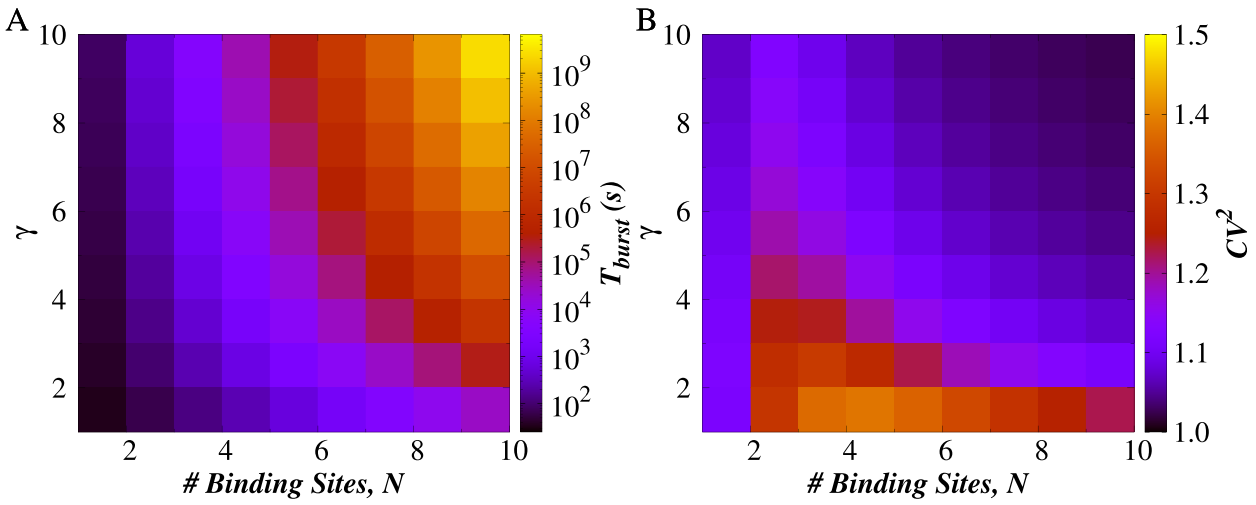


Figure 6. 2D contour plot of (A) average burst times and (B) temporal noise in transcriptional bursting as a function of the equilibrium constant γ and the number of binding sites N. Other parameter used for the calculations is k_{off} = 0.04 s^{-1}.

of bursting times. This unexpected result can be explained in the following way. For experimental equilibrium constant $\gamma \simeq 1.2$ the most probable situation is with two or three bound TFs: see Figure 2A. Then, to end the transcription activities, the system will have multiple pathways to exit. For example, the following pathways $2 \rightarrow 1 \rightarrow 0$, $2 \rightarrow 3 \rightarrow 2 \rightarrow 1 \rightarrow 0$ are possible. But the possibility of so many pathways will produce very different bursting times, leading to a large noise in the system. Another possible reason to have a wide distribution of bursting times is that it might be also beneficiary for the system since it can now easily adjust to changes in external conditions, making transcription process more flexible and robust. It will be interesting to explore this aspect of the coupling between TFs association and transcriptional activities in more detail.

Our theoretical calculations found that two properties, number of binding sites and the strength of TFs interactions with the promoter region, are critically important for regulating transcription processes. However, these features are not independent from each other as one can see from Figure 6. While increasing both equilibrium constant γ and the number of sites N always leads to very long bursting times (Figure 6A), the effect on the distribution of these times is more complex (Figure 6B). This observation allows us to answer the question on why different transcription systems might have different numbers of binding sites. It seems that for every specific value of γ there is an optimal value of N. If the affinity of TFs to their binding sites is large, then it is enough to have only a single binding site to achieve the most efficient transcription. However, if the affinity is not too strong then having multiple binding sites allows the system to support transcription more effectively.

To summarize, we developed a theoretical method to understand the coupling between TFs association to DNA promoter's region and transcriptional activities. It allowed us to clarify the correlations between TFs exchange and transcription bursting at the molecular level. It is argued that the burst lasts while TFs are associated with the promoter region, and it ends as soon as all TFs depart from DNA. Based on these arguments, a discrete-state stochastic model is proposed and successfully applied to analyze experimental observations on Gal4 system in yeast. It is found that two features, affinity, and number of binding sites, are critically important in regulating transcriptional activities. Our explicit analysis, supported by extensive Monte Carlo computer simulations, indicates that experimental systems are optimized to have bursting times not too short and not too long, while the variations in the bursting dynamics are predicted to be maximal. It is argued that this is the result of having multiple bound TFs that lead to multiple pathways to dissociation. It is also suggested that this might be beneficial for biological systems allowing them to be more flexible and robust with respect to variations in external conditions. In addition, our theoretical method allows us to explain the existence of transcriptional systems with variable affinities and variable numbers of binding sites. The inverse correlations between the affinities and the number of binding sites are predicted.

While our theoretical framework presents a plausible microscopic picture of the coupling of TFs associations to the promoter regions and transcriptional activities, it is important to discuss its limitations. Our theoretical model is rather oversimplified, and it may not fully encompass the complexities of *in vivo* transcriptional regulation. For instance, factors such as chromatin state, dynamics of transcriptional machinery, and variations in cellular environments are not explicitly incorporated. Furthermore, the model assumes idealized conditions of TFs binding and dissociation kinetics, neglecting potential complexities such as cooperative binding or heterogeneous TF interactions. Additionally, the model assumes that the system quickly reaches stationary conditions, which might not fully agree with the dynamics in real biological systems. However, despite these limitations, our theoretical approach offers a coherent physical-chemical framework for understanding the coupling between TF binding and transcriptional burst dynamics. It fully explains all existing experiments, and it provides specific quantitative predictions that can guide future experimental designs aimed at uncovering detailed microscopic mechanisms of underlying cellular processes. Thus, while there is room for refinement and expansion, our theoretical framework serves as a valuable tool for both experimental validation and more advanced theoretical investigations.

https://doi.org/10.1021/acs.jpclett.4c02050
J. Phys. Chem. Lett. 2024, 15, 8781−8789

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jpclett.4c02050.

Details of analytical calculations for the stationary probabilities and for the first-passage processes to obtain mean burst times and noise in transcriptional bursting (PDF)

AUTHOR INFORMATION

Corresponding Author

Anatoly B. Kolomeisky − *Center for Theoretical Biological Physics, Rice University, Houston, Texas 77005, United States; Department of Chemistry and Department of Chemical and Biomolecular Engineering, Rice University, Houston, Texas 77005, United States*; orcid.org/0000-0001-5677-6690; Email: tolya@rice.edu

Author

Anupam Mondal − *Center for Theoretical Biological Physics, Rice University, Houston, Texas 77005, United States; Department of Chemistry, Rice University, Houston, Texas 77005, United States*; orcid.org/0000-0002-8436-5618

Complete contact information is available at:
https://pubs.acs.org/10.1021/acs.jpclett.4c02050

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The work was supported by the Welch Foundation (C-1559), by the NIH (R01GM148537 and R01HL157714-04), and by the Center for Theoretical Biological Physics sponsored by the NSF (PHY-2019745).

REFERENCES

(1) Inukai, S.; Kock, K. H.; Bulyk, M. L. Transcription factor-DNA binding: beyond binding site motifs. *Curr. Opin Genet Dev* **2017**, *43*, 110−119.

(2) Cramer, P. Organization and regulation of gene transcription. *Nature* **2019**, *573*, 45−54.

(3) Ptashne, M.; Gann, A. Transcriptional activation by recruitment. *Nature* **1997**, *386*, 569−577.

(4) Fuda, N. J.; Ardehali, M. B.; Lis, J. T. Defining mechanisms that regulate RNA polymerase II transcription in vivo. *Nature* **2009**, *461*, 186−192.

(5) Rodriguez, J.; Larson, D. R. Transcription in living cells: Molecular mechanisms of bursting. *Annu. Rev. Biochem.* **2020**, *89*, 189−212.

(6) Tunnacliffe, E.; Chubb, J. R. What is a transcriptional burst? *Trends Genet* **2020**, *36*, 288−297.

(7) Brouwer, I.; Lenstra, T. L. Visualizing transcription: key to understanding gene expression dynamics. *Curr. Opin Chem. Biol.* **2019**, *51*, 122−129.

(8) Lenstra, T. L.; Rodriguez, J.; Chen, H.; Larson, D. R. Transcription dynamics in living cells. *Annu. Rev. Biophys* **2016**, *45*, 25−47.

(9) Dar, R. D.; Razooky, B. S.; Singh, A.; Trimeloni, T. V.; McCollum, J. M.; Cox, C. D.; Simpson, M. L.; Weinberger, L. S. Transcriptional burst frequency and burst size are equally modulated across the human genome. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 17454−17459.

(10) Larsson, A. J. M.; Johnsson, P.; Hagemann-Jensen, M.; Hartmanis, L.; Faridani, O. R.; Reinius, B.; Segerstolpe, A.; Rivera, C. M.; Ren, B.; Sandberg, R. Genomic encoding of transcriptional burst kinetics. *Nature* **2019**, *565*, 251−254.

(11) Chong, S.; Chen, C.; Ge, H.; Xie, X. S. Mechanism of transcriptional bursting in bacteria. *Cell* **2014**, *158*, 314−326.

(12) Golding, I.; Paulsson, J.; Zawilski, S. M.; Cox, E. C. Real-time kinetics of gene activity in individual bacteria. *Cell* **2005**, *123*, 1025−1036.

(13) So, L. H.; Ghosh, A.; Zong, C.; Sepúlveda, L. A.; Segev, R.; Golding, I. General properties of transcriptional time series in Escherichia coli. *Nat. Genet.* **2011**, *43*, 554−560.

(14) Larson, D. R.; Zenklusen, D.; Wu, B.; Chao, J. A.; Singer, R. H. Real-time observation of transcription initiation and elongation on an endogenous yeast gene. *Science* **2011**, *332*, 475−478.

(15) Zenklusen, D.; Larson, D. R.; Singer, R. H. Single-RNA counting reveals alternative modes of gene expression in yeast. *Nat. Struct Mol. Biol.* **2008**, *15*, 1263−1271.

(16) Suter, D. M.; Molina, N.; Gatfield, D.; Schneider, K.; Schibler, U.; Naef, F. Mammalian genes are transcribed with widely different bursting kinetics. *Science* **2011**, *332*, 472−474.

(17) Raj, A.; Peskin, C. S.; Tranchina, D.; Vargas, D. Y.; Tyagi, S. Stochastic mRNA synthesis in mammalian cells. *PLoS Biol.* **2006**, *4*, No. e309.

(18) Wan, Y.; Anastasakis, D. G.; Rodriguez, J.; Palangat, M.; Gudla, P.; Zaki, G.; Tandon, M.; Pegoraro, G.; Chow, C. C.; Hafner, M.; Larson, D. R. Dynamic imaging of nascent RNA reveals general principles of transcription dynamics and stochastic splice site selection. *Cell* **2021**, *184*, 2878−2895.

(19) Chen, H.; Levo, M.; Barinov, L.; Fujioka, M.; Jaynes, J. B.; Gregor, T. Dynamic interplay between enhancer-promoter topology and gene activity. *Nat. Genet.* **2018**, *50*, 1296−1303.

(20) Luo, S.; Wang, Z.; Zhang, Z.; Zhou, T.; Zhang, J. Genome-wide inference reveals that feedback regulations constrain promoter-dependent transcriptional burst kinetics. *Nucleic Acids Res.* **2023**, *51*, 68−83.

(21) Nicolas, D.; Zoller, B.; Suter, D. M.; Naef, F. Modulation of transcriptional burst frequency by histone acetylation. *Proc. Natl. Acad. Sci. U. S. A.* **2018**, *115*, 7153−7158.

(22) Donovan, B. T.; Huynh, A.; Ball, D. A.; Patel, H. P.; Poirier, M. G.; Larson, D. R.; Ferguson, M. L.; Lenstra, T. L. Live-cell imaging reveals the interplay between transcription factors, nucleosomes, and bursting. *EMBO J.* **2019**, *38*, No. e100809.

(23) Cho, W. K.; Jayanth, N.; English, B. P.; Inoue, T.; Andrews, J. O.; Conway, W.; Grimm, J. B.; Spille, J. H.; Lavis, L. D.; Lionnet, T.; Cisse, I. I. RNA Polymerase II cluster dynamics predict mRNA output in living cells. *Elife* **2016**, *5*, No. e13617.

(24) Kawasaki, K.; Fukaya, T. Functional coordination between transcription factor clustering and gene activity. *Mol. Cell* **2023**, *83*, 1605−1622.

(25) Patel, H. P.; Coppola, S.; Pomp, W.; Aiello, U.; Brouwer, I.; Libri, D.; Lenstra, T. L. DNA supercoiling restricts the transcriptional bursting of neighboring eukaryotic genes. *Mol. Cell* **2023**, *83*, 1573−1587.

(26) Mondal, A.; Sangeeta; Bhattacherjee, A. Torsional behaviour of supercoiled DNA regulates recognition of architectural protein Fis on minicircle DNA. *Nucleic Acids Res.* **2022**, *50*, 6671−6686.

(27) Senecal, A.; Munsky, B.; Proux, F.; Ly, N.; Braye, F. E.; Zimmer, C.; Mueller, F.; Darzacq, X. Transcription factors modulate c-Fos transcriptional bursts. *Cell Rep* **2014**, *8*, 75−83.

(28) Loffreda, A.; Jacchetti, E.; Antunes, S.; Rainone, P.; Daniele, T.; Morisaki, T.; Bianchi, M. E.; Tacchetti, C.; Mazza, D. Live-cell p53 single-molecule binding is modulated by C-terminal acetylation and correlates with transcriptional activity. *Nat. Commun.* **2017**, *8*, 313.

(29) Clauß, K.; Popp, A. P.; Schulze, L.; Hettich, J.; Reisser, M.; Escoter Torres, L.; Uhlenhaut, N. H.; Gebhardt, J. C. M. DNA residence time is a regulatory factor of transcription repression. *Nucleic Acids Res.* **2017**, *45*, 11121−11130.

(30) Popp, A. P.; Hettich, J.; Gebhardt, J. C. M. Altering transcription factor binding reveals comprehensive transcriptional kinetics of a basic gene. *Nucleic Acids Res.* **2021**, *49*, 6249−6266.

(31) Stavreva, D. A.; et al. Transcriptional bursting and co-bursting regulation by steroid hormone release pattern and transcription factor mobility. *Mol. Cell* **2019**, *75*, 1161−1177.
(32) Trojanowski, J.; Frank, L.; Rademacher, A.; Mücke, N.; Grigaitis, P.; Rippe, K. Transcription activation is enhanced by multivalent interactions independent of phase separation. *Mol. Cell* **2022**, *82*, 1878−1893.
(33) Mehta, G. D.; Ball, D. A.; Eriksson, P. R.; Chereji, R. V.; Clark, D. J.; McNally, J. G.; Karpova, T. S. Single-molecule analysis reveals linked cycles of RSC chromatin remodeling and Ace1p transcription factor binding in Yeast. *Mol. Cell* **2018**, *72*, 875−887.
(34) Phair, R. D.; Misteli, T. High mobility of proteins in the mammalian cell nucleus. *Nature* **2000**, *404*, 604−609.
(35) McNally, J. G.; Müller, W. G.; Walker, D.; Wolford, R.; Hager, G. L. The glucocorticoid receptor: rapid exchange with regulatory sites in living cells. *Science* **2000**, *287*, 1262−1265.
(36) Gebhardt, J. C.; Suter, D. M.; Roy, R.; Zhao, Z. W.; Chapman, A. R.; Basu, S.; Maniatis, T.; Xie, X. S. Single-molecule imaging of transcription factor binding to DNA in live mammalian cells. *Nat. Methods* **2013**, *10*, 421−426.
(37) Ochiai, H.; Hayashi, T.; Umeda, M.; Yoshimura, M.; Harada, A.; Shimizu, Y.; Nakano, K.; Saitoh, N.; Liu, Z.; Yamamoto, T.; Okamura, T.; Ohkawa, Y.; Kimura, H.; Nikaido, I. Genome-wide kinetic properties of transcriptional bursting in mouse embryonic stem cells. *Sci. Adv.* **2020**, *6*, No. eaaz6699.
(38) Mazzocca, M.; Colombo, E.; Callegari, A.; Mazza, D. Transcription factor binding kinetics and transcriptional bursting: What do we really know? *Curr. Opin Struct Biol.* **2021**, *71*, 239−248.
(39) Lammers, N. C.; Kim, Y. J.; Zhao, J.; Garcia, H. G. A matter of time: Using dynamics and theory to uncover mechanisms of transcriptional bursting. *Curr. Opin Cell Biol.* **2020**, *67*, 147−157.
(40) de Jonge, W. J.; Patel, H. P.; Meeussen, J. V. W.; Lenstra, T. L. Following the tracks: How transcription factor binding dynamics control transcription. *Biophys. J.* **2022**, *121*, 1583−1592.
(41) Meeussen, J. V. W.; Lenstra, T. L. Time will tell: comparing timescales to gain insight into transcriptional bursting. *Trends Genet* **2024**, *40*, 160−174.
(42) Pomp, W.; Meeussen, J. V. W.; Lenstra, T. L. Transcription factor exchange enables prolonged transcriptional bursts. *Mol. Cell* **2024**, *84*, 1036−1048.
(43) Klindziuk, A.; Meadowcroft, B.; Kolomeisky, A. B. A mechanochemical model of transcriptional bursting. *Biophys. J.* **2020**, *118*, 1213−1220.
(44) Klindziuk, A.; Kolomeisky, A. B. Understanding the molecular mechanisms of transcriptional bursting. *Phys. Chem. Chem. Phys.* **2021**, *23*, 21399−21406.

https://doi.org/10.1021/acs.jpclett.4c02050
J. Phys. Chem. Lett. 2024, 15, 8781−8789