### Title
How does supercoiling regulation on a battery of RNA polymerases impact on bacterial transcription bursting?




### Journal
Physical Biology, 15(6)

### ISSN
1478-3967



### Authors
Jing, Xiaobo
Loskot, Pavel
Yu, Jin



### Publication Date
2018

### DOI
10.1088/1478-3975/aad933

Peer reviewed

IOP Publishing Phys. Biol. 15 (2018) 066007 https://doi.org/10.1088/1478-3975/aad933

Physical Biology

PAPER

# How does supercoiling regulation on a battery of RNA polymerases impact on bacterial transcription bursting?

Xiaobo Jing[1], Pavel Loskot[1,2] and Jin Yu[1]

[1] Beijing Computational Science Research Center, 100193, Beijing, People's Republic of China
[2] Systems and Process Engineering Centre, Swansea University, Swansea, SA28PP, United Kingdom

**E-mail:** jinyu@csrc.ac.cn

**RECEIVED**
12 February 2018

**REVISED**
5 August 2018

**ACCEPTED FOR PUBLICATION**
9 August 2018

**PUBLISHED**
6 September 2018

**Keywords:** transcription burst, DNA supercoiling, RNA polymerase (RNAP)

Supplementary material for this article is available online

## Abstract

Transcription plays an essential role in gene expression. The transcription bursting in bacteria has been suggested to be regulated by positive supercoiling accumulation in front of a transcribing RNA polymerase (RNAP) together with gyrase binding on DNA to release the supercoiling. In this work, we study the supercoiling regulation in the case of a battery of RNAPs working together on DNA by constructing a multi-state quantitative model, which allows gradual and stepwise supercoiling accumulation and release in the RNAP transcription. We solved for transcription characteristics under the multi-state bursting model for a single RNAP transcription, and then simulated for a battery of RNAPs on DNA with T7 and *Escherichia coli* RNAP types of traffic, respectively, probing both the average and fluctuation impacts of the supercoiling regulation. Our studies show that due to the supercoiling accumulation and release, the number of RNAP molecules loaded onto the DNA vary significantly along time in the traffic condition. Though multiple RNAPs in transcription promote the mRNA production, they also enhance the supercoiling accumulation to suppress the production. In particular, the fluctuations of the mRNA transcripts become highly pronounced for a battery of RNAPs transcribing together under the supercoiling regulation, especially for a long process of transcription elongation. In such an elongation process, though a single RNAP can work at a high duty ratio, multiple RNAPs are hardly able to do so. Our multi-state model thus provides a systematical characterization of the quantitative features of the bacterial transcription bursting; it also supports improved physical examinations on top of this general modeling framework.

## Introduction

Transcription is the first key step of gene expression. It is directed by RNA polymerases (RNAPs), the protein enzymes that move along double-stranded (ds) DNA and catalyze polymerization chain reactions to synthesize messenger RNA (mRNA) based on DNA template [1, 2]. The complete transcription includes initiation, elongation, and termination. A transcription bubble forms during the initiation, in which the RNAP enzyme starts unwinding dsDNA downstream. The unwound DNA strands reanneal afterwards upstream to the transcribing RNAP. It had been recognized in a twin-supercoiled-domain model that in front of and behind an elongating RNAP, positive and negative DNA supercoiling build up [3–5], respectively, as the dsDNA under transcription is subject to topological constrains at two ends to be prevented from freely rotating. Notably, it has been found in bacterial transcription systems, when the negative supercoiling can be quickly resolved by sufficiently abundant topoisomerase I (Topo I), the positive supercoiling built in front of an RNAP would persist to slow down further RNAP elongation and eventually turn off the transcription initiation [3–6]. Meanwhile, participation of DNA gyrase that resolves the positive supercoiling would allow recovery of the transcription [6–8] (see figure 1).

Recently, the dynamical processes of the positive supercoiling generation and gyrase binding to DNA to resolve the positive supercoiling have been sug-

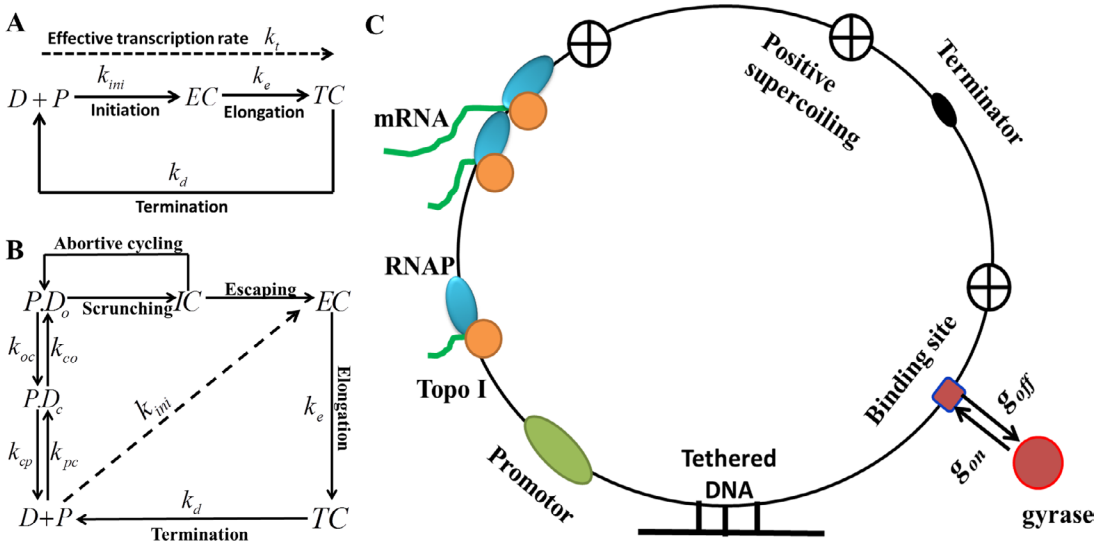


**Figure 1.** (A) A simplified transcription scheme including initiation, elongation, and termination. *D*, *P*, *EC*, and *TC* refer to DNA, RNAP, the elongation complex, and the termination complex, respectively. The effective initiation rate, the elongation rate, and the termination/dissociation rate are denoted as $k_{init}$, $k_e$, and $k_d$, respectively, while $k_t$ refers to an overall effective rate of transcription. These rates were used in the analytic model shown in SI. (B) A complete RNAP transcription scheme including additional steps in initiation, elongation, and termination. $P \cdot D_c$, $P \cdot D_o$, and *IC* refer to the closed RNAP-DNA complex, the open complex, and the initiation complex, respectively. $k_{pc}$, $k_{cp}$, $k_{co}$ and $k_{oc}$ are the RNAP binding rate, RNAP unbinding rate, DNA bubble opening rate, and its closing rate, respectively. (C) The schematics of the positive supercoiling accumulation during transcription on a tethered DNA loop. The green, orange, blue, black, and red ellipsoids represent the promoter, topo I, RNAP, terminator, and gyrase. The green line shows the mRNA in production. The black circle with the plus sign inside and the red square are on behalf of the positive supercoiling and the gyrase binding site, respectively. The gyrase binding and unbinding rates are denoted as $g_{on}$ and $g_{off}$, respectively.

gested to directly contribute to the bacterial transcription bursting [6]. The transcription bursting had been reported for individual *E. coli* cells [9], as gene ON and OFF events were recorded from real-time transcription activity measurements. Accordingly, one could attribute the gene transcription ON to OFF transition to relatively slow gyrase unbinding or disassociation from the DNA followed by fast supercoiling accumulation, while assign the OFF to ON transition to slow gyrase binding or association onto the DNA along with fast supercoiling releasing. As the gyrase binding and unbinding events are usually slow (with characteristic time ~$10^3$ s), the above scenario is consistent with a two-state model [10, 11], in which the gyrase bound and unbound states are assigned to the gene ON and OFF periods, respectively.

However, one can find that the transcription ON and OFF status do not necessarily correspond to the respective gyrase bound and unbound states to DNA, as the supercoiling accumulation and release can also proceed slowly (lasting from ~$10^2$ s to $10^3$ s) as well as in multiple kinetic steps. For example, when there is no gyrase bound to the DNA (treated as OFF in the two-state model), the transcription elongation can still proceed for a while (appear as ON in reality), during which the positive supercoiling gradually builds up and slows down the transcription activities until a final shut down of the activities. The corresponding process may take up to minutes over and do not reduce to a single kinetic event such as the gyrase unbinding transition in the two-state model. Meanwhile, gyrase can bind to DNA and once it binds and reacts, the supercoiling accumulation stops (treated as ON in the two-state model). However, the transcription cannot turn ON immediately upon the gyrase reaction, as it may still take several minutes for the accumulated supercoiling to fully release, so that the transcription activities only gradually recover toward the original level. In brief, the two-state model does not work well when the supercoiling accumulation and relaxation also proceed slowly, comparable to the gyrase binding and unbinding events. In such cases, a multi-state model is demanded to correctly characterize the kinetic features.

In this work, we provide a quantitative framework to accommodate in general that the supercoiling accumulation and releasing take place gradually and via multiple steps. To do that, we constructed a multi-state transcriptional bursting model, in which each configuration of the transcription system is characterized by the gyrase bound status as well as an accumulated supercoiling density level on DNA, which presumably increases one level at a time for each transcript production and decreases one by one as well during the supercoiling releasing. In the model, the mRNA production proceeds with time-dependent elongation and initiation rates, which are determined by the supercoiling density level. Overall, the supercoiling accumulation and release are regulated by the gyrase binding and unbinding events at constant rates (see figure 2). We abolished the assumptions in the two-state model that requires only one slow transition to be dominant to switch between the gene ON and OFF periods. Current multiple-state model is similar to a recent development where stepwise supercoiling accumulation is also

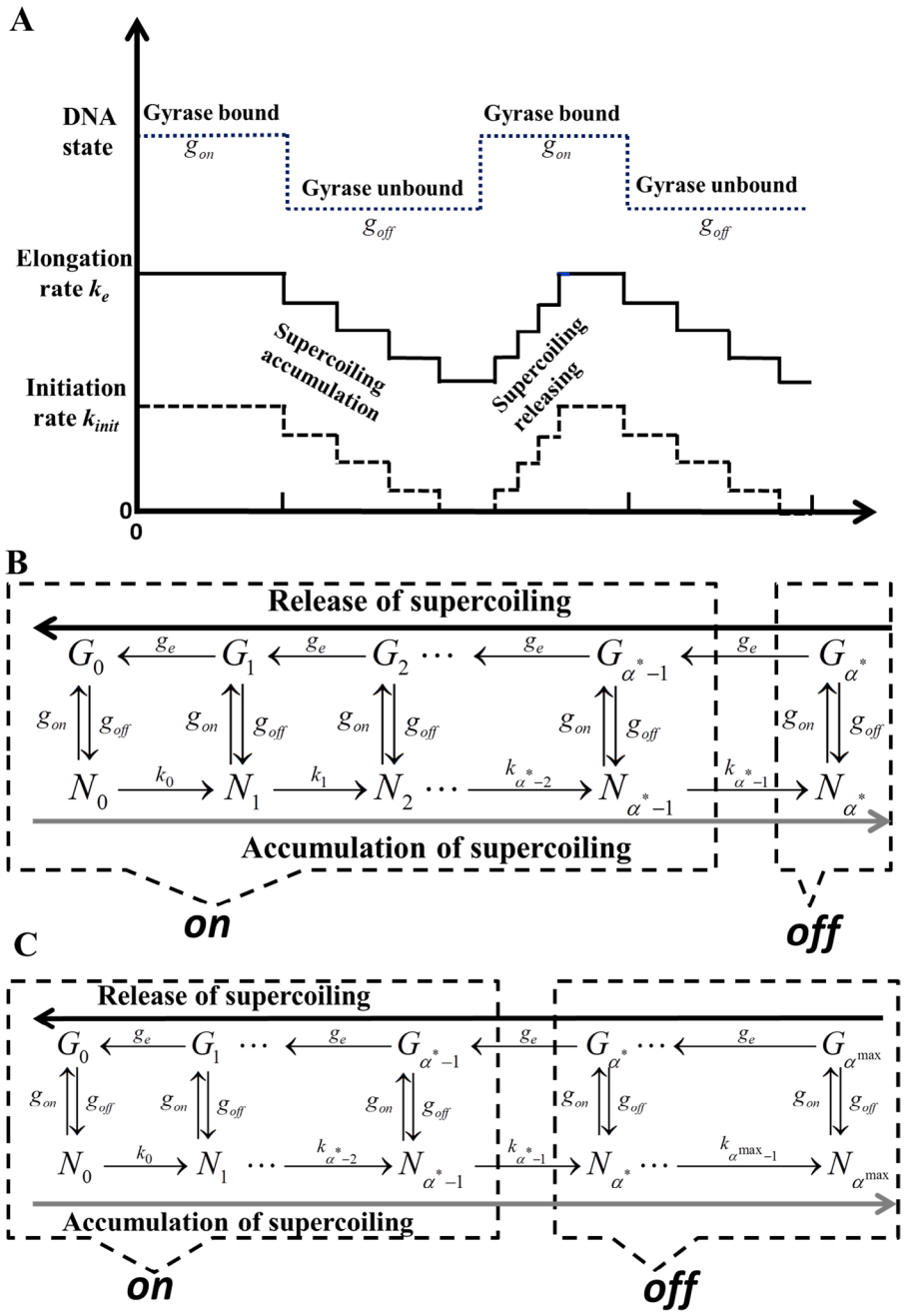


**Figure 2.** A summary schematics of the multi-state transcription-bursting model for both the single RNAP and traffic case. The changes of the effective transcription initiation rate ($k_{init}$) and the elongation rate ($k_e$) due to the positive supercoiling accumulation and relaxation during an RNAP transcription are shown in (A), in the presence of the gyrase binding and unbinding (at rates of $g_{on}$ and $g_{off}$). The simplified schemes of the multi-state transcription-bursting model for the single RNAP (B) and the RNAP traffic case (C) are also shown. The top and bottom layers represent the gyrase bound ($G$) and unbound ($N$) status, respectively, so that the switching in between the two status proceeds at $g_{on}$ and $g_{off}$. Without the gyrase bound or $N$, the supercoiling accumulates along with the transcript production at a rate of $k_\alpha$, with $\alpha^*$ the maximum supercoiling level formed in the single RNAP case; in the traffic case, there are still $Np$ RNAPs left on the DNA at the initiation stall ($\alpha = \alpha^*$), accordingly, the maximum supercoiling level reaches to $\alpha^{\max} = \alpha^* + Np$. Upon the gyrase binding or $G$, the supercoiling releases stepwise at a rate of $g_e$.

explicitly modeled and simulated to address mechanical bounds to transcription noises [12]. Nevertheless, in that model, immediate supercoiling release was employed, while the transcription kinetic was modeled without explicitly considering the gyrase association/dissociation events. A biophysical model was also built recently to take into account the supercoiling energy state of the transcription [13], focusing on the promoter impacts and correlated gene expression within a supercoiling domain. In addition, a general framework of stochastic transcription has been built to address both the uncorrelated transcription events and bursting behaviors [14]. Besides, the bursting transcription is suggested to lead to protein localization to facilitate target search in the transcription regulation, according to another recent modeling work [15].

In particular, we consider a battery of RNAPs forming elongation traffic on DNA, as it is an efficient way to achieve the supercoiling regulation when there are only a limited amount of gyrase molecules present

in the cell [8]. The multiple RNAPs can elongate in synchrony on average, so that the positive supercoiling generated in front of a trailing RNAP can be immediately canceled by the negative supercoiling produced behind a leading RNAP downstream in the traffic. In that case, a battery of RNAPs move in tandem and only accumulate net positive supercoiling in front of the first leading RNAP, so that only one gyrase molecule at a time rather than several are needed to resolve the accumulated supercoiling [8]. Meanwhile, the negative supercoiling produced behind the very last RNAP in the traffic can be resolved by Topo I that is comparatively abundant. Within this context, one might expect that multiple RNAPs moving in tandem would demonstrate similar or additive behaviors over individual RNAPs. Nevertheless, we show that a battery of RNAPs promotes both the mRNA and supercoiling accumulation, so that transcription noises are amplified and the transcription bursting becomes highly significant.

To model the inhibition impacts of the supercoiling on the RNAP elongation and initiation, we considered that the supercoiling generates tension on DNA to slow down both the elongation and initiation, with the initiation being impacted more significantly than the elongation. Recent measurements showed that the transcription initiation could be fully stopped while the elongation persisted at a low rate under the positive supercoiling inhibition [6]. Early studies had also indicated that the transcription initiation is sensitive to the tension on DNA [16], while the elongation rate is not highly force sensitive for T7 RNAP as well as *E. coli* RNAP [17–20]. The tension responses thus allow an existing RNAP traffic to continue elongating albeit slowly on DNA while no further RNAP transcription initiation is supported.

In this work, we first describe how the multi-state transcriptional bursting model was constructed for a single RNAP with repetitive runs of the transcription, along with analytic solutions of the model in the steady state condition. Next we utilized this model to simulate stochastically on how multiple RNAP molecules are loaded onto DNA and elongate in tandem for the same gene transcription. We chose the fast elongating T7 RNAP and the comparatively slow *E. coli* RNAP as two representative systems, which have also been directly monitored in the recent transcriptional bursting study, both *in vitro* and *in vivo* [6]. Consequently, we are able to show how the transcript productions along with the fluctuations are impacted by the supercoiling regulation for a battery of the RNAPs transcribing together on short and long DNA, respectively, based on the multi-state model.

## Methods & model construction

### Transcription and supercoiling accumulation simulation setup

The transcription process includes initiation, elongation, and termination (see figures 1(A) and (B)). For simplicity, one may treat the initiation as one step, with an effective rate of $k_{init}$. In our analytic model construction (see figure 1(A) and supplementary information or SI) (stacks.iop.org/PhysBio/15/066007/mmedia), the overall transcription was treated as one effective step. In the numerical simulation, however, detailed steps in the initiation and elongation were modeled: The promoter binding of an RNAP from free solution (with forward and backward rates of $k_{pc}$ and $k_{cp}$), a reversible close-to-open transition of the promoter along with formation of a transcription bubble (with forward and backward rates of $k_{co}$ and $k_{oc}$), and the RNAP complex together with the transcription bubble undergoing a series of conformational transitions via scrunching, abortive cycling, and escaping to elongation (see figure 1(B)) [21–23] (with detailed kinetics found in [21]). The mRNA is then produced during the elongation (at a rate of $k_e$) until termination or towards the end of DNA (with a dissociation rate $k_d$), and the length of gene ($L$) is usually set to hundreds to thousands of base pairs.

Studies had shown that the accumulation of positive supercoiling in front of an elongation RNAP inhibits the transcription, while the gyrase enzyme binding and action onto DNA release the positive supercoiling (see figure 1(C)), in particular, reduce the writhe number for two during each enzymatic cycle [24–26]. Besides, the elongating RNAP also produces negative supercoiling behind, which can be quickly resolved by Topo I that may interact directly with RNAP [27] and populate abundantly in the *E. coli* cell [6, 28] (see figure 1(C)). Otherwise, the negative supercoiling produced upstream of a leading RNAP can also be canceled by the positive supercoiling produced downstream of a trailing RNAP in the RNAP traffic, prior to the Topo I binding.

The *in vitro* experiments for single T7 and *E. coli* RNAP transcribing repetitively on the circular DNA template were actually conducted in three conditions (see figure 4 in [6]): in the absence of Topo I and gyrase, in the presence of Topo I and gyrase; and in the presence of Topo I only. There was no transcription slowing down nor bursting in the first two conditions. Only in the third case, however, the transcription was suppressed, revealing the substantial role of the positive supercoiling. In the experiment, the transcription initiation slowing down and recovery after adding the gyrase lasted for ~6000 s and 3000 s (see figure 5(B) in [6]), respectively. However, the gyrase binding from bulk takes ~1000 s at the corresponding condition ($g_{on} = \widetilde{g}_{on}[G]$ with $\widetilde{g}_{on} = 10^4\ \mathrm{M}^{-1}\ \mathrm{s}^{-1}$ and the solution gyrase concentration $[G] \sim 100$ nM), while the dissociation of the gyrase from DNA also takes ~1000 s ($g_{off} = 10^{-3}\ \mathrm{s}^{-1}$) [6]. Hence, there must be other events that account for a significant amount of time, which we propose as the slow accumulation of the positive supercoiling and the gradual release of the accumulated supercoiling.

To quantify the positive supercoiling accumulation on the DNA, we count the supercoiling density level

additively upon each transcript production, when the DNA loop or domain is tethered under a topological constraint and there is no gyrase bound to the DNA. The supercoiling density on the DNA with a linking number $Lk_0$ in the relaxed state is defined as

$$\sigma \equiv \Delta Lk/Lk_0, \tag{1}$$

in which the linking number $Lk$ is counted as the sum of a DNA helical twist number $Tw$ and a coiling or writhe number $Wr$

$$Lk = Tw + Wr. \tag{2}$$

In the relaxed state, the writhe number $Wr_0 = 0$, so that $Lk_0 = Tw_0 = L/10.5$, e.g. for the B-form DNA of length $L$.

For the downstream DNA under the topoligical constrain at the end, the linking number change upon the RNAP action is $\Delta Lk = l/10.5$ for a transcript of length $l$. Hence, upon production of each mRNA transcript of length $l$ ($l \leqslant L$), one has

$$\sigma_l = \frac{l/10.5}{Lk_0} = l/L. \tag{3}$$

In accumulation, we obtain the overall supercoiling density level as $\alpha = \sum_{i=1}^{i_{\max}} \sigma_l^i$ for the multiple transcript production, where $i_{\max}$ is the maximum number of transcripts being generated in the absence of the gyrase association.

As the positive supercoiling is built up in front of RNAP to shrink the length of the DNA that is either tethered or with the two ends restrained, we consider that the tension built up along DNA impedes the RNAP translocation and ultimately slows down the elongation or initiation. Dynamic impact from supercoiling propagation along DNA has not been considered yet in current model. We assumed that both the initiation and elongation rates (as well as the effective transcription rate) are determined only by the accumulated supercoiling density level $\alpha$ on the DNA as defined above, and these rates decrease linearly with $\alpha$ for multiple runs of the transcription. A close-to-linear dependence of the transcription rate on the supercoiling accumulation has actually been demonstrated in the recent biophysical modeling [13]. In addition, it has been shown that the positive supercoiling can slow down T7 or *E. coli* RNAP elongation while fully stop the transcription initiation [6]. Hence, when we separately treated the initiation and elongation processes in the numerical simulations (described below), we assumed that the effective initiation rate $k_{init}$ would drop all the way to zero upon the positive supercoiling accumulation, while the elongation rate $k_e$ would be lowered to a small but non-zero value ($k_e^{sm}$).

### Construction of a multi-state transcription-bursting model for single RNAP

First, we considered the case when there is only one RNAP during transcription. As the gyrase dissociates from DNA (at a rate $g_{off}$), the supercoiling density increases by one level upon each full-length transcript production. Hence, the supercoiling accumulation rate can be equal to the effective production rate of the mRNA transcript, which combines the initiation and elongation together, and was assumed to decrease linearly with the number of transcripts in production

$$k_\alpha = k_t\left(\alpha^* - \alpha\right)/\alpha^*, \tag{4}$$

with $k_t$ the transcription rate in the absence of the supercoiling. Note that the supercoiling level $\alpha$ rises along with the number of transcripts $i$ when there is no gyrase bound on the DNA. Once the supercoiling level reaches to a maximal value at $\alpha^*$, the transcription activity suspends. Note that when we treated the initiation and elongation separately in the numerical simulations, only the transcription initiation but not the elongation suspends at $\alpha = \alpha^*$.

Consequently, one sees that the supercoiling level increases in synchrony with the mRNA production in the absence of the gyrase. Once the gyrase binds onto DNA, the mRNA production can then proceed without coupling to the supercoiling accumulation; instead, the accumulated supercoiling starts releasing at a rate of $g_e$, as being estimated below. The detailed scheme of the multi-state model is presented in SI figure S1.

For T7 RNAP transcibing on the 12 kbp DNA template, e.g. the *in vitro* experiment recorded that the transcription initiation stops after generating 9 transcripts, and recovers back ~2000 s after adding gyrase in solution (see figure 5(B) in [6]). To count from the initiation *off* status, one has $\alpha = 9$ at $t = 0$. From $t = 0$ to 1000 s on average, gyrase has not yet bound to DNA ($g_{on} = 0.001\,\mathrm{s}^{-1}$ *in vitro*), and the transcript production rate was set low at ~$0.001\,\mathrm{s}^{-1}$ so that $\alpha$ further increases to 10 by e.g. $t = 1000$ s. Next, from $t = 1000$ to 2000 s on average, the gyrase molecule remains bound on the DNA ($g_{off} = 0.001\,\mathrm{s}^{-1}$), and all the supercoiling is then released (from $\alpha = 10$ to 0). Hence, the rate of releasing one supercoiling level is estimated as $g_e^0 = 10/1000/s = 0.01/s$. Since $g_e^0$ is for a DNA length of 12 000 bp ($L_0$), we scaled supercoiling releasing rate for a DNA length of $L$ as

$$g_e = \frac{L_0}{L} g_e^0. \tag{5}$$

The key kinetic parameters are summarized in table 1, while additional parameters can be found in SI table S1.

In SI, we show the chemical master equation for the above multi-state model, and then use the generating function method [33] to solve the equation at the steady-state condition. Consequently, the interested quantities including the average number of mRNA ($\langle m \rangle$), the Fano factor (*Fano*) [34], and the effective duty cyle ratio (*DCR*) [35] in the single RNAP transcription case were obtained.

### The multi-state transcription-bursting model for RNAP traffic

Then we implemented the multi-state model to describe a battery of RNAPs transcribing in traffic

**Table 1.** Parameter values based on references or estimation.

| Parameter | Symbol | Value | Units | Reference |
|---|---|---|---|---|
| DNA loop length | $L_0$ | 12 000 | bp | [6] |
| Effective transcription initiation rate | $k_{init}$ | 0.1–0.2 | $s^{-1}$ | *Estimated* |
| T7 RNAP elongation rate | $k_e$ | 50–100 | bp $s^{-1}$ | [29] |
| *E. coli* RNAP elongation rate | $k_e$ | 10–20 | bp $s^{-1}$ | [30] |
| Effective transcription rate without supercoiling ($L = 4.5$ kbp DNA) | $k_t$ | 0.02 for T7 0.002 for *E. coli* | $s^{-1}$ | *Estimated* |
| Termination rate | $k_d$ | 1 | $s^{-1}$ | [31] |
| mRNA degradation rate | $\gamma$ | 0.1 | $min^{-1}$ | [32] |
| First-order rate constant for gyrase-DNA binding | $\tilde{g}_{on}$ | $10^4$ | $M^{-1}$ $s^{-1}$ | [6] |
| Gyrase binding rate *in vitro* | $g_{on}$ | 0.001 | $s^{-1}$ | [6] |
| Gyrase binding rate *in vivo* | $g_{on}$ | 0.003 | $s^{-1}$ | [6] |
| Gyrase dissociation rate | $g_{off}$ | 0.001 | $s^{-1}$ | [6] |
| Supercoiling releasing rate for one density level and $L_0 = 12$ kbp DNA | $g_e^0$ | 0.01 | $s^{-1}$ | *Estimated* |

conditions, as one RNAP after another initiating from a same promoter and then elongating in tandem [8]. In between any two adjacent RNAPs on the DNA, the locally generated positive and negative supercoiling cancel each other. In a typical case, we considered that Topo I binds (e.g. ~$0.1\,s^{-1}$ to $0.01\,s^{-1}$) much more frequently than the gyrase ($g_{on} \sim 0.001\,s^{-1}$), but less frequently than the initiation of an individual RNAP ($k_{init} \sim 0.1\,s^{-1}$), and only bind to the upstream of the last RNAP in the traffic to resolve the negative supercoiling. In current setting, the initiation process usually lasts about several seconds (i.e. 5–10 s on average), Topo I binding thus happens after several RNAPs have been loaded onto the promoter for the transcription initiation. Indeed, prior to the Topo I binding, the negative supercoiling generated upstream to an elongating RNAP loaded last in the traffic might also assist the promoter opening to facilitate the next RNAP initiation.

Meanwhile, we still assume that the transcription initiation and elongation rates are mainly regulated by the tension on the DNA, which can be largely determined by the overall supercoiling density level $\alpha$. Hence, even though the positive supercoiling is accumulated only in front of the very leading RNAP, due to the tension generated on the DNA, all elongating RNAPs can be slowed down with $k_e$ dropping simultaneously. Similarly, the effective initiation rate $k_{init}$ drops as the supercoiling density or DNA tension builds up, until a full suspension of the initiation. A summary schematics of the transcription rate changes upon the supercoiling accumulation and release is shown in figure 2(A), while the simplified schemes of the multi-state model for the single RNAP and traffic cases are provided in figures 2(B) and (C), respectively.

In the RNAP traffic, although individual RNAP elongates at the same rate, on average, there is still some chance that two RNAPs collide stochastically. We consider two types of collisions: (i) For the viral T7 RNAP which does not usually pause, a trailing RNAP can kick a leading RNAP off the DNA track if the leading one slows down and gets in the way of the trailing one [36]; (ii) For the *E. coli* RNAP, a trailing RNAP likely assists a pausing RNAP in front to recover back to the elongation [37]. We recorded the number of simultaneously elongating RNAPs on the DNA as the working #RNAP, which is determined by the length of transcript $l$, the effective rates of initiation $k_{init}(\alpha)$ and elongation $k_e(\alpha)$ at a supercoiling level $\alpha$, as well as by the pausing interval time $\tau_{pause}$ if there is any (pause may occur during elongation of *E. coli* RNAP with $\tau_{pause} \sim 2$ s). Hence, the average number of working RNAP can be estimated by

$$\langle \#\mathrm{RNAP} \rangle = l/\left[\left(1/k_{init}(\alpha) + \tau_{pause}\right) k_e(\alpha)\right], \quad (6)$$

with the effective rates of initiation and elongation determined by the supercoiling density level as

$$k_{init}(\alpha) = k_{init}(\alpha^* - \alpha)/\alpha^*, \quad (7)$$

and

$$k_e(\alpha) = k_e(\alpha^* - \alpha)/\alpha^* + k_e^{sm}. \quad (8)$$

Note that now the maximal supercoiling level can reach up to $\alpha^{\max} = \alpha^* + Np$ before the gyrase binds, in which $Np$ is the number of RNAPs left on the DNA at $\alpha = \alpha^*$. That says, even though the transcription initiation stops at $\alpha = \alpha^*$, there can be $Np$ RNAPs left on the DNA, which have been loaded prior to the transcription initiation stall and can continue transcribing until the termination. Comparing with the single RNAP case (see figure 2(B)), the release of supercoiling or recovery of the transcription initiation takes a longer time on average in the traffic case (figure 2(C)). Note that in this work, we attributed the gene *off* status to those transcription configurations with $\alpha \geqslant \alpha^*$, i.e. when the positive supercoiling fully turns off the transcription initiation; otherwise, as along as $\alpha < \alpha^*$, the transcription status remains *on*.

### Numerical simulations on the RNAP traffic under supercoiling regulation

We use T7 and *E. coli* RNAPs as two representative model systems to show the multi-state transcription-bursting model. T7 RNAP can elongate comparatively

fast (up to 100–200 nt $s^{-1}$) during the elongation [29]. *E. coli* RNAP elongates slowly, and can pause or backtrack from time to time [38]. For simplicity, the initiation kinetics of the two RNAP species was treated similarly as from T7 RNAP in our simulation [21]. The particularly slow initiation, e.g. of the *E. coli* RNAP, was then considered similar to the single RNAP transcription case examined in *ii*. As mentioned above, we assume that the collision between two *E. coli* RNAPs leads to a velocity (i.e. the elongation rate) switch between the two RNAPs, while the collision between two T7 RNAPs causes the leading RNAP to drop off the track.

In this work, we used the kinetic Monte Carlo (KMC) algorithm [39] to simulate the RNAP transcription traffic. We simulated for the RNAP transcription, without the supercoiling first, and then in the presence of the supercoiling accumulation and releasing, according to the reaction schemes in figure 2(C). By default, we let the transcription initiation stop at a supercoiling density level of $\alpha^* = 4$. To compare, we used a short gene or transcript length (~300 bp) and a long one (~4500 bp) in the simulation systems. The model parameters are provided in table 1 and SI table S1.

In the KMC simulation we grouped the states into the '*on*' and '*off*' sets according to the initiation status. In the '*off*' set, the transcription initiation stops due to the high supercoiling density level ($\alpha \geqslant \alpha^*$). Otherwise, the states are in the '*on*' set ($\alpha < \alpha^*$). The *DCR* is calculated as the probability of keeping the transcription initiation *on* (see detailed calculations in SI). Note that the *on* and *off* sets defined here, according to the transcription initiation status in the simulations, do not match exactly with the ON and OFF periods measured experimentally or directly from time series of the mRNA production.

## Results

According to the multi-state transcription-bursting model presented in figure 2(B), we solved the steady-state chemical master equations analytically to obtain transcription characteristic of single RNAP (see SI and figure S1). Then we conducted KMC simulations to numerically show the steady-state characteristics in the RNAP traffic case (see figure 2(C)), comparing the situations with and without the supercoiling regulation.

### Characteristics of the single RNAP transcription in the multi-state model

Based on the chemical master equations of the multi-state model, we obtained analytic solutions in the steady state of single RNAP transcription. The characteristics are shown in figure 3 for T7 RNAP transcribing on a comparatively long piece of DNA ($L = 4500$ bp), including the average number of mRNA transcript ($\langle m \rangle$), the Fano factor (*Fano*), and the effective duty cycle ratio (*DCR*). These quantities are demonstrated by varying the gyrase binding/unbinding rate ($g_{on/off}$, 0.001 $s^{-1}$ by default), the supercoil-releasing rate ($g_e$, 0.027 $s^{-1}$), the effective transcription rate ($k_t$, 0.02 $s^{-1}$), the mRNA degradation rate ($\gamma$, 0.0017 $s^{-1}$ or 0.1 $min^{-1}$), and the supercoiling level for the initiation suspension ($\alpha^* = 4$). In the default setting, we have $\langle m \rangle \sim 6$, *Fano* ~ 3.5, and *DCR* ~ 0.66.

First, one can see that an increase of $g_{on}$ for ten times (to 0.01 $s^{-1}$) leads to about twice the amount of mRNA transcripts ($\langle m \rangle \sim 12$) along with a low value of the *Fano* (~1), indicating an enhanced transcription and quenched fluctuation when gyrases are abundant to allow sufficiently fast actions to remove the accumulated supercoiling. Lowering the gyrase unbinding rate $g_{off}$ shows similar impacts. The *Fano* remains low even when $g_{on}$ rises (>0.01 $s^{-1}$) or $g_{off}$ drops (<0.0001 $s^{-1}$) further. On the other hand, a decrease of $g_{on}$ for one order of magnitude (to 0.0001 $s^{-1}$) lowers the mRNA production close to zero while allowing a rise of the *Fano* factor (>5), as having fewer gyrases or slower gyrase binding encourages the system to stay longer in the gyrase unbound state, with inhibited transcript productions across multiple supercoiling density levels. In comparison, a significant rise of $g_{off}$ (e.g. to 0.01 $s^{-1}$) can lower both the mRNA production and the Fano factor, since the configuration with particularly high supercoiling level transits too fast to the gyrase unbound state, without being able to release the supercoiling. Nevertheless, the currently adopted supercoiling release rate upon the gyrase action ($g_e$) appears to be large enough such that a further increase of it does not improve the mRNA production, nor it leads to any decrease of the *Fano* or increase of the *DCR*. Reducing $g_e$, however, would significantly reduce the mRNA production, raise the *Fano*, and lower the *DCR*, making the transcription difficult and versatile.

In addition, one can see that a continuing rise of the effective transcription rate ($k_t$) would lead to a steep increase of the mRNA production, while the *Fano* also rises sharply. The *Fano* drops low as $k_t$ reduces to very low values. Lowering the mRNA degradation rate ($\gamma$) promotes both the mRNA production and the *Fano*, with no impacts on the *DCR*. Finally, if one varies $\alpha^*$, the maximum supercoiling density level for the transcription initiation suspension, neither the mRNA production nor the Fano factor changes much. However, the *DCR* increases from ~0.5 to approaching to 1 as $\alpha^*$ increases from 1 to 9.

### Verifying the multi-state characteristics in the time series of the mRNA production

To verify that our model captures multi-state characteristics that are experimentally detectable, we recorded the duration time of the ON and OFF periods according to the simulated time series of the mRNA production in our model, for both the single RNAP and traffic cases. Note that the current model is intrinsically multi-state, yet one could group various

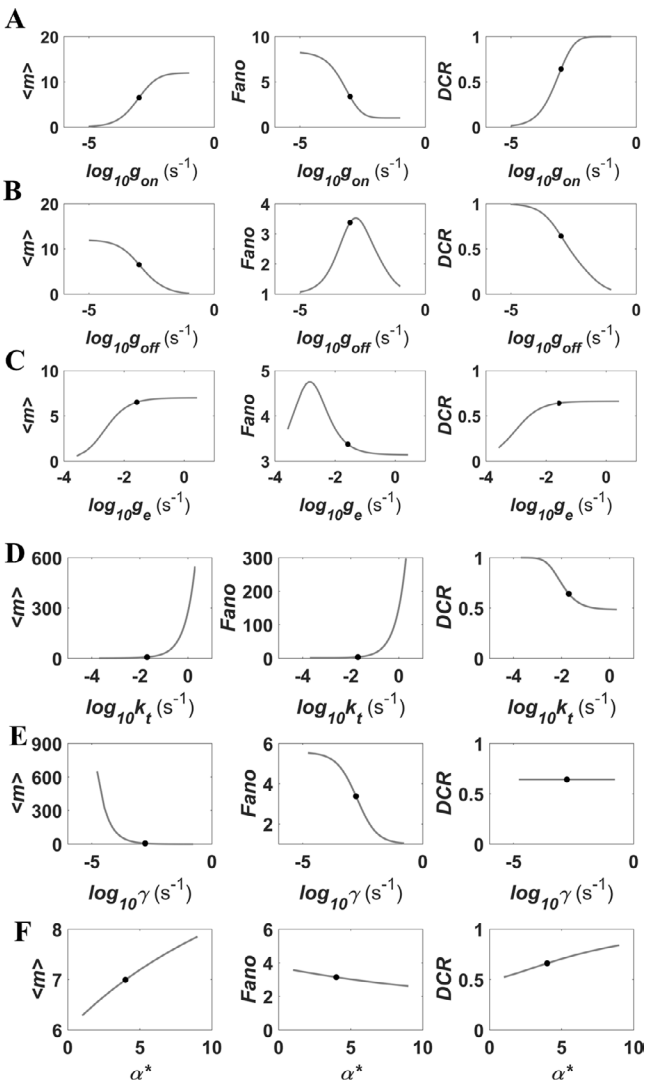


**Figure 3.** The steady-state characteristics in the single T7 RNAP transcription ($L = 4500$ bp) as obtained analytically from the multi-state transcription-bursting model in figure 2(B) (and SI figure S1). The changes of $\langle m \rangle$, *Fano*, and *DCR* upon variations of the gyrase binding and unbinding rates (A) $g_{on}$ and (B) $g_{off}$, the supercoiling release rate (C) $g_e$, the effective transcription rate (D) $k_t$, the mRNA degradation rate (E) $\gamma$, and the maximum supercoiling level (F) $\alpha^*$ are presented, respectively. The values labeled by the black points were calculated under the default setting.

configurations into the *on* and *off* sets, intrinsically, according to the transcription initiation status. To be compatible with experimental detections, however, we determined the transcription OFF period phenomenologically when there was no mRNA production within a certain time window. The time window was chosen at ~100 s, approximately an upper bound value for the time interval in between two consecutively generated transcripts during the transcription ON period.

Figures 4(A) and (B) show the time series of the mRNA production obtained in the single T7 RNAP transcription and the corresponding traffic case, respectively. The DNA length is $L = 4500$ bp, and there are 14 RNAPs in the cell (~38 nM) in the traffic case. In comparison, one notices that the average transcript production ON time is longer in the single T7 RNAP transcription ($T_{\rm ON} = 19.5 \pm 0.9$ min) than in the traffic case ($T_{\rm ON} = 15.2 \pm 0.6$ min); the OFF time appears shorter in the single RNAP case ($T_{\rm OFF} = 4.7 \pm 0.3$ min) than in the traffic case ($T_{\rm OFF} = 6.8 \pm 0.8$ min). Note that the system parameters (see table 1 and SI table S1) were incorporated to be consistent with that from the *in vitro* measurements of T7 RNAP [6], except that we set $\alpha^* = 4$ instead of $\alpha^* = 9$ then to mimic the *in vivo* condition. The observations show that the RNAP traffic biases the transcription into the OFF period, and the transcription bursts appear more significant in the traffic case.

Notably, one sees that the distributions of the duration times of the ON and OFF periods, in particular that in the OFF period, do not fit well in the log linear fittings (see figures 4(C) and (D)), which indicate that the underlying processes deviate from the Poisson process, or say, the two-state model do not work well for the above transcription activities. The results thus confirm that the simulated time series from the current model contain multi-state characteristics easy for experimental detections.

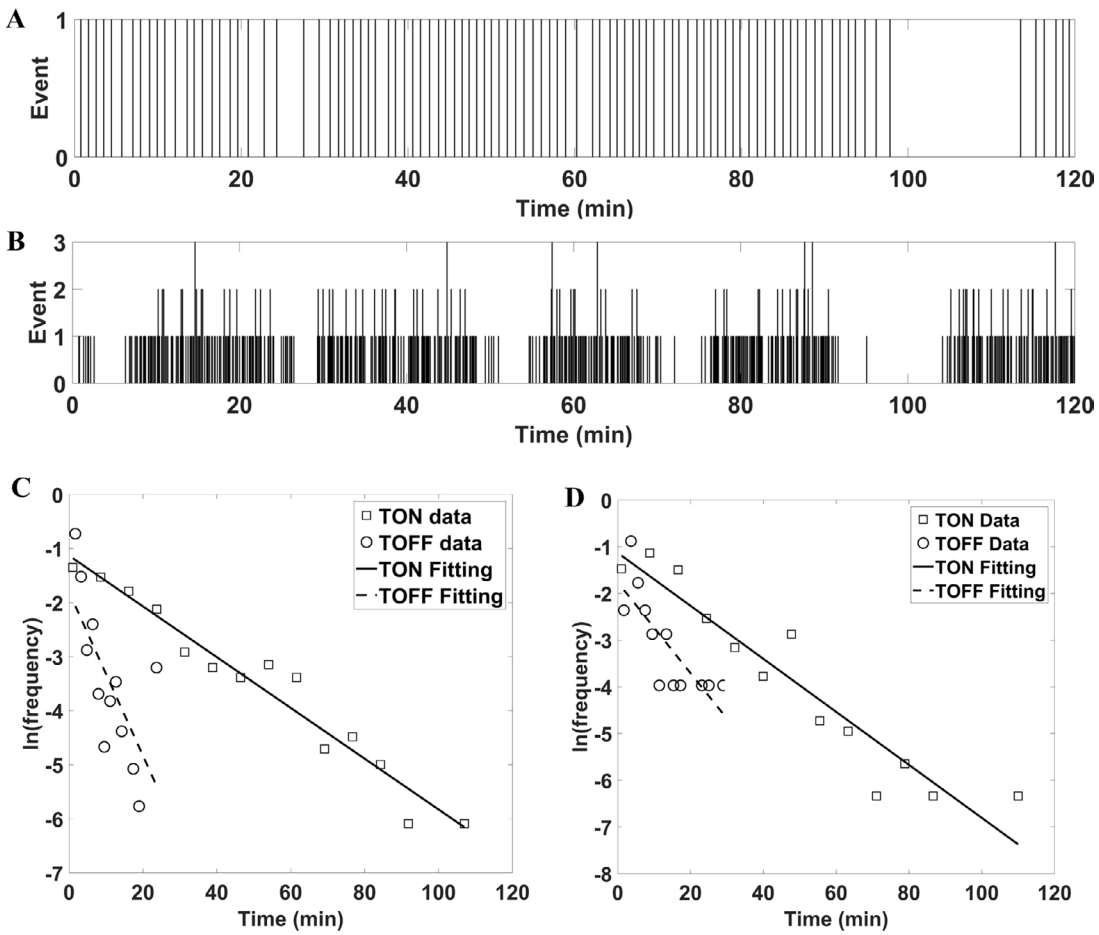


**Figure 4.** Measuring the ON and OFF duration times in the simulated time series of the mRNA production. The time series in the transcription of single T7 RNAP and RNAP traffic (with 14 RNAPs in the cell) are shown in (A) and (B), respectively. The DNA length is $L = 4500$ bp and $\alpha^* = 4$. (C) The fittings of the ON and OFF duration times from the trajectories as shown in (A), with the duration time frequencies shown in logarithmic values. The $R^2$ of the linear fittings for the log values of frequencies of the ON and OFF times are 0.95 and 0.44, respectively, with the average duration times $T_{ON} = 19.5 \pm 0.25$ min and $T_{OFF} = 4.7 \pm 0.17$ min. (D) The fittings for the log values of frequencies of the ON and OFF duration time from the trajectories as shown in (B). The $R^2$ of the respective linear fittings are 0.89 and 0.60, with the average duration times $T_{ON} = 15.2 \pm 0.20$ min and $T_{OFF} = 6.8 \pm 0.28$ min (with the standard errors). The bin size used in the frequency counting is 2 min. 100 trajectories of 200 min each were used to obtain the statistics.

### Supercoiling regulation increases the variations of the number of working RNAPs for long DNA transcription

In the RNAP traffic case, the number of RNAPs working simultaneously on DNA, i.e. the working #RNAP, is a stochastic variable, which is affected by the RNAP concentration ([RNAP]) in the cell, how fast an RNAP opens the promoter and then escapes to elongation (effectively by $k_{init}$), by how fast the RNAP elongates ($k_e$) along with the DNA length ($l$), and also by the dimension of the RNAP ($r$). Without the DNA supercoiling, the working #RNAP demonstrates a single peak distribution (gray histograms in figure 5). The higher the [RNAP] in the cell, the larger the values of $k_{init}$ and $l$, the more the working RNAPs on the DNA. On the other hand, the larger values of $k_e$ and $r$ lead to the fewer working RNAPs.

In the presence of the supercoiling accumulation and releasing, the distributions of the working #RNAP expand to low value ranges. Except for a peak at zero (i.e. in the OFF period), the distribution expansion is not significant for the short or fast DNA transcription (see figures 5(A)–(C)). In the slow transcription with comparatively long DNA, e.g. $L = 4500$ bp for the *E. coli* RNAP, the working #RNAP without the supercoiling is ~40 ± 5, while the distribution expands to a full range of <45 under the supercoiling regulation (see figure 5(D)). Note that in the default setting, there are 100 RNAPs per cell (or [RNAP] ~ 270 nM) in our simulation. It turns out that only a portion of these RNAPs are able to be simultaneously working on the DNA, i.e. the maximum working #RNAP is still significantly lower than the total #RNAP in the cell. In the presence of the supercoiling accumulation, there would be fewer RNAPs loaded onto DNA due to the initiation inhibition, though the elongation inhibition would accommodate slightly more RNAPs simultaneously working. Overall, a significant variation of the working #RNAP appears under the supercoiling regulation in the long DNA transcription.

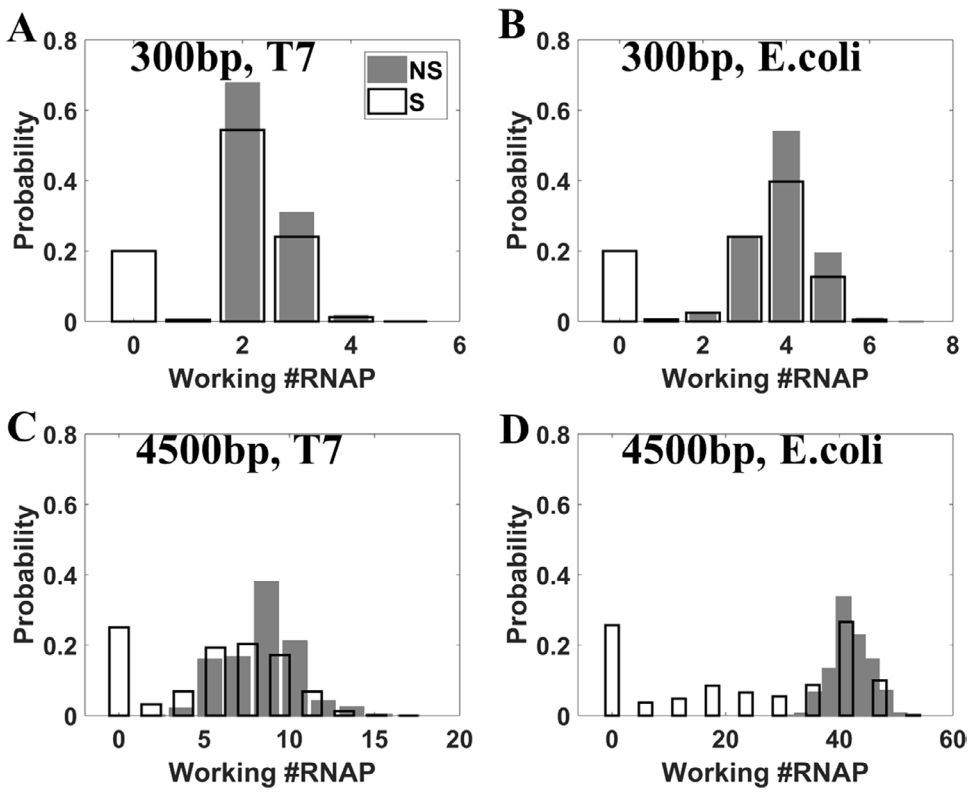


**Figure 5.** The distributions of the working #RNAPs on DNA ([RNAP] = 270 nM, or ~100 RNAPs in cell), without the supercoiling (NS, gray bars) and with supercoiling (S, black-line bars) for four settings: (A) 300 bp DNA, T7 RNAP; (B) 300 bp DNA, *E. coli* RNAP; (C) 4500 bp DNA, T7 RNAP; (D) 4500 bp DNA, *E. coli* RNAP.

**The RNAP traffic increases the mRNA production and also enhances the supercoiling accumulation to inhibit the mRNA production**

As expected, the average #mRNA produced ($\langle m \rangle$) increases with the #RNAPs in the cell or [RNAP], but soon reaches a plateau once the working #RNAP on DNA does not increase any further (see figure 6). T7 RNAP transcribes much faster than *E. coli* RNAP, hence, leading to higher $\langle m \rangle$. The presence of the supercoiling inhibits the mRNA production particularly in the RNAP traffic case, as long as there are more than ~10 RNAPs in the cell.

Specifically, one notices that in the absence of the supercoiling, the same RNAP species transcribing short and long DNA leads to the same amount of mRNA or $\langle m \rangle$ at the steady state, despite different working #RNAP in the short and long gene cases (e.g. 2 and 10 T7 RNAPs on the short and long DNA at saturation and both produce ~75 transcripts; 4 and 40 *E. coli* RNAPs on the short and long DNA and both produce ~50 transcripts). This is because the same species of RNAPs use the same amount of the initiation time on average, or say, the mean duration time between two adjacently produced transcripts keeps the same for the same RNAP species. The presence of the supercoiling, however, inhibits the mRNA production a bit more and thus leads to a slightly lower $\langle m \rangle$ in the long DNA transcription, owning to the higher level of supercoiling accumulation on the long DNA.

**RNAP traffic significantly enhances the noises in the presence of supercoiling**

For the stochastic production of mRNA, one can measure the Fano factor as the ratio between the variance and the average of #mRNA, so that to quantify the fluctuation level and determine how much the production resembles the Poisson process. In case that the transcription is modeled by a single rate-limiting transition, as for the Poisson process, the *Fano* equals to 1. In the simulations of stepwise transcription initiations followed by repetitive elongation cycles (see figure 1(B)), without considering the supercoiling, one obtains the *Fano* consistently smaller than 1 no matter in the single RNAP or the traffic transcription condition (e.g. ~0.5–0.7 for T7 RNAP, ~0.5 for *E. coli* RNAP). Accordingly, the transcript mRNA production follows the sub-Poissonian distribution.

In the presence of the supercoiling, however, the *Fano* can become much larger. The *Fano* also increases with the [RNAP] in the cell until it reaches a plateau, as the working #RNAP on DNA cannot increase any more (see figure 7(A)). Besides, one notices that the longer the DNA, and the faster the transcription, the larger the *Fano* converges to (e.g. 6 and 8 for the short and long DNA transcription by T7 RNAP, respectively; 5 and 7 for the short and long DNA transcription by *E. coli* RNAP), as the supercoiling accumulation becomes more significant. We also demonstrate the *Fano* versus $\langle m \rangle$ relationship in figure 7(B). One can see that the Fano factor always increases with the average #mRNA

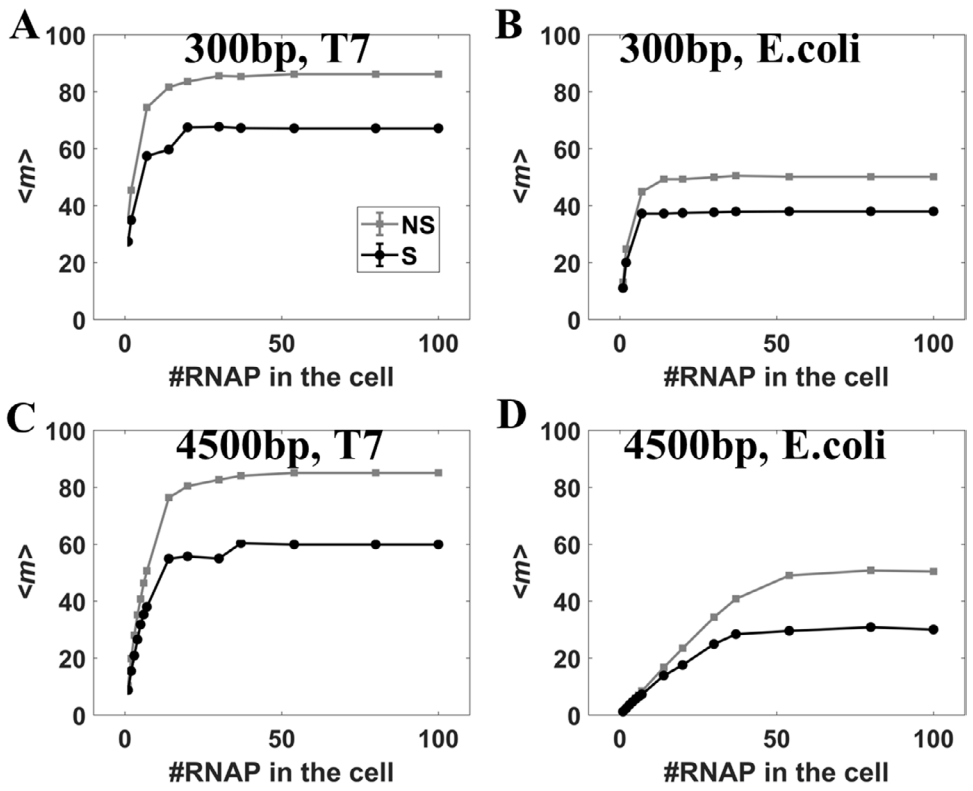


**Figure 6.** The impacts of RNAP traffic on the average mRNA production $\langle m \rangle$ in the no supercoiling case (NS, gray curve) and the supercoiling case (S, black curve) for four settings: (A) 300 bp DNA, T7 RNAP; (B) 300 bp DNA, *E. coli* RNAP; (C) 4500 bp DNA, T7 RNAP; (D) 4500 bp DNA, *E. coli* RNAP. 100 trajectories were used with standard errors provided.

in production, as the #RNAPs increases in the cell and becomes then saturated on the DNA.

### RNAP traffic on long gene transcription bias the transcription to the transcription initiation off status

For the RNAP transcription, one can also measure the intrinsic duty cycle ratio *DCR* (see detailed calculations in SI), which characterizes how much percentile the RNAP engages in the transcription initiation *on* status. In the presence of the supercoiling, the *DCR* decreases with [RNAP] before reaching to a plateau, where the working #RNAP on DNA cannot increase further. Hence, the RNAP traffic essentially biases the transcription toward the *off* status due to the enhanced supercoiling inhibition. Accordingly, the long DNA transcription would bias more toward the *off* status in the high traffic condition as a comparatively large population of working RNAPs are accommodated on the DNA (e.g. the *DCR* in the short DNA case is larger than $g_{on}/(g_{on}+g_{off}) \sim 0.75$, in the long DNA case becoming smaller than 0.75; see figure 8(A)). When there are a small number of working RNAPs, the *DCR* can reach high: For example, in the case of *E. coli* RNAP transcribing $L = 4500$ bp DNA, $DCR > 0.90$ for #RNAP $< 5$ in the cell. When [RNAP] keeps low, the long DNA transcription actually leads to larger *DCR* than the short DNA transcription, as supercoiling accumulation proceeds comparatively slower in the long DNA case.

By varying $\alpha^*$, the *DCR* obtained at different conditions change in similar trends (see figure 8(B)): the larger $\alpha^*$ leads to the higher *DCR*. With the increasing of $\alpha^*$, it becomes harder to accumulate high enough supercoiling level to inhibit the transcription initiation. That is to say, the less sensitive the initiation rate is to the supercoiling inhibition, the larger the bias is set toward the initiation *on*.

## Discussion

In current work we developed a multi-state transcription-bursting model in order to accommodate potentially slow and stepwise supercoiling accumulation and releasing processes for bacterial transcription regulation. When topoisomerase I enzymes are present comparatively abundant, only positive supercoiling accumulates substantially in front of the transcribing RNAP(s) on the DNA that is topologically constrained. Consequently, the transcription activities are inhibited by the positive supercoiling accumulation. We consider that the transcription slowing down proceeds in synchrony with the supercoiling density building up on the DNA, which rises for one level at a time upon each mRNA transcript production. As a result, the supercoiling accumulation proceeds stepwise and takes a comparable amount of time as the RNAP elongation. The process can thus be regarded slow (e.g. from several to tens of minutes), especially for long transcript or slow elongation.

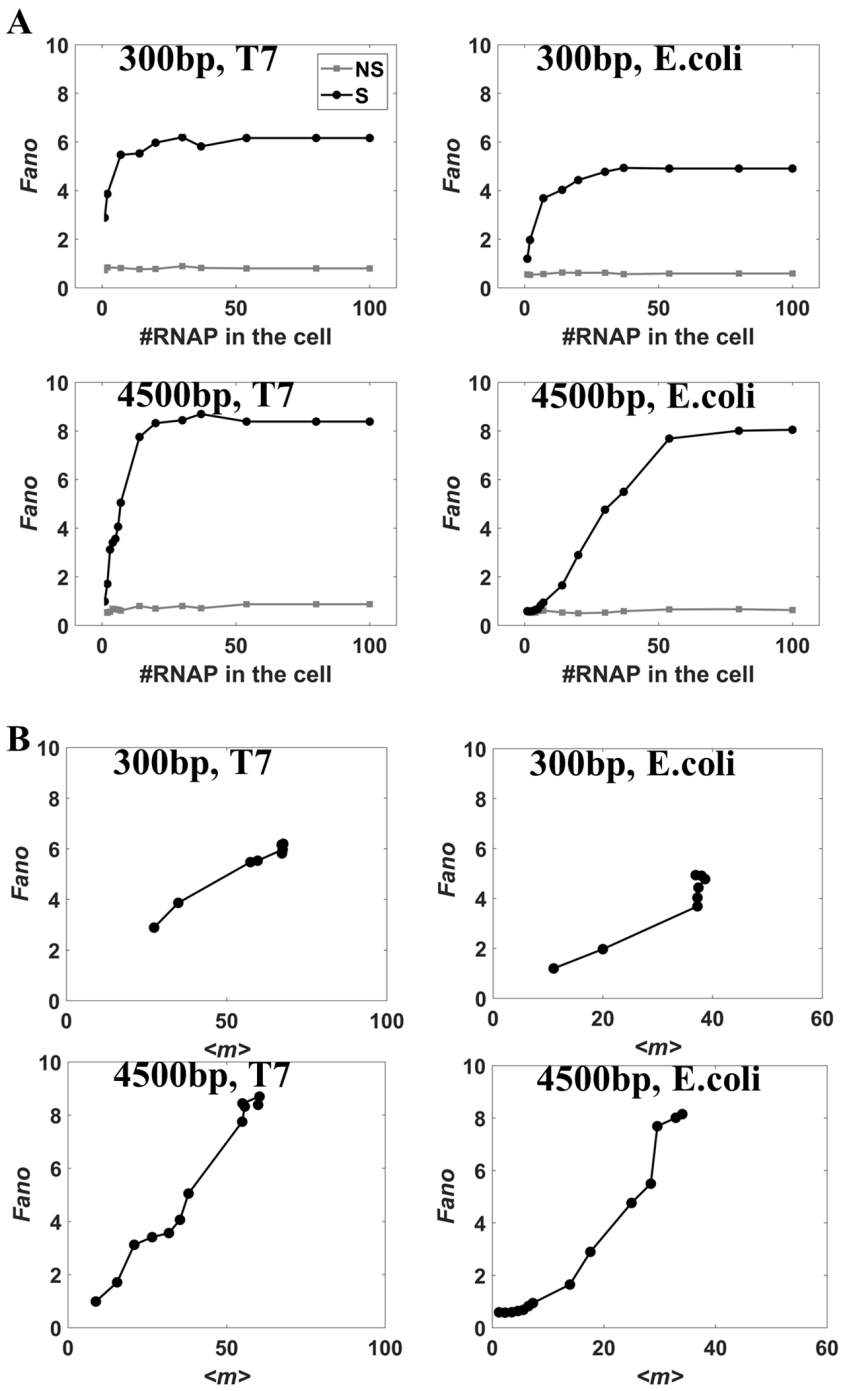


**Figure 7.** (A) The RNAP traffic impacts on the Fano factor in the no supercoiling (NS, gray) and the supercoiling case (S, black) for four settings: 300 bp DNA, T7 RNAP; 300 bp DNA, *E. coli* RNAP; 4500 bp DNA, T7 RNAP; 4500 bp DNA, *E. coli* RNAP. (B) The relationship between the *Fano* and $\langle m \rangle$ (as obtained in figure 6, i.e. by increasing the #RNAP in the cell) in the corresponding settings with the supercoiling regulation.

In the case of high RNAP traffic, the supercoiling impacts become even more significant. The release of the supercoiling also proceeds comparatively slowly for processive gyrase actions on the DNA, e.g. over minutes for a long transcript [25]. As the timescales of supercoiling accumulation and relaxation approach to that of the gyrase binding and unbinding (e.g. minutes to tens of minutes), the two-state model that requires single slow event to dominate the gene ON/OFF switching does not work well any more. Our multi-state model accordingly accommodates the more general situation.

The roles of RNAP convoys in transcription and bursting have been concerned lately [40, 41]. According to our study, the convoy or traffic of RNAPs turns out to substantially enhance the fluctuation of the mRNA production as the transcription subjects to the supercoiling regulation. Although multiple simultaneously working RNAPs improve the mRNA production, they also promote the supercoiling accumulation to inhibit the transcription, and the inhibition becomes more significant with an increasing number of RNAPs working together on the DNA until saturation. Accordingly, the Fano factor of the mRNA production,

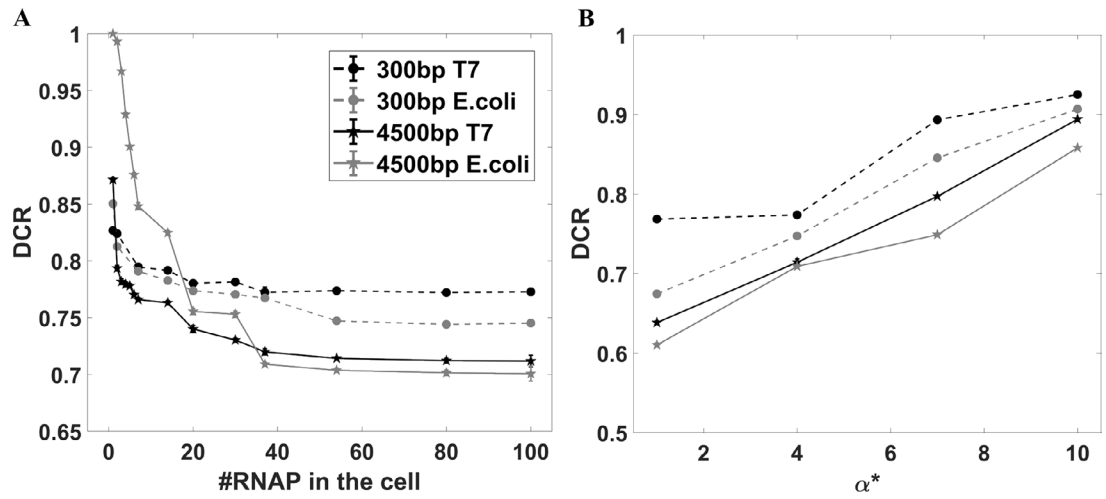


**Figure 8.** The variations of the duty cycle ratio *DCR*. (A) The *DCR* versus the RNAP traffic in the supercoiling case for four settings. (B) The *DCR* variations with $\alpha^*$ in the traffic cases. The #RNAP in the cell is set at 100 or say [RNAP] ~ 270 nM. The standard errors of *DCR* for four settings were provided.

which is boosted over one due to the transcription bursting under supercoiling regulation, can be maximally enhanced (e.g. to 5–10) for a battery of RNAPs in dense traffic.

Underlying, the supercoiling regulation can substantially enhance the variations of the working #RNAPs on DNA. The working #RNAPs at saturation on the DNA can be significantly lower than the total amount of RNAP molecules in the cell and is affected by several factors: It is controlled by how fast an RNAP is loaded and occupies on the promoter for initiation, i.e. the larger the effective initiation rate ($k_{init}$), the more RNAPs simultaneously work on the same DNA; it is also affected by how long the transcription elongation lasts, i.e. the larger the length of gene ($L$), and the lower the elongation rate ($k_e$), the more RNAPs can sustain on DNA. Without considering the pausing activities, one finds the average working #RNAPs scales with $L\,k_{init}/k_e$. In the presence of the topological constraint and supercoiling accumulation, both the rates of the initiation and elongation drop accordingly. Since the initiation appears to be more sensitive to the tension on DNA due to the supercoiling accumulation, $k_{init}$ decreases more quickly than $k_e$ till a full initiation suspension. Correspondingly, the average working #RNAP decreases upon the supercoiling accumulation, while the overall distributions of the working #RNAP can be significantly expanded for the long transcript or slow elongation.

Putting together, one can see that the supercoiling and RNAP traffic impact on each other to regulate the overall mRNA productions and fluctuations. It can be counter intuitive since a large number of RNAPs working together does not quench the transcription noises. Instead, due to the correlation between the RNAP traffic and the supercoiling accumulation that serves as a DNA mechanical feedback to the transcription, the transcription productivities become more versatile in the traffic condition. Consequently, by varying the #RNAPs on DNA or in the cell, one obtains the positive correlation between the Fano factor and the average amount of mRNA in the steady-state production, which turns out to be another gene-nonspecific feature in the mRNA production [42]. The positive correlation between the mRNA production and the Fano factor or fluctuation can also reveal by directly speeding up the transcription or slowing down the mRNA degradation. Modulating only the supercoiling regulation kinetics, e.g. by accelerating the gyrase association, deaccelerating the gyrase dissociation with the DNA, or enabling sufficiently fast supercoiling releasing, however, one would obtain improved mRNA productions with quenched fluctuations. Note that the *Fano* and $\langle m\rangle$ characteristics for the single RNAP transcription (in figure 3(A)) maintain the same trends for the traffic condition (results not shown).

For a battery of T7 or *E. coli* RNAPs transcribing in tandem considered in current setting, occasional collisions due to stochasticity or RNAP pausing (*E. coli*) do not seem to interfere much the overall traffic. One may actually regard a battery of RNAPs in dense traffic as an integrated transcription machinery. On average, these RNAPs move one behind another to avoid supercoiling building up within the pack, so that only a small amount of gyrase molecules are needed to remove the positive supercoiling in front of the very leading RNAP. In current model, we assume that the overall tension generated on DNA under the supercoiling condition essentially slows down the transcription initiation and elongation. Hence, even the positive supercoiling is geometrically restricted to the very downstream region in front of the RNAP battery, the tension persists all along the DNA to impact on the RNAP transcription.

The duty cycle ratio *DCR* turns out to be insensitive to the #RNAP in the dense traffic, while in the low traffic condition, it decrease largely with an increasing #RNAP, in particular, for the slow transcription

of long DNA. Note that the *DCR* in current work was evaluated according to the probability of the RNAP transcription initiation *on* status. For a small number of RNAPs transcribing slowly to generate a long transcript, supercoiling hardly accumulates soon to inhibit the transcription, hence, there is almost no *off* status and the *DCR* reaches quite high (e.g. lager than 95%). Note that the comparatively low *DCR* for the dense RNAP traffic can improve, however, with an increasing value of the maximum supercoiling level ($\alpha^*$) to suspend the transcription initiation, which corresponds to lowering the sensitivity of the promoter to the supercoiling inhibition. That says, a less tension-sensitive promoter encourages a high *DCR* for the RNAP traffic machinery.

Note that some of properties demonstrated in current model rely on how the RNAP elongation and initiation react to the DNA tension during the supercoiling regulation. We assume that the RNAP elongation rate is less sensitive to the tension than the initiation rate, according to experimental measurements so far. During an elongation cycle of T7 or *E. coil* RNAP, it had been found that the translocation is unlikely a rate limiting step, so that the force implementation on DNA that deters the RNAP translocation cannot significantly change the overall elongation rate. On the other hand, the tension on DNA appears to be easily detected by the promoter, such that a comparatively small force ($<$10 pN) can hinder the RNAP initiation, due to the force sensitive promoter opening or the RNAP 'abortive escaping'. Accordingly, in current model, the elongation rate drops to a small but nonzero value at the time of the initiation suspension or the *off* status. As a result, there exists a period of time during which a small number of RNAPs loaded before the initiation suspension continue elongating on the DNA, prior to a complete transcription OFF. The low-rate elongation persisting after the initiation turning *off* can nevertheless be detected as ON, due to a small amount of transcripts produced by those left RNAPs on the DNA.

Besides, we assume in current model that the elongation and initiation slow down constantly upon each transcript production, as the supercoiling density keeps building up. Indeed, how fast the supercoiling generates and propagates on DNA during the transcription, how exactly an elongating or initiating RNAP reacts to the supercoiling accumulation, and how exactly the supercoiling is resolved upon the gyrase action, remain physically elusive. In addition, for high concentrations of topoisomerase I, some of RNAPs may have negative supercoiling upstream being resolved immediately, so that positive supercoiling may prevail in front of quite many RNAPs in the traffic. Consequently, there can be several groups of RNAPs occupying differently localized supercoiling regions on the same DNA. The supercoiling regulation on individual RNAPs in a battery or convoy of RNAPs may become more involved. All these aspects deserve further investigations.

## Conclusion

By building up a multi-state model of bacteria transcription bursting, we are able to describe potentially slow and stepwise supercoiling accumulation and releasing processes in a general quantitative framework. In particular, we address for a battery of RNAPs transcribing in tandem under the positive supercoiling regulation by employing the multi-state model. It is found that supercoiling regulation enhances variations of the number of RNAPs working together on DNA, so that the traffic of RNAPs amplifies the supercoiling impacts to the transcription. Consequently, the RNAP traffic promotes the mRNA production as well as enhances the mRNA fluctuations, leading to the highly pronounced transcriptional bursting. The positive correlation between the average mRNA production and fluctuations under the RNAP traffic condition thus reveals as one of generic features in the gene transcription. In comparison, modulating gyrase kinetic impacts on the DNA supercoiling may improve the mRNA production while quench the fluctuations. Furthermore, the RNAPs in dense traffic resemble an integrated transcription machinery with a duty cycle ratio between the initiation *on* and *off* maintained low but independent of the number of RNAPs. To verify current model, it is essential to physically determine the supercoiling dynamics along DNA together with the mechanical responses of RNAPs in the transcription bursting process.

## Acknowledgments

This work is supported by NSFC #11775016, #11635002, and #11275022. We acknowledge the computational support from the Beijing Computational Science Research Center (CSRC).

## ORCID iDs

Xiaobo Jing https://orcid.org/0000-0002-2360-0258
Jin Yu https://orcid.org/0000-0001-8224-1374

## References

[1] Borukhov S and Nudler E 2008 RNA polymerase: the vehicle of transcription *Trends Microbiol.* **16** 126–34
[2] Buc H C and Strick T 2009 *RNA Polymerases as Molecular Motors* vol 16 (Cambridge: Royal Society of Chemistry) (https://doi.org/10.1039/9781847559982)
[3] Liu L F and Wang J C 1987 Supercoiling of the DNA template during transcription *Proc. Natl Acad. Sci. USA* **84** 7024–7
[4] Kouzine F and Levens D 2007 Supercoil-driven DNA structures regulate genetic transactions *Frontiers Biosci.* **12** 4409–23
[5] Dorman C J and Dorman M J 2016 DNA supercoiling is a fundamental regulatory principle in the control of bacterial gene expression *Biophys. Rev.* **8** 89–100
[6] Chong S, Chen C, Ge H and Xie X S 2014 Mechanism of transcriptional bursting in bacterial *Cell* **158** 314–26

[7] Baranello L, Kouzine F and Levens D 2013 DNA topoisomerases beyond the standard role *Transcription* **4** 232–7
[8] Guptasarma P 1996 Cooperative relaxation of supercoils and periodic transcriptional intitiation with polymerase batteries *BioEssays* **18** 325–32
[9] Golding I, Paulsson J, Zawilski S M and Cox E C 2005 Real-time kinetics of gene activity in individual bacteria *Cell* **123** 1025–36
[10] Friedman N, Cai L and Xie X S 2006 Linking stochastic dynamics to population distribution: an analytical framework of gene expression *Phys. Rev. Lett.* **97** 168302
[11] Ge H, Qian H and Xie X S 2015 Stochastic phenotype transition of a single cell in an intermediate region of gene state switching *Phys. Rev. Lett.* **114** 078101
[12] Sevier S A, Kessler D A and Levine H 2016 Mechanical bounds to transcriptional noise *Proc. Natl Acad. Sci. USA* **113** 13983–8
[13] Bohrer C H and Roberts E 2016 A biophysical model of supercoiling dependent transcription predicts a structural aspect to gene regulation *BMC Biophys.* **9** 2
[14] Brackley C A, Johnson J, Bentivoglio A, Corless S, Gilbert N, Gonnella G and Marenduzzo D 2016 Stochastic model of supercoiling-dependent transcription *Phys. Rev. Lett.* **117** 018101
[15] Kar P, Cherstvy A G and Metzler R 2018 Acceleration of bursty multiprotein target search kinetics on DNA by colocalisation *Phys. Chem. Chem. Phys.* **20** 7931–46
[16] Skinner G M, Kalafut B S and Visscher K 2011 Downstream DNA tension regulates the stability of the T7 RNA polymerase initiation complex *Biophys. J.* **100** 1034–41
[17] Thomen P, Lopez P J, Bockelmann U, Guillerez J, Dreyfus M and Heslot F 2008 T7 RNA polymerase studied by force measurements varying cofactor concentration *Biophys. J.* **95** 2423–33
[18] Wang M D, Schnitzer M J, Yin H, Landick R, Gelles J and Block S M 1998 Force and velocity measured for single molecules of RNA polymerase *Science* **282** 902–7
[19] Forde N R, Izhaky D, Woodcok G R, Wuite G J L and Bustamante C 2002 Using mechanical force to probe the mechanism of pausing and arrest during continuous elongation by *Escherichia coli* RNA polymerase *Proc. Natl Acad. Sci. USA* **99** 11682–7
[20] Abbondanzieri E A, Greenleaf W J, Shaevitz J W, Landick R and Block S M 2005 Direct observation of base-pair stepping by RNA polymerase *Nature* **438** 460–5
[21] Tang G Q, Roy R, Bandwar R P, Ha T and Patel S S 2009 Real-time observation of the transition from transcription initiation to elongation of the RNA polymerase *Proc. Natl Acad. Sci. USA* **106** 22175–80
[22] Skinner G M, Baumann C G, Quinn D M, Molloy J E and Hoggett J G 2004 Promoter binding, initiation, and elongation by bacteriophage T7 RNA polymerase *J. Biol. Chem.* **279** 3239–44
[23] Saecker R M, Record M T Jr and Dehaseth P L 2011 Mechanism of bacterial transcription initiation: RNA polymerase-promoter binding, isomerization to initiation-competent open complexes, and initiation of RNA synthesis *J. Mol. Biol.* **412** 754–71
[24] Levens D, Baranello L and Kouzine F 2016 Controlling gene expression by DNA mechanics: emerging insights and challenges *Biophys. Rev.* **8** S23–32
[25] Gore J, Bryant Z, Stone M D, Nollmann M, Cozzarelli N R and Bustamante C 2006 Mechanochemical analysis of DNA gyrase using rotor bead tracking *Nature* **439** 100–4
[26] Brown P O and Cozzarelli N R 1979 A sign inversion mechanism for enzymatic supercoiling of DNA *Science* **206** 1081–3
[27] Cheng B, Zhu C-X, Ji C, Ahumada A and Tse-Dinh Y-C 2003 Direct interaction between *Escherichia coli* RNA polymerase and the zinc ribbon domains of DNA topoisomerase I *J. Biol. Chem.* **278** 30705–10
[28] Zechiedrich E L, Khodursky A B, Bachellier S, Schneider R, Chen D, Lilley D M and Cozzarelli N R 2000 Roles of topoisomerases in maintaining steady-state DNA supercoiling in *Escherichia coli J. Biol. Chem.* **275** 8103–13
[29] Thomen P, Lopez P J and Heslot F 2005 Unravelling the mechanism of RNA-polymerase forward motion by using mechanical force *Phys. Rev. Lett.* **94** 128102
[30] Ma J, Bai L and Wang M D 2013 Transcription under torsion *Science* **340** 1580–3
[31] Czyz A, Mooney R A, Iaconi A and Landick R 2014 Mycobacterial RNA polymerase requires a U-tract at intrinsic terminators and is aided by NusG at suboptimal terminators *MBio* **5** e00931
[32] Reznikoff W and Gold L 1986 *Maximizing Gene Expression* vol 9 (Amsterdam: Elsevier)
[33] Kim P and Lee C H 2014 Fast probability generating function method for stochastic chemical reaction networks *MATCH Commun. Math. Comput. Chem.* **71** 57–69
[34] Thattai M and Van Oudenaarden A 2001 Intrinsic noise in gene regulatory networks *Proc. Natl Acad. Sci.* **98** 8614–9
[35] Howard J 2001 *Mechanics of Motor Proteins and the Cytoskeleton* (Sunderland, MA: Sinauer) ch 13 pp 221–5
[36] Zhou Y and Martin C T 2006 Observed instability of T7 RNA polymerase elongation complexes can be dominated by collision-induced 'bumping'∗ *J. Biol. Chem.* **281** 24441–8
[37] Epshtein V and Nudler E 2003 Cooperation between RNA polymerase molecules in transcription elongation *Science* **300** 801–5
[38] Herbert K M, Greenleaf W J and Block S M 2008 Single-molecule studies of RNA polymerase: motoring along *Annu. Rev. Biochem.* **77** 149–76
[39] Battaile C C 2008 The kinetic Monte Carlo method: foundation, implementation, and application *Comput. Methods Appl. Mech Eng.* **197** 3386–98
[40] Lesne A, Victor J-M, Bertrand E, Basyuk E and Barbi M 2018 The role of supercoiling in the motor activity of RNA polymerases *Molecular Motors* (New York: Humana Press) pp 215–32
[41] Tantale K, Mueller F, Kozulicpirher A, Lesne A, Victor J M, Robert M, Capozi S, Chouaib R, Bäcker V and Mateoslangerak J 2016 A single-molecule view of transcription reveals convoys of RNA polymerases and multi-scale bursting *Nat. Commun.* **7** 12248
[42] Sanchez A and Golding I 2013 Genetic determinants and cellular constraints in noisy gene expression *Science* **342** 1188–93

# Supplementary Information for "How does supercoiling regulation on a battery of RNA polymerases impact on bacterial transcription bursting"

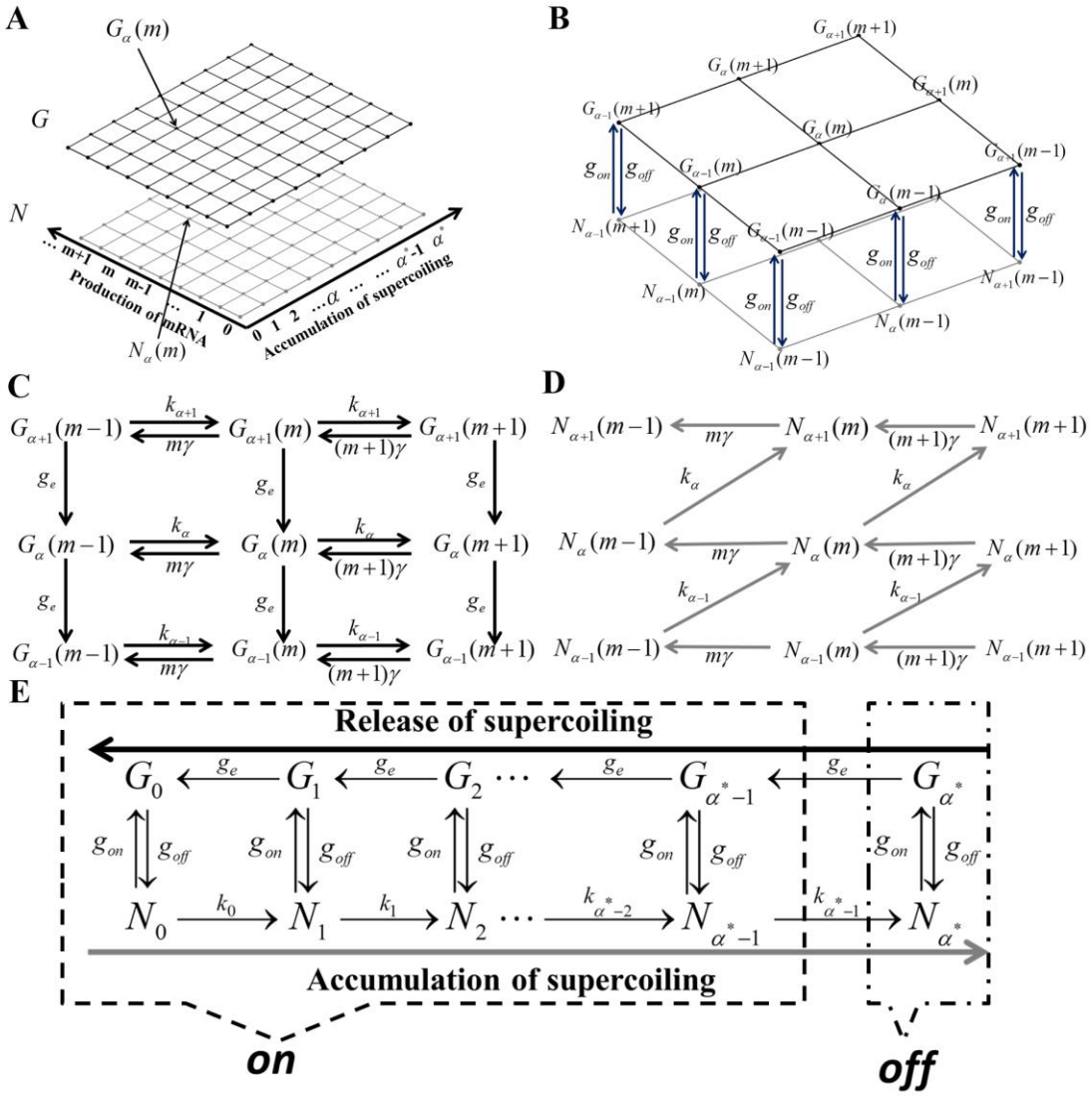


**Fig S1** A detailed scheme for the multi-state transcription-bursting model of single RNAP with stepwise supercoiling accumulation and relaxation. (A) We split the scheme into two layers, with the upper and lower layers representing the gyrase bound (*G*) and gyrase unbound states (*N*), respectively. For each state in the model, the superscript is for the gyrase bound $G_a$ or unbound $N_a$ state, while the subscripti is for the supercoil level, with $a$ ranging from 0, 1, 2, 3 … to $a^*$. $a^*$ is the supercoiling density level at which the transcription initiation is

fully suspended. $G_a(m)$ and $N_a(m)$ denote two status with the copy number *m* of the mRNA transcription produced. (B) The two layers are connected by the gyrase binding and unbinding transitions, with binding/association rate $g_{on}$ and unbinding/dissociation rate $g_{off}$. The gyrase releases supercoiling with a relaxation rate $g_e$ on the *top* layer (C) while supercoiling accumulation builds up in syncrhony with the mRNA production on the *bottom* layer upon the gyrase unbinds (D). $g$ is the degradation rate of mRNA, $k_\alpha$ represents an effective transcription rate at a supercoiling density level $a$. (E) A simplified scheme showing the supercoiling accumulation and releasing via the gyrase unbinding and binding kinetics for single RNAP transcription.

***Model building and analytic solutions to obtain the average number of mRNA (<m>), the Fano factor (Fano), and the effective duty cycle rate (DCR) in the single RNAP transcription***

We used chemical master eqution to describe the dynamic of the above multi-state model (see **Fig S1**), and then solved it at the steady-state condition, so that the interested quantities could be calculated including the average number of mRNA ($\langle m \rangle$), the Fano factor (*Fano*), and the duty cycle ratio (*DCR*).

Here we define $P_{a,b}(m,t), 0 \le a \le a^*, b = 0 \; or \; 1$as the probability for each state in the scheme above with *m* mRNA at time *t* (see **Fig S1E**), where $a$ denotes the supercoiling density level and $b$ labels the gyrase bound status (1 for the bound state *G* , 0 for the unbound state *N*). $P_{a,b}^{\mathbf{SS}}$ represents the steady state value. The corresponding chemical master equations are given by:

$$\begin{cases} \dfrac{\partial P_{a,1}(m,t)}{\partial t} = -g_{off}P_{a,1}(m,t) + g_{on}P_{a,0}(m,t) - (1-\delta_{a0})g_e P_{a,1}(m,t) + (1-\delta_{aa^*})g_e P_{a+1,1}(m,t) + \\ \qquad (1-\delta_{aa^*})k_a P_{a,1}(m-1,t) - (1-\delta_{aa^*})k_a P_{a,1}(m,t) + (m+1)gP_{a,1}(m+1,t) - mgP_{a,1}(m,t) \\ \dfrac{\partial P_{a,0}(m,t)}{\partial t} = -\left((1-\delta_{aa^*})k_a + g_{on}\right)P_{a,0}(m,t) + g_{off}P_{a,1}(m,t) + (1-\delta_{a0})k_{a-1}P_{a-1,0}(m-1,t) + \\ \qquad (m+1)gP_{a,0}(m+1,t) - mgP_{a,0}(m,t) \end{cases} \quad \textbf{(1)}$$

in which $1 \le a \le a^*$ and $\delta_{nm}$ is the Kronecker delta (if $m = n$, $\delta_{nm} = 1$, or $\delta_{nm} = 0$). The first equation of eq (1) suggests that the changing rate of the probability of state

$G_a(m)$ comes from two parts (see **FigS1 B and C**): One is 'death' and the other is 'birth'. The 'death' part includes gyrase unbinding events from $G_a(m)$ to $N_a(m)$ with probability flux $-g_{off}P_{a,1}(m,t)$, supercoiling relaxation event from $G_a(m)$ to $G_{a-1}(m)$ with probability flux $-(1-\delta_{a0})g_eP_{a,1}(m,t)$, new transcription event from $G_a(m)$ to $G_a(m+1)$ with probability flux $-(1-\delta_{aa^*})k_aP_{a,1}(m,t)$, and degradation event from $G_a(m)$ to $G_a(m-1)$ with probability flux $-m\gamma P_{a,1}(m,t)$; the 'birth' part of the state $G_a(m)$ includes the gyrase binding event from $N_a(m)$ to $G_a(m)$ with probability flux $g_{on}P_{a,0}(m,t)$, supercoiling relaxation event from $G_{a+1}(m)$ to $G_a(m)$ with probability flux $(1-\delta_{aa^*})g_eP_{a+1,1}(m,t)$, transcription event from $G_a(m-1)$ to $G_a(m)$ with probability flux $(1-\delta_{aa^*})k_aP_{a,1}(m-1,t)$, and degradation event from $G_a(m+1)$ to $G_a(m)$ with probability flux $(m+1)\gamma P_{a,1}(m+1,t)$. Then the summation of all fluxes is thus the changing rate of the probability of state $G_a(m)$. Similarly, we can obtain the second equation in eq (1).

We next use the generating function method to solve the above equations [1]. We define

$$F_{\alpha,\beta}(z,t)=\sum_{m=0}^{\infty}z^mP_{\alpha,\beta}(m,t), \qquad \textbf{(2)}$$

where $F_{a,b}(z,t)$ is the generating function. Substituting eq (2) into eq (1), we can get

$$\begin{cases} \dfrac{\partial F_{a,1}(z,t)}{\partial t}=-g_{off}F_{a,1}(z,t)+g_{on}F_{a,0}(z,t)-(1-\delta_{a0})g_eF_{a,1}(z,t)+(1-\delta_{aa^*})g_eF_{a+1,1}(z,t)+ \\ \qquad (1-\delta_{aa^*})zk_aF_{a,1}(z,t)-(1-\delta_{aa^*})k_aF_{a,1}(z,t)+\gamma(1-z)\partial_zF_{a,1}(z,t) \\ \dfrac{\partial F_{a,0}(z,t)}{\partial t}=-\left((1-\delta_{aa^*})k_a+g_{on}\right)F_{a,0}(z,t)+g_{off}F_{a,1}(z,t)+z(1-\delta_{a0})k_{a-1}F_{a-1,0}(z,t)+ \\ \qquad \gamma(1-z)\partial_zF_{a,0}(z,t) \end{cases} \qquad \textbf{(3)}$$

We now introduce $F_{a,b}^{\mathbf{SS}}$ as the steady state value of $F_{a,b}$, which corrsponds to

$\frac{\partial F_{\alpha,\beta}(z,t)}{\partial t}=0$. Then we have

$$\begin{cases} -g_{off}F^{SS}_{a,1}(z)+g_{on}F^{SS}_{a,0}(z)-(1-d_{a0})g_eF^{SS}_{a,1}(z)+(1-d_{aa^*})g_eF^{SS}_{a+1,1}(z)+ \\ \quad (1-d_{aa^*})zk_aF^{SS}_{a,1}(z)-(1-d_{aa^*})k_aF^{SS}_{a,1}(z)+g(1-z)\partial_zF^{SS}_{a,1}(z)=0 \\ -\left((1-d_{aa^*})k_a+g_{on}\right)F^{SS}_{a,0}(z)+g_{off}F^{SS}_{a,1}(z)+(1-d_{a0})zk_{a-1}F^{SS}_{a-1,0}(z)+ \\ \quad g(1-z)\partial_zF^{SS}_{a,0}(z)=0 \end{cases} \quad \textbf{(4)}$$

where $1\le a\le a^*$. At $z=1$, we will obtain the corresponding probabilities from eq (4) by using the relationship：

$$P^{SS}_{\alpha,\beta}=F^{SS}_{\alpha,\beta}(1)=\sum_{m=0}^{\infty}P^{SS}_{\alpha,\beta}(m), \quad \textbf{(5)}$$

and

$$\sum_{\beta=0}^{1}\sum_{\alpha=0}^{\alpha^*}P^{SS}_{\alpha,\beta}=1. \quad \textbf{(6)}$$

Consequently, one obtains the steady state probabilities as:

$$\begin{cases} -g_{off}P^{SS}_{a,1}+g_{on}P^{SS}_{a,0}-(1-d_{a0})g_eP^{SS}_{a,1}+(1-d_{aa^*})g_eP^{SS}_{a+1,1}=0 \\ -\left((1-d_{aa^*})k_a+g_{on}\right)P^{SS}_{a,0}+g_{off}P^{SS}_{a,1}+(1-d_{a0})k_{a-1}P^{SS}_{a-1,0}=0 \end{cases} \quad \textbf{(7)}$$

The solution is dependent on $a^*$, and we need to solve $2a^*+2$ equations:

$$\begin{cases} -(k_0+g_{on})P^{SS}_{0,0}+g_{off}P^{SS}_{0,1}=0 \\ -g_{off}P^{SS}_{0,1}+g_{on}P^{SS}_{0,0}+g_eP^{SS}_{1,1}=0 \\ -\left(k_a+g_{on}\right)P^{SS}_{a,0}+k_{a-1}P^{SS}_{a-1,0}+g_{off}P^{SS}_{a,1}=0, \quad 0<a<a^* \\ -g_{off}P^{SS}_{a,1}+g_{on}P^{SS}_{a,0}-g_eP^{SS}_{a,1}+g_eP^{SS}_{a+1,1}=0, \quad 0<a<a^* \\ -g_{on}P^{SS}_{a^*,0}+g_{off}P^{SS}_{a^*,1}+k_{a^*-1}P^{SS}_{a^*-1,0}=0 \\ g_{on}P^{SS}_{a^*,0}-g_{off}P^{SS}_{a^*,1}-g_eP^{SS}_{a^*,1}=0 \end{cases} \quad \textbf{(8)}$$

We can then solve the above equation numerically upon setting the rates and $a^*$.

After obtaining the probabilities of each state in the system, we now introduce

$F(z,t) \equiv \sum_{\beta=0}^{1}\sum_{\alpha=0}^{\alpha^*} F_{\alpha,\beta}(z,t)$ to calculate the average number of mRNA.

Summing up all equations in eq (4), we get

$$\partial_z F^{\mathbf{SS}}(z) = \sum_{m=0}^{\infty} m z^{m-1} P_{a,b}^{\mathbf{SS}}(m) = \sum_{a=0}^{a^*} \frac{k_a}{g}\left(\sum_{b=0}^{1} F_{a,b}^{\mathbf{SS}}(z)\right). \quad \textbf{(9)}$$

Hence, the average #mRNA, denoted as <*m*> here, is obtained as

$$\langle m \rangle = \partial_z F^{\mathbf{SS}}(z)|_{z=1} = \sum_{a=0}^{a^*} \frac{k_a}{g}\left(\sum_{b=0}^{1} P_{a,b}^{\mathbf{SS}}\right). \quad \textbf{(10)}$$

The Fano factor (*Fano*) is an important index to measure the deviation of the distribution of the mRNA number from the Poisson distribution [2]. The value of the *Fano* is defined as the ratio of the variance and the average value. Since we have obtained the average value $\langle m \rangle$, we need to have the variance $\mathrm{Var} \equiv \langle m^2 \rangle - \langle m \rangle^2$ to obtain the *Fano*. To get the second moment of the #mRNA, we want to calculate the second derivative of $F(z,t)$ from eq (9), which is

$$\partial_z^2 F^{\mathbf{SS}}(z) = \sum_{a=0}^{a^*} \frac{k_a}{g}\left(\sum_{b=0}^{1} \partial_z F_{a,b}^{\mathbf{SS}}(z)\right). \quad \textbf{(11)}$$

At $z=1$, we have the below relationship,

$$\partial_z^2 F^{\mathbf{SS}}(z)|_{z=1} = < m >^2 - \langle m \rangle = \sum_{a=0}^{a^*} \frac{k_a}{g}\left(\sum_{b=0}^{1} < m_{a,b} >\right), \quad \textbf{(12)}$$

where

$$< m_{\alpha,\beta} > = \partial_z F_{\alpha,\beta}^{\mathbf{SS}}(z)|_{z=1}, \quad \textbf{(13)}$$

which is an intermediate variable for computing. Then the *Fano* can be calculated by

$$
\begin{aligned}
Fano &= \frac{\mathrm{Var}}{<m>} = \frac{<m^2> - <m>^2}{\langle m \rangle} = \frac{<m^2> - <m> + <m> - <m>^2}{\langle m \rangle} \\
&= \frac{\sum_{a=0}^{ST} \frac{k_a}{g} \left( \sum_{b=0}^{1} <m_{a,b}> \right) + <m> - <m>^2}{<m>} \\
&= 1 + \left( \frac{\sum_{a=0}^{ST} \frac{k_a}{g} \left( \sum_{b=0}^{1} <m_{a,b}> \right)}{<m>} - <m> \right)
\end{aligned}
\tag{14}
$$

As we have $<m>$, it is then vital that we obtain $<m_{a,b}>, 0 \le a \le a^*, b = 0 \; or \; 1$.

Using the partial derivative of eq (4) for $z$ and eq (10) and (11), $<m_{\alpha,\beta}>$ can be sloved then from the following equations:

$$
\begin{cases}
-g_{off} \partial_z F^{SS}_{a,1}(z,t) + g_{on} \partial_z F^{SS}_{a,0}(z,t) - (1-\delta_{a0}) g_e \partial_z F^{SS}_{a,1}(z,t) + (1-\delta_{aa^*}) g_e \partial_z F^{SS}_{a+1,1}(z,t) \\
\quad + (1-\delta_{aa^*}) z k_a \partial_z F^{SS}_{a,1}(z,t) + (1-\delta_{aa^*}) k_a F^{SS}_{a,1}(z,t) - (1-\delta_{aa^*}) k_a \partial_z F^{SS}_{a,1}(z,t) \\
\quad + g(1-z) \partial_z^2 F^{SS}_{a,1}(z,t) - g \partial_z F^{SS}_{a,1}(z,t) = 0 \\
-\left( (1-\delta_{aa^*}) k_a + g_{on} \right) \partial_z F^{SS}_{a,0}(z,t) + g_{off} \partial_z F^{SS}_{a,1}(z,t) + (1-\delta_{a0}) z k_{a-1} \partial_z F^{SS}_{a-1,0}(z,t) \\
\quad + (1-\delta_{a0}) k_{a-1} F^{SS}_{a-1,0}(z,t) + g(1-z) \partial_z^2 F^{SS}_{a,0}(z,t) - g \partial_z F^{SS}_{a,0}(z,t) = 0
\end{cases}
\tag{15}
$$

At $z = 1$, we have

$$
\begin{cases}
-g_{off} <m_{\alpha,1}> + g_{on} <m_{\alpha,0}> - (1-\delta_{\alpha 0}) g_e <m_{\alpha,1}> + (1-\delta_{\alpha\alpha^*}) g_e <m_{\alpha+1,1}> + \\
\quad (1-\delta_{\alpha\alpha^*}) z k_\alpha <m_{\alpha,1}> + (1-\delta_{\alpha\alpha^*}) k_\alpha P_{\alpha,1} - (1-\delta_{\alpha\alpha^*}) k_\alpha <m_{\alpha,0}> - \gamma <m_{\alpha,1}> = 0 \\
-\left( (1-\delta_{\alpha\alpha^*}) k_\alpha + g_{on} \right) <m_{\alpha,0}> + g_{off} <m_{\alpha,1}> + (1-\delta_{\alpha 0}) z k_{\alpha-1} <m_{\alpha-1,0}> + (1-\delta_{\alpha 0}) k_{\alpha-1} P_{\alpha-1,0} \\
\quad - \gamma <m_{\alpha,0}> = 0
\end{cases}
\tag{16}
$$

Solving the linear equations and substuting the average values into eq (14), we can finally obtain the *Fano*. The Matlab codes solving the above equations are available upon request.

Note that for the duty cycle ratio (*DCR*) [3], it is calculated as

$$
DCR = \sum_{\beta=0}^{1} \sum_{\alpha=0}^{\alpha^*-1} P^{SS}_{\alpha,\beta}, \tag{17}
$$

in which we count the intrinsic *on* status of our transcription system for all

configurations with $a < a^*$, and count the transcription *off* status for $a = a^*$ only (for single RNAP case) as the initiation is suspended at the supercoiling level $a^*$.

| Parameter | Symbol | value | Units | Reference |
|---|---|---|---|---|
| RNAP binding rate constant ($k_{pc} = k_{pc}^0$ [RNAP]) | $k_{pc}^0$ | 190 | $\mu M^{-1}s^{-1}$ | [4] |
| RNAP dissociaiton rate | $k_{cp}$ | 1 | $s^{-1}$ | [4] |
| Transition rate from close complex to open complex | $k_{co}$ | 30 | $s^{-1}$ | [4] |
| Transition rate from open complex to close complex | $k_{oc}$ | 150 | $s^{-1}$ | [4] |
| Length of T7 RNAP | $r$ | 10 | bp | [5] |
| Length of *E.coli* RNAP | $r$ | 35 | bp | [6] |
| Pausing density | $\gamma$ | 0.03 | $nt^{-1}$ | [7] |
| Duration time of pausing | $t$ | 2 | $s^{-1}$ | [7] |

**Table S1** Parameter values based on references or estimation.

## References


[1] Kim P and Lee C H 2014 Fast probability generating function method for stochastic chemical reaction networks *MATCH Commun. Math. Comput. Chem* **71** 57-69

[2] Thattai M and Van Oudenaarden A 2001 Intrinsic noise in gene regulatory networks *Proceedings of the National Academy of Sciences* **98** 8614-9

[3] Howard J 2001 *Mechanics of motor proteins and the cytoskeleton* page 221 Sinauer Associates, Inc

[4] Tang G Q, Roy R, Bandwar R P, Ha T and Patel S S 2009 Real-time observation of the transition from transcription initiation to elongation of the RNA polymerase *Proc Natl Acad Sci U S A* **106** 22175-80

[5] Zhou Y and Martin C T 2006 Observed Instability of T7 RNA Polymerase Elongation Complexes Can Be Dominated by Collision-induced "Bumping" *The journal of Biological Chemistry* **281** 24441-8

[6] Proshkin S, Rahmouni A R, Mironov A and Nudler E 2010 Cooperation between translating ribosomes and RNA polymerase in transcription elongation *Science* **328** 504-8

[7] Mejia Y X, Mao H, Forde N R and Bustamante C 2008 Thermal Probing of E. coli RNA Polymerase Off-Pathway Mechanisms *Journal of Molecular Biology* **382** 628-37