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

pubs.acs.org/JPCB

Article

Long-Range Supercoiling-Mediated RNA Polymerase Cooperation in Transcription

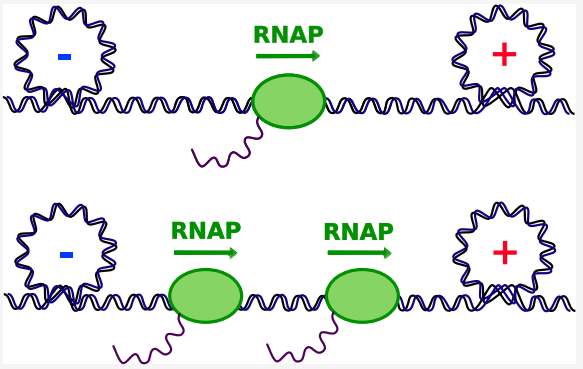
Alena Klindziuk and Anatoly B. Kolomeisky*

Cite This: *J. Phys. Chem. B* 2021, 125, 4692–4700

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: It is widely believed that DNA supercoiling plays an important role in the regulation of transcriptional dynamics. Recent studies show that it could affect transcription not only through the buildup and relaxation of torsional strain on DNA strands but also via effective long-range supercoiling-mediated interactions between RNA polymerase (RNAP) molecules. Here, we present a theoretical study that quantitatively analyzes the effect of long-range RNAP cooperation in transcription dynamics. Our minimal chemical-kinetic model assumes that one or two RNAP molecules can simultaneously participate in the transcription, and it takes into account their binding to and dissociation from DNA. It also explicitly accounts for competition between the supercoiling buildup that reduces the RNA elongation speed and gyrase binding that rescues the RNA synthesis. The full analytical solution of the model accompanied by Monte Carlo computer simulations predicts that the system should exhibit transcriptional bursting dynamics, in agreement with experimental observations. The analysis also revealed that when there are two polymerases participating in the elongation rather than one, the transcription process becomes much more efficient since the level of stochastic noise decreases while more RNA transcripts are produced. Our theoretical investigation clarifies molecular aspects of the supercoiling-mediated RNAP cooperativity during transcription.

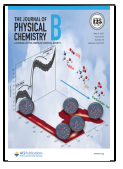


INTRODUCTION

Transcription is the first step in a fundamental process of the transfer of genetic information where specific enzymatic molecules, known as RNA polymerases (RNAPs), copy segments of DNA, transforming them into messenger RNA transcripts.[1−3] Because of its critical importance for all living organisms, this process has been extensively studied using multiple biochemical and biophysical methods.[4−9] Recent experimental advances allowed researchers to monitor transcription processes both *in vivo* and *in vitro*, with unprecedented spatial and temporal resolutions.[6−8,10] These studies revealed that in bacteria, highly expressed genes are transcribed in stochastic bursts of activity,[9−12] which are called transcriptional bursting. It has been suggested that this phenomenon is the main source of gene expression variability and it might be an important contributing factor to phenotype diversity, helping organisms to adapt and survive under variable environmental conditions.[13−15]

The molecular mechanisms responsible for the transcriptional bursting remain unclear. Most likely, it is a result of multiple independent factors.[8,9,15] At the same time, there are a large number of experimental pieces of evidence pointing to the importance of the mechanics of chromosomal DNA and its influence on transcription regulation that might lead to the bursting phenomenon.[16−21] Specifically, *in vitro* single-molecule studies revealed that during the transcription elongation, as the RNAP moves along the DNA strand producing RNA copies of DNA, it applies a torque on DNA, generating twists known as DNA supercoiling.[16−18] The twisting of the DNA behind the RNAP is called negative supercoiling, while the twisting in front of the RNAP is known as positive supercoiling. These processes can be quantitatively described, for example, by the twin-supercoiled-domain model.[22] The accumulation of supercoiling can gradually become a mechanical barrier for the polymerase, eventually stalling it and hindering the elongation.[5,16,23] The stalled RNAP is rescued once the supercoiling is released either by the binding of topoisomerases[5,16,22,23] or via mechanical relaxation.[5] In particular, *in vitro* single-molecule studies in bacteria have shown that negative supercoiling is quickly released by an enzyme Topo I, a prokaryotic type IA topoisomerase, which is abundant under typical cellular conditions, while the positive supercoiling is released by a less-abundant type II topoisomerase molecule known as a gyrase. This leads to a buildup of the

Received: March 1, 2021
Revised: April 15, 2021
Published: April 29, 2021



https://doi.org/10.1021/acs.jpcb.1c01859
J. Phys. Chem. B 2021, 125, 4692–4700

positive supercoiling and gradual reduction in the RNA synthesis.[5,16,23] Note that the bacterial Topo IA should not be confused with the eukaryotic topoisomerase Topo IB which can release both positive and negative superhelical strains. The recent chemical-kinetic mechanochemical model has shown that such interplay between the mechanical process of supercoiling generation and chemical enzymatic process of supercoiling release can lead to noisy or "bursting" transcription dynamics.[24] Thus, the single-molecule studies and closely related theoretical investigations have elucidated the various aspects of how single RNAP participates in the elongation and influences the transcription dynamics.[6,9,16,19,24−26]

Under realistic cellular conditions, however, most genes are typically transcribed simultaneously by multiple RNAP molecules. Moreover, a recent *in vitro* study revealed that the amount of DNA twists generated by transcription of mere 150 bases, which is much less than a typical length of a gene, is enough to stall an RNAP elongation complex.[27] It is unclear then whether the tremendous amount of the torsional stress generated by many RNAPs transcribing DNA simultaneously could be resolved by Topo IA alone.[23] In other words, the supercoiling created by multiple RNAPs would prevent them from synthesizing the RNA molecules, but this does not happen in real cells. Therefore, it was suggested that there is another *in vivo* mechanism, such as the long-range RNAP cooperation, that likely plays a role in making sure that RNAPs can continue transcription despite the buildup of torsional stress.[20,21,23] In addition, a recent study found that in *Escherichia coli* bacteria, the transcriptional elongation by multiple RNAPs is even faster than that by a single RNAP and the elongation rate is independent of the RNAP density as long as two or more RNAPs are present.[20,21] These results suggest that polymerases probably can effectively interact over long distances, and transcription-induced DNA supercoiling might be one of the possible sources for such effective interactions. A possible mechanism of RNAP cooperation during transcription, where the negative supercoiling of the leading polymerase and the positive supercoiling of the trailing polymerase annihilate each other resulting in an efficient elongation by multiple polymerases, has been introduced, but only qualitative arguments were given.[23] Other ideas of potential cooperation mechanisms have been also presented.[28] However, the molecular mechanisms of cooperativity between RNAPs during the transcription remain not well understood.[21]

In this work, we developed a discrete-state stochastic model of RNAP cooperation in transcription, which is driven by the long-range supercoiling-mediated effective interactions. In our approach, transcription might happen when one or two RNAP molecules are present on DNA. The model assumes that upon encountering supercoiling, the leading polymerase relaxes the mechanical strain on the DNA behind itself such that the trailing polymerase can proceed unimpeded. It also includes the effects of RNAP binding and dissociation, RNA production and degradation, and the mechanochemical coupling between the supercoiling buildup and the synthesis of RNA transcripts. Our explicit analytical calculations, supported by Monte Carlo computer simulations, show that multiple RNAPs are more effective in producing RNAs than a single polymerase while at the same time lowering the level of stochastic fluctuations in the system. Thus, our theoretical model quantitatively explains the need for multiple RNAPs in the transcription processes.

THEORETICAL MODEL

To understand the microscopic mechanism of cooperativity between RNAPs, we develop a discrete-state stochastic model that incorporates the most relevant physicochemical processes that take place during transcription. They include synthesis and degradation of RNA transcripts, supercoiling buildup due to the activity of RNAPs that move in a helical fashion along the DNA chains, supercoiling release due to the enzymatic action of gyrases, and polymerase binding and dissociation events. In the model, we consider a situation when either one or two RNAPs transcribe the DNA, as illustrated in Figure 1. This

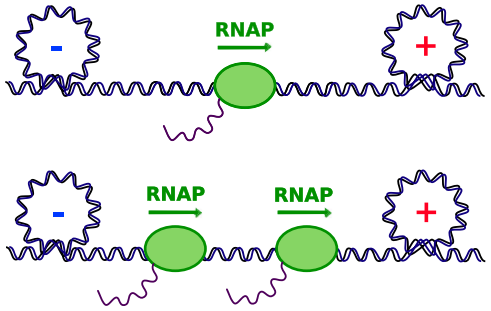


Figure 1. Schematic view of the model to evaluate the role of cooperativity between RNAPs during transcription. Either one or two RNAPs transcribe the DNA, while creating negative and positive supercoiling regions. The top figure shows an instance when a single polymerase is present on the DNA, while the bottom figure depicts the case when two polymerases are present.

allows us to explicitly quantify the effect of cooperativity when two polymerases are present in comparison with the single-RNAP transcription. It is assumed that Topo IA binds to the region of DNA upstream of the trailing RNAP molecule releasing the negative supercoiling, while the gyrase binds downstream of the leading RNAP releasing the positive supercoiling: see Figure 1. All other possible attachments of enzyme molecules to DNA are neglected in our theoretical approach. As a result of these assumptions, only the variation in effective RNA synthesis rate from positive supercoiling fluctuations is taken into account, while the synthesis rate-altering effects of the underlying negative superhelical density of DNA are not included in the model.

Physicochemical transformations during the transcription process lead to the existence of multiple states, and in Figure 2, we present a detailed chemical-kinetic scheme for the model. It has two planes of chemical states: a plane where there is one polymerase on DNA (1 RNAP plane) shown on the top of Figure 2 and a plane where there are two polymerases on DNA (2 RNAP plane) shown on the bottom of Figure 2. The transitions between the corresponding states on different planes can happen when the RNAP attaches to DNA with a rate r_{on} or dissociates from the strand with a rate r_{off}. The produced RNA molecule can be degraded with a rate β, which is the same in all states of the system. In the absence of DNA supercoiling, when the gyrase is bound, a synthesis rate for a single RNA by one polymerase is equal to α. However, when the gyrase is detached from DNA, each additional produced RNA leads to a buildup of DNA supercoiling in front of the RNAP, and this slows down the RNA synthesis rate by a factor $1/y^{j+1}$. Here, j is the number of RNA molecules produced after the gyrase detached from DNA, and $y = \exp(\varepsilon/k_BT)$ with ε being the additional energy needed to produce one transcript

https://doi.org/10.1021/acs.jpcb.1c01859
J. Phys. Chem. B 2021, 125, 4692−4700

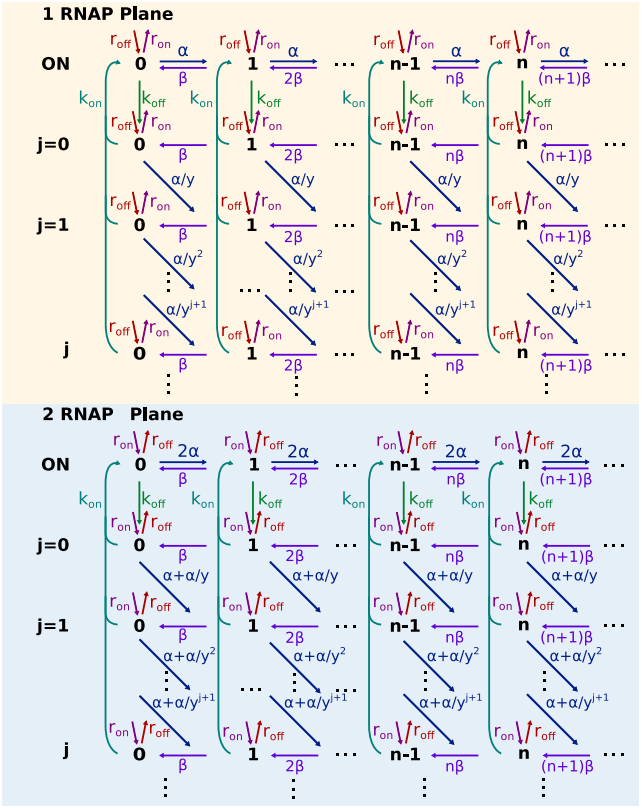


Figure 2. Chemical kinetic scheme for the model. The number of RNAs produced in the system is labeled as 0, 1, ... n, $n+1$, ... The RNA synthesis rate by one RNAP is α, the RNA degradation rate is β, the gyrase binding to DNA rate is k_{on}, the gyrase dissociation rate is k_{off}, the RNAP binding to DNA rate is r_{on}, the RNAP dissociation rate is r_{off}, and y is how much supercoiling influences the elongation rate. The model describes a scenario where either one or two RNAPs participate in the transcription. 1 RNAP plane represents the case when there is one polymerase, and 2 RNAP plane is the case when there are two polymerases producing RNA transcripts. The model has the following states: two ON states where the gyrase is bound—one in each plane, as well as two sets of j states where (j = 0, 1, ..., ∞) is the number of transcripts made after the gyrase detached—a set in each plane.

under supercoiling conditions. However, only the leading polymerase is slowed down due to the supercoiling. For the trailing RNAP, the positive buildup in supercoiling is canceled by the negative contribution from the leading RNA, and this polymerase can proceed then unhindered. Therefore, the trailing RNAP has the original RNA synthesis rate α. The gyrase binding rate is defined as k_{on}, and the gyrase detachment rate is defined as k_{off}. Thus, in the 1 RNAP plane, the active transcription ON states specify the situation when the gyrase is attached to DNA and a single RNAP transcribes with the rate α. When the gyrase detaches, the system moves into a series of j states with the reduced RNA production rate α/y^j [Figure 2 (top)]. In the 2 RNAP plane, the two polymerases synthesize RNAs with a combined rate of 2α when supercoiling is suppressed by gyrase (ON states). However, without the gyrase, the combined RNA production rate is $\alpha + \alpha/y^{j+1}$, where j is the number of RNAs produced since the dissociation of the gyrase [Figure 2 (bottom)].

An important advantage of our theoretical approach is that it explicitly takes into account the coupling between chemical processes of RNA synthesis and mechanical processes of supercoiling buildup in a thermodynamically consistent way. It is done by utilizing the parameter $y = \exp(\varepsilon/k_BT)$, which reduces the synthesis rate according to the degree of supercoiling that can be measured using the variable j. The energetic parameter ε describes an additional work that the RNAP molecule needs to do in order to catalyze the formation of an RNA molecule in the presence of mechanical stress on DNA. Because in the presence of supercoiling the RNA synthesis rate gets reduced by an exponentially growing factor, any number of RNA molecules can be produced in our model, but the probability of having many RNA transcripts exponentially decreases with the degree of the mechanical stress in the system. This eliminates the need for introducing an artificial cutoff on the number of synthesized RNA molecules.[19,24]

The dynamic processes in the system can be described by a system of master equations for each individual chemical state

$$\frac{dP_{on}(1;n,t)}{dt} = \alpha P_{on}(1;n-1,t) + (n+1)\beta P_{on}(1;n+1,t) + k_{on}\sum_{j=0}^{\infty} P_j(1;n,t) + r_{off}P_{on}(2;n,t) - (\alpha + n\beta + k_{off} + r_{on})P_{on}(1;n,t) \quad (1)$$

$$\frac{dP_0(1;n,t)}{dt} = (n+1)\beta P_0(1;n+1,t) + k_{off}P_{on}(1;n,t) + r_{off}P_0(2;n,t) - (\alpha/y + n\beta + k_{on} + r_{on})P_0(1;n,t) \quad (2)$$

$$\frac{dP_j(1;n,t)}{dt} = \alpha/y^j P_{j-1}(1;n-1,t) + (n+1)\beta P_j(1;n+1,t) + r_{off}P_j(2;n,t) - (\alpha/y^{j+1} + n\beta + k_{on} + r_{on})P_j(1;n,t) \quad (3)$$

$$\frac{dP_{on}(2;n,t)}{dt} = 2\alpha P_{on}(2;n-1,t) + (n+1)\beta P_{on}(2;n+1,t) + k_{on}\sum_{j=0}^{\infty} P_j(2;n,t) + r_{on}P_{on}(1;n,t) - (2\alpha + n\beta + k_{off} + r_{off})P_{on}(2;n,t) \quad (4)$$

$$\frac{dP_0(2;n,t)}{dt} = (n+1)\beta P_0(2;n+1,t) + k_{off}P_{on}(2;n,t) + r_{on}P_0(1;n,t) - ((\alpha + \alpha/y) + n\beta + k_{on} + r_{off})P_0(2;n,t) \quad (5)$$

$$\frac{dP_j(2;n,t)}{dt} = (\alpha + \alpha/y^j)P_{j-1}(2;n-1,t) + (n+1)\beta P_j(2;n+1,t) + r_{on}P_j(1;n,t) - ((\alpha + \alpha/y^{j+1}) + n\beta + k_{on} + r_{off})P_j(2;n,t) \quad (6)$$

where $P_{on}(1;n,t)$ is defined as the probability density function of being in the ON state on the 1 RNAP plane with n RNAs at time t, and $P_j(1;n,t)$ is a probability density function to find the system on the 1 RNAP plane and the state j (j = 0, 1, ...) with n RNA molecules at the time t. Similarly, $P_{on}(2;n,t)$ and $P_j(2;n,t)$

https://doi.org/10.1021/acs.jpcb.1c01859
J. Phys. Chem. B 2021, 125, 4692−4700

have analogous definitions but for the 2 RNAP plane. For simplicity, the t will be dropped in the subsequent discussions. These master equations reflect the dynamic evolution of molecular fluxes between different chemical states in transcription. Assuming that the system can reach the steady state ($t \to \infty$), the dynamic quantities in this model can be explicitly evaluated using the generating functions method,[29,30] as explained in detail in the Supporting Information.

To understand better some dynamic features of the system, it is convenient to consider a reduced chemical-kinetic scheme presented in Figure 3. We combine all ON states on the 1

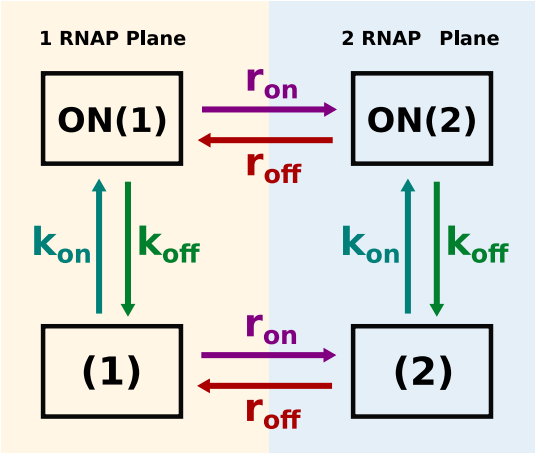


Figure 3. Reduced chemical-kinetic scheme for the model of collective dynamics of RNAPs during transcription. Macrostate (ON,1) describes all chemical states when the gyrase is bound and only one RNAP transcribes. Macrostate (ON,2) describes all chemical states when the gyrase is bound and two RNAPs transcribe. Macrostate (1) describes all chemical states when the gyrase is unbound and only one RNAP transcribes. Macrostate (2) describes all chemical states when the gyrase is unbound and two RNAPs transcribe.

RNAP plane as a new macrostate labeled as ON(1) (see Figures 2 and 3). Now, for each j, we can also combine all chemical states with the fixed j and label them as $j(1)$. After that, we combine together all j states into a new macrostate (1) (for all j). Similarly, for the 2 RNAP plane, we create a new macrostate ON(2) and states $j(2)$ that we eventually combine in a new macrostate (2). Then, in the stationary state, the probability to find the system in the macrostate ON(1) is $P_{on}(1) = \sum_{n=0}^{\infty} P_{on}(1;n)$, and the probability to be found in the state $j(1)$ is $P_j(1) = \sum_{n=0}^{\infty} P_j(1;n)$, while the probability to have the system in the macrostate (1) is $P(1) = \sum_{j=0}^{\infty} P_j(1)$. Similarly, the stationary state probabilities for the macrostate ON(2) and $j(2)$ are $P_{on}(2) = \sum_{n=0}^{\infty} P_{on}(2;n)$ and $P_j(2) = \sum_{n=0}^{\infty} P_j(2;n)$, respectively, while the probability for the macrostate (2) is $P(2) = \sum_{j=0}^{\infty} P_j(2)$. We note that once the system reaches the steady state, there is going to be zero net flux between the macrostates (see Figure 3), although there still could be nonzero molecular fluxes between individual chemical states illustrated in Figure 2. Also, the normalization requires that

$$\sum_{j=0}^{\infty} [P_j(1) + P_j(2)] + P_{on}(1) + P_{on}(2) = 1 \tag{7}$$

Using this result and zero-net-flux conditions, we can determine the stationary state probabilities for these macrostates

$$P_{on}(1) = \frac{\sigma}{(1+\gamma)(1+\sigma)} \tag{8}$$

$$P(1) = \frac{\gamma\sigma}{(1+\gamma)(1+\sigma)} \tag{9}$$

$$P_{on}(2) = \frac{1}{(1+\gamma)(1+\sigma)} \tag{10}$$

$$P(2) = \frac{\gamma}{(1+\gamma)(1+\sigma)} \tag{11}$$

where $\gamma = k_{off}/k_{on}$ is the equilibrium constant for the gyrase dissociation and $\sigma = r_{off}/r_{on}$ is the equilibrium constant for the polymerase dissociation.

As shown in detail in the Supporting Information, the stationary state probabilities for individual j states can be calculated, producing

$$P_0(1) = \frac{k_{off}(k_{on} + r_{off} + \alpha + \alpha/y)P_{on}(1)}{\Gamma_0} + \frac{r_{off}k_{off}P_{on}(2)}{\Gamma_0} \tag{12}$$

$$P_0(2) = \frac{k_{off}(k_{on} + r_{on} + \alpha/y)P_{on}(2)}{\Gamma_0} + \frac{r_{on}k_{off}P_{on}(1)}{\Gamma_0} \tag{13}$$

where

$$\Gamma_0 = (k_{on} + \alpha/y)(k_{on} + \alpha + \alpha/y) + r_{on}(\alpha + \alpha/y + k_{on}) + r_{off}(\alpha/y + k_{on}) \tag{14}$$

Also, it can be shown that

$$P_j(1) = \frac{\alpha/y^j(\alpha + \alpha/y^{j+1} + k_{on} + r_{off})P_{j-1}(1)}{\Gamma_j} + \frac{r_{off}(\alpha + \alpha/y^j)P_{j-1}(2)}{\Gamma_j} \tag{15}$$

$$P_j(2) = \frac{(\alpha + \alpha/y^j)(\alpha/y^{j+1} + k_{on} + r_{on})P_{j-1}(2)}{\Gamma_j} + \frac{r_{on}\alpha/y^j P_{j-1}(1)}{\Gamma_j} \tag{16}$$

with

$$\Gamma_j = (k_{on} + \alpha/y^{j+1})(k_{on} + \alpha + \alpha/y^{j+1}) + r_{off}(\alpha/y^{j+1} + k_{on}) + r_{on}(k_{on} + \alpha + \alpha/y^{j+1}) \tag{17}$$

These expressions allow us to exactly evaluate stationary probabilities for all chemical states, from which the dynamic properties can be derived.

Theoretical calculations also yield the average number of the produced RNA molecules in the system under the stationary conditions (see the Supporting Information)

https://doi.org/10.1021/acs.jpcb.1c01859
J. Phys. Chem. B 2021, 125, 4692−4700

$$\langle n\rangle = x\left[P_{on}(1) + 2P_{on}(2) + \sum_{j=0}^{\infty}\frac{1}{y^{j+1}}P_j(1) + \sum_{j=0}^{\infty}\left(1+\frac{1}{y^{j+1}}\right)P_j(2)\right] \quad (18)$$

where $x = \alpha/\beta$ is the equilibrium constant for the RNA synthesis. Moreover, the second moment of the probability distribution function with respect to n can be found, allowing us to calculate a Fano factor F, which is a dimensionless ratio of the variance and mean

$$F = \frac{\langle n^2\rangle - \langle n\rangle^2}{\langle n\rangle} \quad (19)$$

This is an important quantity because it provides a convenient measure of the stochastic fluctuations and noise in the system.

For the second moment, we have (see the Supporting Information)

$$\langle n^2\rangle = \langle n\rangle + x\left[f_{on}(1) + 2f_{on}(2) + \sum_{j=0}^{\infty}\frac{1}{y^{j+1}}f_j(1) + \sum_{j=0}^{\infty}\left(1+\frac{1}{y^{j+1}}\right)f_j(2)\right] \quad (20)$$

where the coefficients f are given by

$$f_{on}(1) = \frac{\alpha P_{on}(1)(r_{off}+r_{on})(k_{on}+k_{off}+\beta+r_{off}) + 2\alpha P_{on}(2)r_{off}(r_{off}+r_{on})}{(r_{off}+r_{on})(\beta+k_{off}+k_{on}+r_{off}+r_{on})(\beta+k_{off}+k_{on})} + \frac{\langle n\rangle r_{off}k_{on}(\beta+k_{off}+k_{on}+r_{off}+r_{on})}{(r_{off}+r_{on})(\beta+k_{off}+k_{on}+r_{off}+r_{on})(\beta+k_{off}+k_{on})} \quad (21)$$

$$f_{on}(2) = \frac{\alpha P_{on}(1)r_{on}(r_{off}+r_{on}) + 2\alpha P_{on}(2)(r_{off}+r_{on})(k_{on}+k_{off}+\beta+r_{on})}{(r_{off}+r_{on})(\beta+k_{off}+k_{on}+r_{off}+r_{on})(\beta+k_{off}+k_{on})} + \frac{\langle n\rangle r_{on}k_{on}(k_{on}+k_{off}+\beta+r_{on}+r_{off})}{(r_{off}+r_{on})(\beta+k_{off}+k_{on}+r_{off}+r_{on})(\beta+k_{off}+k_{on})} \quad (22)$$

$$f_0(1) = \frac{(\beta+k_{on}+r_{off}+\alpha+\alpha/y)k_{off}f_{on}(1) + r_{off}k_{off}f_{on}(2)}{(\beta+k_{on}+r_{on}+\alpha/y)(\beta+k_{on}+r_{off}+\alpha+\alpha/y) - r_{off}r_{on}} \quad (23)$$

$$f_0(2) = \frac{(\beta+k_{on}+r_{on}+\alpha/y)k_{off}f_{on}(2) + r_{on}k_{off}f_{on}(1)}{(\beta+k_{on}+r_{on}+\alpha/y)(\beta+k_{on}+r_{off}+\alpha+\alpha/y) - r_{off}r_{on}} \quad (24)$$

In addition, we have

$$f_j(1) = \frac{(\beta+k_{on}+r_{off}+\alpha+\alpha/y^{j+1})(\alpha/y^j)P_{j-1}(1) + r_{off}(\alpha+\alpha/y^j)P_{j-1}(2)}{(\beta+k_{on}+r_{on}+\alpha/y^{j+1})(\beta+k_{on}+r_{off}+\alpha+\alpha/y^{j+1}) - r_{off}r_{on}} + \frac{(\beta+k_{on}+r_{off}+\alpha+\alpha/y^{j+1})(\alpha/y^j)f_{j-1}(1) + r_{off}(\alpha+\alpha/y^j)f_{j-1}(2)}{(\beta+k_{on}+r_{on}+\alpha/y^{j+1})(\beta+k_{on}+r_{off}+\alpha+\alpha/y^{j+1}) - r_{off}r_{on}} \quad (25)$$

and

$$f_j(2) = \frac{(\beta+k_{on}+r_{on}+\alpha/y^{j+1})(\alpha+\alpha/y^j)P_{j-1}(2) + r_{on}(\alpha/y^j)P_{j-1}(1)}{(\beta+k_{on}+r_{on}+\alpha/y^{j+1})(\beta+k_{on}+r_{off}+\alpha+\alpha/y^{j+1}) - r_{off}r_{on}} + \frac{(\beta+k_n+r_n+\alpha/y^{j+1})(\alpha+\alpha/y^j)f_{j-1}(2) + r_{on}(\alpha/y^j)f_{j-1}(1)}{(\beta+k_{on}+r_{on}+\alpha/y^{j+1})(\beta+k_{on}+r_{off}+\alpha+\alpha/y^{j+1}) - r_{off}r_{on}} \quad (26)$$

Please also note that explicit expressions for all coefficients f have been obtained assuming that polymerase bindings and unbindings take place quite frequently in comparison with other processes. To supplement the analytical calculations, we also investigated the RNAP cooperation model using Monte Carlo computer simulations based on a Gillespie algorithm.[31]

RESULTS AND DISCUSSION

Our theoretical approach allows us to evaluate the distributions of produced RNA molecules and the degree of supercoiling. The results are presented in Figure 4. One can see that

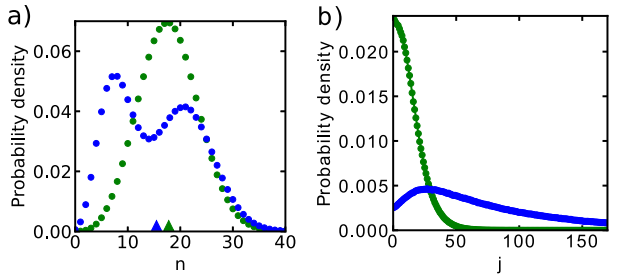


Figure 4. Probability distributions of (a) produced RNA transcripts, n, and (b) degree of supercoiling, which is measured as the number of transcripts j produced after gyrase dissociated from the DNA. Symbols are from computer simulations, and solid curves are from analytical calculations. In (a), the triangles indicate the means of the distributions of the corresponding color. The parameters used in computations are $\alpha = 15\ s^{-1}$, $\beta = 1\ s^{-1}$, $r_{on} = r_{off} = 10\ s^{-1}$, $y = 1.1$, $k_{on} = k_{off} = 1\ s^{-1}$ (green), and $k_{on} = k_{off} = 0.1\ s^{-1}$ (blue).

depending on the parameter choice, two drastically different behaviors might be observed. When the rates of the gyrase association and dissociation are slow in comparison with other processes, the system has enough time to spend in the ON states (with one or two RNAPs) where the RNA synthesis is quite fast as well as in the j states (with one or two RNAPs) where the RNA synthesis is slower due to supercoiling buildup. This leads to bimodal distribution in the production of RNA transcripts (blue curves in Figure 4a). The peak at a higher value of n corresponds to the system being in one of the ON states, while the peak at a smaller n describes the elongation with the resistive supercoiling action (j states). However, when the binding/unbinding dynamics of the gyrase molecules becomes quite fast, the distribution of produced RNA becomes unimodal (green curve in Figure 4a). In this case, the system is not able to explore fully ON states and j states. Because of the fast transitions of the gyrase, the system can be viewed as operating in a combined equilibrium biochemical state where all the transition rates are averaged out over the states with the gyrase and without, leading to a single peak in n.

The dynamics of the gyrase binding/unbinding also influences the distribution of the degree of supercoiling that can be reached in the system, as shown in Figure 4b. We specifically associate the parameter j with the degree of supercoiling. This is because it counts the number of produced RNA transcripts after gyrase dissociates, while the synthesis of every new RNA molecule adds the same amount of supercoiling to the DNA template. One can see that when the gyrase association/dissociation dynamics is slow, a wide distribution in j is observed (blue curve, Figure 4b). For fast gyrase dynamics, however, the distribution is much narrower (green curve, Figure 4b). These observations can be explained

https://doi.org/10.1021/acs.jpcb.1c01859
J. Phys. Chem. B 2021, 125, 4692−4700

using the following arguments. For slow association/dissociation rates of the gyrase, significant times are spent in the j states when the RNA is synthesized, but the degree of supercoiling also increases. This allows the system to reach much larger values of j. However, for fast gyrase dynamics, the system does not have enough time to explore the large j region before resetting to the ON state, and this leads to the narrow distribution of the degree of supercoiling with the peak at $j = 0$.

An important role in our model is played by the parameter y, which determines how the RNA elongation rate decreases with supercoiling. One can think of y as a mechanochemical coupling parameter that quantifies how the mechanical energy of supercoiling buildup affects the chemical reaction of RNA synthesis. When $y = 1$, the supercoiling does not influence the RNA production, but for $y > 1$, supercoiling dampens the transcription elongation. Thus, as y increases, the average number of RNA transcripts in the system is expected to decrease, as shown in Figure 5a. The effect is stronger if the initial synthesis rate α is larger. Eventually, for very large values of y, the saturation value of the mean number of produced RNA transcripts is achieved. It corresponds to a balance between the ON states when the synthesis takes place unopposed and the j states where only the RNA degradation takes place for $y \gg 1$.

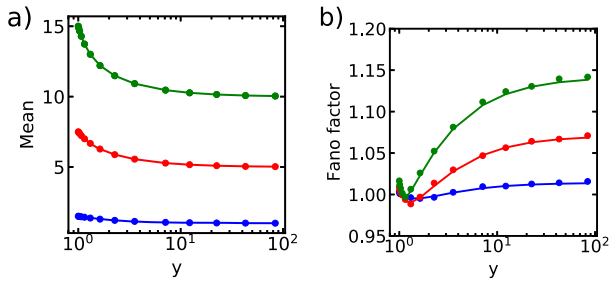


Figure 5. (a) Average number of transcripts and (b) Fano factor vs the level of mechanochemical coupling, y. A system with a fast RNA production rate is shown in green ($\alpha = 10$), a medium rate in red ($\alpha = 5$), and a slow rate in blue ($\alpha = 1.0$). Simulation results are represented by circles and analytical calculations by solid lines. Other parameters used are $\beta = 1$, $k_{on} = k_{off} = 10$, and $r_{on} = r_{off} = 50$.

A more complex non-monotonic behavior is observed for the dependence of the Fano factor as a function of the mechanochemical coupling. As the parameter y starts to increase, the Fano factor F first goes down slightly before increasing at larger values of y. Surprisingly, the value of F can even go below 1. This can be understood by recalling the definition of the Fano factor given in eq 19. When the supercoiling starts to affect the elongation rates (for y slightly above 1), there is a strong initial decrease in the mean number of produced RNA molecules, as shown in Figure 5a. This lowers the numerator in eq 19 much more than the denominator in the same expression, lowering the overall noise in the system. This result leads to an important prediction that there is an optimal level of mechanochemical coupling that allows the system to minimize noise while maintaining a significant production of RNA transcripts. It is interesting to note that for the real bacterial systems,[16] the estimated values of the parameter $y \sim 1.6-1.9$ are very close to this optimal range of parameters.[24] We can speculate then that one of the possible roles of collective dynamics of RNAPs might be not only to increase the production of RNA transcripts but also to use the supercoiling to minimize the level of noise in the system.

The gyrase molecules play an important role in our model since they occasionally release supercoiling, lowering the dampening effect that a mechanochemical coupling has on RNA production. In Figure 6, we present the results of our calculations on how varying the equilibrium constant of gyrase dissociation $\gamma = k_{off}/k_{on}$ influences the mean number of RNA transcripts produced and the Fano factor. If γ is small, the gyrase binding rate is faster than the dissociation rate and the system is mostly found in the ON states, resulting in a higher average number of RNA transcripts. On the other hand, if γ is high, the gyrase dissociation rate dominates and the system tends to explore more j states, leading to a lower mean number of produced RNA molecules. However, if there is no coupling between supercoiling and RNA production ($y = 1$), the average number will not change regardless of the value of γ. Also, as expected, the effect of varying γ is larger when the mechanochemical coupling is stronger (large y): see Figure 6a.

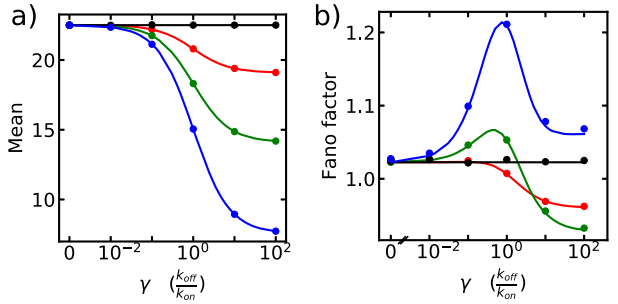


Figure 6. (a) Average number of RNA transcripts and (b) Fano factor as a function of the equilibrium constant for gyrase dissociation, $\gamma = k_{off}/k_{on}$. A system with no mechanochemical coupling is shown in black ($y = 1$), a low level of coupling is shown in red ($y = 1.1$), a medium level is shown in green ($y = 1.5$), and the high level is shown in blue ($y = 100$). Symbols represent simulation results, and solid lines are analytical predictions. Other parameters used in the analysis are $\alpha = 15$, $\beta = 1$, and $r_{on} = r_{off} = 50$.

The dependence of the gyrase association/dissociation dynamics on the Fano factor is more complex (Figure 6b). One can see that the stochastic fluctuations are the largest when the gyrase binding and dissociation rates are comparable ($\gamma \approx 1$). This happens because in this case, the system has access to the greatest number of chemical states. On the other hand, when the gyrase binding rate is greater than the dissociation rate, the system can be found mostly in the ON states with y-independent RNA synthesis rates. Thus, as γ approaches 0, the Fano factor curves for different values of y converge to the same value $F = 1$. When γ is large, the system mostly explores the y-dependent j states. As expected, the level of stochastic noise in the system depends on the value of the mechanochemical coupling y. When y is large, the mean is small compared to the variance in the distribution, leading to higher Fano factors. When y is small, the mean is large compared to the variance, leading to lower Fano factors. For realistic values of the gyrase equilibrium constant, we expect $\gamma \ll 1$, suggesting that the transcription will be very efficient with relatively low stochastic noise.

To test the importance of collective dynamics of RNAPs, we calculated the dynamic properties of the system as a function of the RNAP dissociation equilibrium constant $\sigma = r_{off}/r_{on}$, as

https://doi.org/10.1021/acs.jpcb.1c01859
J. Phys. Chem. B 2021, 125, 4692−4700

illustrated in Figure 7. When $\sigma \to 0$, two polymerases participate in transcription most of the time, and the average

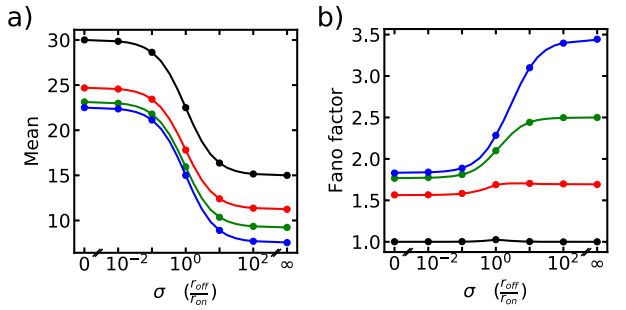


Figure 7. (a) Average number of the RNA transcripts and (b) Fano factor as a function of the RNAP dissociation equilibrium constant, $\sigma = r_{off}/r_{on}$. A system with no mechanochemical coupling is shown in black ($y = 1$), a low level of coupling is shown in red ($y = 1.1$), a medium level is shown in green ($y = 1.5$), and a high level is shown in blue ($y = 100$). The parameters used in our analysis are $\alpha = 15\ s^{-1}$, $\beta = 1\ s^{-1}$, and $k_{on} = k_{off} = 1\ s^{-1}$.

RNA production is quite high (Figure 7a). However, the situation is the opposite in the limit of high σ when mostly one RNAP participates in transcription elongation. This significantly lowers the production of RNA molecules. The calculations also show that increasing the level of the mechanochemical coupling generally lowers the average synthesis rate.

One interesting finding of our theoretical analysis is that for non-zero mechanochemical coupling ($y > 1$), a system with mostly two active polymerases has less stochastic noise than a system with mostly one active polymerase: see Figure 6b. This appears to be counter-intuitive because transcription with two polymerases is more complex due to a larger number of chemical states involved, and therefore, it is expected to have more fluctuations. However, in the two-polymerase system ($\sigma \ll 1$), the cooperation between polymerases ensures that the variance in the number of produced RNA transcripts is similar to the mean number of transcripts. This is because the synthesis rates vary between α and 2α. On the other hand, in the system with only one RNAP ($\sigma \gg 1$), the variance is kept on the same level (the synthesis rates are between zero and α), but the mean number of RNA transcripts is significantly smaller, resulting in a greater Fano factor. However, this effect disappears as y approaches 1 because in this case, there is no range in the synthesis rates in either limit. In the one-polymerase case, the rate is α in all states, while in the case with two polymerases, all states have a synthesis rate 2α. When this happens, the system can be effectively viewed as a single biochemical state, resulting in Poisson dynamics with a signature Fano factor being equal to 1 (Figure 7b).

The important conclusion from our analysis, as illustrated by Figure 7, is that increasing the number of cooperative events (decreasing σ) has two beneficial effects on transcription. First, it increases the overall production rate of RNA transcripts, and second, it lowers the stochastic noise. Both features make the transcription processes in the presence of multiple RNAPs more efficient.

To better understand the effect of collective dynamics of RNAPs in transcription, in Figure 8, we plot the Fano factor versus the mean number of RNA transcripts for models with and without cooperativity for different levels of the

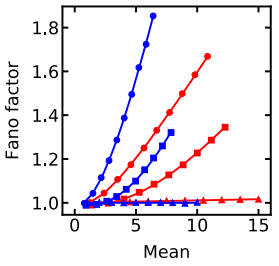


Figure 8. Relations between the Fano factor and the average number of produced RNA transcripts for two models: with RNAP cooperation and without it. The cooperative systems are shown in red, while the systems with a single RNAP are shown in blue. Triangles correspond to the system with no mechanochemical coupling ($y = 1$), the low level of the coupling is represented by squares ($y = 1.1$), and the higher level of the coupling is shown in circles ($y = 1.5$). Symbols are from computer simulations, and solid curves are from analytical calculations. The parameters used in our calculations are $\beta = 1\ s^{-1}$, $k_{on} = k_{off} = 1\ s^{-1}$, and $r_{on} = r_{off} = 50\ s^{-1}$.

mechanochemical coupling. It is found that there is little difference between the cooperative and zero-cooperativity models when there is no mechanochemical coupling ($y = 1$). However, for more realistic situations of weak or moderate mechanochemical couplings, the cooperative models significantly outperform the non-cooperative models: see Figure 8. For the same level of stochastic noise, the cooperativity leads to a larger production of RNA molecules. In addition, for the same level of RNA synthesis, cooperativity reduces the noise in the system. These observations suggest that during transcription in cells, the polymerases might effectively interact with each other over long distances, which allows them to make transcription a very efficient biological process.

To show the relevance of our theoretical model for real biological systems, in Figure 9, we present the experimental

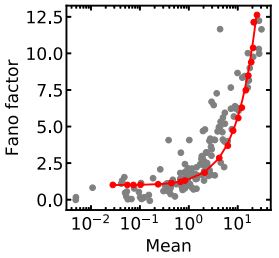


Figure 9. Comparison of the Fano factor vs mean experimental data from different genes in *E. coli* bacteria with theoretical predictions from the cooperative model. Experimental data for different promoters are in gray.[11] Theoretical predictions are in red. Symbols are from computer simulations, and curves are due to analytical calculations. The following parameters were used in computations: $\beta = 1.0 \times 10^{-2}\ s^{-1}$, $k_{on} = k_{off} = 1.0 \times 10^{-4}\ s^{-1}$, $r_{on} = r_{off} = 1.0 \times 10^{-4}\ s^{-1}$, $y = 1.1$, and variable values of α.

data of the Fano factor and the mean RNA production for a variety of different genes in *E. coli* bacteria.[11] These data show that the Fano factor is more or less constant for relatively low mean numbers, but it starts to grow fast for a larger mean. Our theoretical model can capture this trend, as shown in Figure 9. We found a set of reasonable parameters that can well describe the dependence between F and $\langle n \rangle$ for *E. coli* bacteria. This is another indication that cooperativity might be an important feature of transcription in real biological systems.

https://doi.org/10.1021/acs.jpcb.1c01859
J. Phys. Chem. B 2021, 125, 4692−4700

SUMMARY AND CONCLUSIONS

We developed a discrete-state stochastic model of the collective motion of RNAPs that exhibit long-range supercoiling-mediated interactions. It allowed us to quantitatively investigate the effect of RNAP's cooperativity on transcription dynamics. Our minimal theoretical model takes into account the most relevant physicochemical processes in transcription elongation such as the polymerase binding and dissociation, the supercoiling buildup and release due to the reversible attachments of the gyrases, and the RNA synthesis and degradation. The model is solved analytically, and theoretical calculations are fully supported by Monte Carlo computer simulations. Our theoretical analysis shows that the system might exhibit transcriptional bursting dynamics as evidenced by the bimodal probability distributions of the produced RNA transcripts for a certain range of parameters. The analysis suggested that there is an optimal level of mechanochemical coupling between the supercoiling buildup and the RNA synthesis slowdown when the level of the stochastic noise is minimized and the average number of the produced transcripts is maximized. It is concluded then that the supercoiling-mediated RNAP cooperativity has two major effects in transcription, that is, the increase in the RNA production accompanied by the reduction in the level of the noise. Furthermore, it is shown that our theoretical model can provide an excellent description of experimental data on transcription in bacterial systems.

Although our theoretical model is able to capture the main features of the RNAP cooperativity during transcription, it is clearly oversimplified and discussions on the validity of utilized approximations are needed. The model approximates the RNA synthesis as a one-step chemical transition, but in reality, it involves multiple biochemical steps and transformations with the assistance of many other protein molecules.[1,2] We also assumed that the gyrase binding to DNA results in the immediate release of positive supercoiling. However, this might not be the case as indicated by recent experimental measurements[16] and theoretical arguments.[32] It was suggested that the supercoiling relaxation might be a slow process and that this might influence the effective interactions between polymerases. Moreover, the model simply assumes that the frequent binding of Topo IA in the negatively supercoiled region of the DNA releases the negative supercoiling and that the less-frequent binding of gyrase in the positively supercoiled region releases the positive supercoiling. Therefore, the model does not account for the effects that might result from Topo IA and gyrase binding to other regions on the DNA. In addition, our model does not consider the fact that *in vivo* DNA generally has a non-neutral superhelical state that helps regulate gene expression.[33,34] Finally, the model assumes that only one or two polymerases participate in transcription, while in a real cellular environment, there may be none or there may be large numbers of RNAPs transcribing at the same time. Our theoretical method, however, can be modified and generalized to take this into account. However, we believe that this will not change the main conclusions of the analysis. In addition, in our analysis, we neglected the possibility of RNAPs bumping into each other because of the effective 1D nature of transcription. It will be important to consider this effect in a more detailed microscopic study of transcription that views the elongation as multiple transitions along the DNA chain. Despite these limitations, our theoretical approach clarifies the microscopic mechanisms of cooperativity during transcription providing explicit theoretical predictions that can be tested in experiments.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jpcb.1c01859.

Explicit analytical calculations of the stationary state probabilities, calculations of the first and second moments of the RNAP cooperation model using the generating functions method, and calculations of the above quantities for the single-RNAP case of the model (PDF)

AUTHOR INFORMATION

Corresponding Author

Anatoly B. Kolomeisky − *Department of Chemistry, Center for Theoretical Biological Physics, Department of Chemical and Biomolecular Engineering, and Department of Physics and Astronomy, Rice University, Houston, Texas 77005, United States;* orcid.org/0000-0001-5677-6690; Phone: +1 713 3485672; Email: tolya@rice.edu

Author

Alena Klindziuk − *Department of Chemistry and Center for Theoretical Biological Physics, Rice University, Houston, Texas 77005, United States*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jpcb.1c01859

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported by the Welch Foundation (C-1559), by the NSF (CHE-1953453 and MCB-1941106), and by the Center for Theoretical Biological Physics sponsored by the NSF (PHY-2019745).

REFERENCES

(1) Lodish, H. F. *Molecular Cell Biology*, 6th ed.; W. H. Freeman and Company, 2008.

(2) Alberts, B. *Molecular Biology of the Cell*; Garland Science, 2008.

(3) Phillips, R.; Kondev, J.; Theriot, J.; Garcia, H. *Physical Biology of the Cell*; Garland Science, 2012.

(4) Sainsbury, S.; Bernecky, C.; Cramer, P. Structural Basis of Transcription Initiation by RNA Polymerase II. *Nat. Rev. Mol. Cell Biol.* **2015**, *16*, 129−143.

(5) Ma, J.; Wang, M. D. DNA Supercoiling During Transcription. *Biophys. Rev.* **2016**, *8*, 75−87.

(6) Larson, M. H.; Mooney, R. A.; Peters, J. M.; Windgassen, T.; Nayak, D.; Gross, C. A.; Block, S. M.; Greenleaf, W. J.; Landick, R.; Weissman, J. S. A Pause Sequence Enriched at Translation Start Sites Drives Transcription Dynamics in Vivo. *Science* **2014**, *344*, 1042−1047.

(7) Chen, H.; Larson, D. R. What Have Single-Molecule Studies Taught Us About Gene Expression? *Genes Dev.* **2016**, *30*, 1796−1810.

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

(9) Munsky, B.; Neuert, G.; van Oudenaarden, A. Using Gene Expression Noise to Understand Gene Regulation. *Science* **2012**, *336*, 183−187.

(10) Golding, I.; Paulsson, J.; Zawilski, S. M.; Cox, E. C. Real-Time Kinetics of Gene Activity in Individual Bacteria. *Cell* **2005**, *123*, 1025−1036.
(11) So, L.-h.; Ghosh, A.; Zong, C.; Sepulveda, L.; Segev, R.; Golding, I. General Properties of Transcriptional Time Series in Escherichia coli. *Nat. Genet.* **2013**, *43*, 554−560.
(12) Sanchez, A.; Choubey, S.; Kondev, J. Regulation of Noise in Gene Expression. *Annu. Rev. Biophys.* **2013**, *42*, 469−491.
(13) Raser, J. M.; O'Shea, E. K. Noise in Gene Expression: Origins, Consequences, and Control. *Science* **2005**, *309*, 2010−2013.
(14) Raj, A.; van Oudenaarden, A. Nature, Nurture or Chance: Stochastic Gene Expression and Its Consequences. *Cell* **2008**, *135*, 216−226.
(15) Paulsson, J. Models of Stochastic Gene Expression. *Phys. Life Rev.* **2005**, *2*, 157−175.
(16) Chong, S.; Chen, C.; Ge, H.; Xie, X. S. Mechanism of Transcriptional Bursting in Bacteria. *Cell* **2014**, *158*, 314−326.
(17) Sevier, S. A.; Levine, H. Properties of Gene Expression and Chromatin Structure With Mechanically Regulated Elongation. *Nucleic Acids Res.* **2018**, *46*, 5924−5934.
(18) Sevier, S. A.; Levine, H. Mechanical Properties of Transcription. *Phys. Rev. Lett.* **2017**, *118*, 268101.
(19) Sevier, S. A.; Kessler, D. A.; Levine, H. Mechanical Bounds to Transcriptional Noise. *Proc. Natl. Acad. Sci. U.S.A.* **2016**, *113*, 13983.
(20) Kim, S.; Beltran, B.; Irnov, I.; Jacobs-Wagner, C. Long-Distance Cooperative and Antagonistic RNA Polymerase Dynamics via DNA Supercoiling. *Cell* **2019**, *179*, 106−119.
(21) Kim, S. Better Together: Co-operation and Antagonism Between RNA Polymerases During Transcription in Vivo. *BioEssays* **2020**, *42*, 1900215.
(22) Liu, L. F.; Wang, J. C. Supercoiling of the DNA Template During Transcription. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 7024−7027.
(23) Guptasarma, P. Cooperative Relaxation of Supercoils and Periodic Transcriptional Initiation Within Polymerase Batteries. *BioEssays* **1996**, *18*, 325−332.
(24) Klindziuk, A.; Meadowcroft, B.; Kolomeisky, A. B. A Mechanochemical Model of Transcriptional Bursting. *Biophys. J.* **2020**, *118*, 1213−1220.
(25) Bai, L.; Santangelo, T. J.; Wang, M. D. Single-molecule Analysis of RNA Polymerase Transcription. *Annu. Rev. Biophys. Biomol. Struct.* **2006**, *35*, 343−360.
(26) Xu, X.; Zhang, Y. Theoretical Model of Transcription Based on Torsional Mechanics of DNA Template. *J. Stat. Phys.* **2019**, *174*, 1316−1326.
(27) Ma, J.; Bai, L.; Wang, M. D. Transcription Under Torsion. *Science* **2013**, *340*, 1580−1583.
(28) Heberling, T.; Davis, L.; Gedeon, J.; Morgan, C.; Gedeon, T. A Mechanistic Model for Cooperative Behavior of Co-transcribing RNA Polymerases. *PLoS Comput. Biol.* **2016**, *12*, No. e1005069.
(29) Peccoud, J.; Ycart, B. Markovian Modeling of Gene-Product Synthesis. *Theor. Popul. Biol.* **1995**, *48*, 222−234.
(30) Klindziuk, A.; Kolomeisky, A. B. Theoretical Investigation of Transcriptional Bursting: A Multistate Approach. *J. Phys. Chem. B* **2018**, *122*, 11969−11977.
(31) Gillespie, D. T. Exact Stochastic Simulation of Coupled Chemical Reactions. *J. Phys. Chem.* **1977**, *81*, 2340−2361.
(32) Jing, X.; Loskot, P.; Yu, J. How Does Supercoiling Regulation on a Battery of RNA Polymerases Impact Bacterial Transcription Bursting? *Phys. Biol.* **2018**, *15*, 066007.
(33) Dorman, C. J. DNA Supercoiling and Transcription in Bacteria: a Two-way Street. *BMC Mol. Cell Biol.* **2019**, *20*, 26.
(34) Muskhelishvili, G.; Travers, A. The Regulatory Role of DNA Supercoiling in Nucleoprotein Complex Assembly and Genetic Activity. *Biophys. Rev.* **2016**, *8*, 5−22.

https://doi.org/10.1021/acs.jpcb.1c01859
J. Phys. Chem. B 2021, 125, 4692−4700