DNA Supercoiling Drives a Transition between Collective Modes of Gene Synthesis

Purba Chatterjee,[†] Nigel Goldenfeld,[‡] and Sangjin Kim[*]
Department of Physics and Center for the Physics of Living Cells, University of Illinois at Urbana-Champaign, Loomis Laboratory of Physics, 1110 West Green Street, Urbana, Illinois 61801, USA
and Carl R. Woese Institute for Genomic Biology, University of Illinois at Urbana-Champaign, 1206 West Gregory Drive, Urbana, Illinois 61801, USA

(Received 19 March 2021; accepted 18 October 2021; published 16 November 2021)

Transcription of genes can be affected by both biochemical and mechanical factors. Recent experiments suggested that the mechanical stress associated with transcription-induced DNA supercoiling is responsible for the transition from cooperative to antagonistic group dynamics of RNA polymerases (RNAPs) upon promoter repression. To underpin the mechanism behind this drastic transition, we developed a continuum deterministic model for transcription under torsion. In our model, the speed of an RNAP is affected by the local DNA supercoiling, as well as two global factors: (i) the number of RNAPs on the gene affecting the torsional stress experienced by individual RNAPs and (ii) transcription factors blocking the diffusion of DNA supercoils. Our minimal model can successfully reproduce the experimental findings and helps elucidate the interplay of mechanical and biological factors in the collective dynamics of molecular machines involved in gene expression.

DOI: 10.1103/PhysRevLett.127.218101

Transcription by the molecular machine, RNA polymerase (RNAP) is the first step of gene expression and is known to proceed through a collective mechanism: RNAPs concurrently transcribing a gene increase their efficiency through cooperative interactions [1–4]. Previous experimental and theoretical studies generally predict that the speed of an RNAP (elongation rate) increases with the density of RNAPs, or the rate at which RNAPs initiate at the promoter (RNAP flux) [1–9]. On the contrary, a recent study on the transcription of *lacZ* gene in the bacterium *Escherichia coli* showed that multiple RNAPs translocate at high speed irrespective of the initiation rate, as long as the RNAP flux is continuous (cooperative mode) [10]. Surprisingly, environmentally induced repression, that is an interruption of the RNAP flux, results in a drastic slowdown of co-transcribing RNAPs (antagonistic mode), and this slowdown is more pronounced for later repression (higher RNAP density on the gene).

The transition from cooperative to antagonistic dynamics is likely mediated by DNA supercoiling because it was observed in topologically constrained DNA (e.g., plasmids and chromosomes), but not in linear DNA, where the two ends can freely rotate to dissipate supercoils [10]. How does DNA supercoiling generate these two contrasting modes of RNAP group dynamics? The coupling between DNA supercoiling and RNAP translocation is known for a single RNAP. Forward translation of an RNAP results in underwinding of the DNA behind (negative supercoiling) and overtwisting of the DNA in front (positive supercoiling) [11]. Also, the accumulation of these supercoils slows down the RNAP due to torsional stress [12–14]. However, the effect of DNA supercoiling on multiple RNAPs concurrently transcribing the same DNA is less clear. The simple assumption that positive and negative DNA supercoils cancel commensurately between RNAPs [11,15,16] cannot explain the existence of both cooperative and antagonistic modes of RNAP dynamics.

The density and flux of RNAPs on a gene are important parameters for gene regulation. They vary widely from gene to gene [17–19] and change dynamically with a changing environment through the binding or unbinding of transcription factors (TFs) [20,21]. How DNA supercoiling regulates RNAP dynamics under a wide range of RNAP flux conditions found in physiological settings remains poorly understood.

The purpose of this Letter is to introduce a minimal deterministic model that is generally applicable for describing RNAP translocation dynamics coupled with transcription-induced DNA supercoiling under different RNAP flux conditions. Our model is based on two novel hypotheses regarding the mechanism of torsional-stress generation during transcription. The first hypothesis is that the stress due to DNA supercoiling is exacerbated by the number of RNAPs on the gene. The second hypothesis posits that TFs, which bind near the promoter, affect transcription initiation as well as the

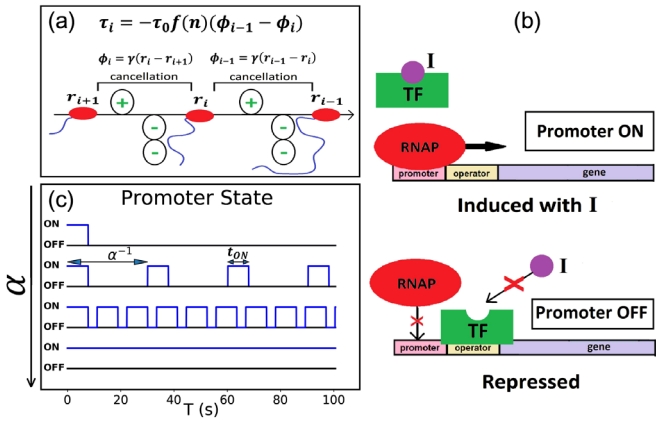

FIG. 1. The model. (a) Excess negative supercoiling in DNA segments and restoring torque on RNAP_i. (b) ON and OFF states of the promoter regulated by a repressor TF. (c) Time series of the promoter state for four different initiation rates α.

diffusion of DNA supercoils. Despite its simplicity, the minimal model accurately recapitulates the experimental observations of Ref. [10] and is a step towards a quantitative understanding of collective effects during gene expression.

Model.—We model the ith RNAP (RNAP_i) as a point particle translocating on a gene of length L with its position on the DNA given by r_i and speed by $v_i = \dot{r}_i$ [Fig. 1(a)]. As it translocates, the RNAP synthesizes an mRNA of increasing length and twists the DNA to generate positive supercoils in front and negative supercoils behind [11]. Positive and negative DNA supercoils are assumed to cancel commensurately between RNAPs [11,15,16]. Hence, the excess negative supercoils introduced into the segment of DNA upstream of RNAP_i, after cancellations with supercoils made by RNAP_{i+1}, is given by

$$\phi_i = \gamma(r_i - r_{i+1}), \tag{1}$$

where γ is the rate of supercoil injection per base pair (bp) transcribed by the RNAP. The theoretical maximum for γ is $1/l_0$ bp^{-1}, where $l_0 = 10.5$ is the number of base pairs in one helical turn of relaxed DNA. However, the actual rate of supercoil accumulation can be much lower (see Discussion), and we choose $\gamma = 0.01$.

The restoring torque applied by a segment of DNA is taken to be proportional to its excess supercoiling ϕ (local effect) [11,22,23]. Also, we hypothesize that this restoring torque depends on the number of RNAPs on the gene, n (global effect). This is because having many bulky RNAP molecules on the gene, along with increasing lengths of nascent mRNA synthesized per RNAP, would make it harder to twist the DNA. The net torque acting on RNAP_i is the difference between the restoring torques applied by its downstream and upstream DNA segments and is given by

$$\tau_i = -\tau_0 f(n)(\phi_{i-1} - \phi_i). \tag{2}$$

Here, τ_0 is a proportionality constant, and $f(n)$ is a monotonically increasing function of n. If the upstream DNA segment is more negatively supercoiled than the downstream DNA segment, the positive net torque provides resistance to translocation [13]. Because of $f(n)$, the number of RNAPs on the gene can exacerbate the torque effect. This model feature is supported by the observation that repressing the promoter earlier, when there are fewer RNAPs on a gene, yields less slowdown of RNAPs, likely due to smaller resisting torque [10].

The speed of RNAP_i decreases with increasing torque as

$$v_i(\tau_i) = \frac{2v_0}{1 + \exp[2(\tau_i/\tau_c)^3]}, \tag{3}$$

where v_0 is the typical RNAP speed, and τ_c is the stalling torque, above which RNAP can be halted. The speed depends on τ^3 to ensure that the torsional stress experienced by an RNAP is more pronounced for higher absolute values of the restoring torque, corresponding to higher levels of supercoiling. This also reflects in part the increasing drag on an RNAP as it synthesizes longer mRNA. While the speed drops to zero at high positive torques, negative torques assist transcription elongation [13]. Thus, we posit that an RNAP can transcribe at its maximum speed of 60 bp/s [1,24] under high negative torques.

In our model, RNAP flux is affected by a repressor TF binding at the promoter, which can sterically hinder initiation [Fig. 1(b)]. The promoter can be turned ON when an inducer (I) binds to TF and causes it to dissociate. The higher the concentration of I, the more frequently TF unbinds, and the higher is the RNAP flux or the transcription initiation rate α. If the inducer disappears, TF rebinds on the DNA, and prevents further loading of RNAPs, interrupting RNAP flux (promoter repression).

Following previous experimental observations [25,26], we hypothesize that TF binding, besides turning the promoter OFF periodically [Fig. 1(c)], can also physically block the diffusion of supercoils (see Discussion). When the promoter is turned ON by TF dissociation, the negative supercoils behind the last loaded RNAP would diffuse out, removing torsional stress on this RNAP. We assume that this diffusion takes place before the next RNAP loading event because the diffusion of DNA supercoils is about 100 times faster than RNAP initiation and elongation dynamics [27]. The promoter remains ON for a duration t_{ON}, which is the average time taken by TF to rebind. As long as the promoter remains ON, no DNA supercoils accumulate behind the last loaded RNAP. When the promoter is turned OFF, the TF blocks both the dissipation of DNA supercoils and the loading of RNAPs till the next time the promoter is turned ON. Repression at time T_{stop} turns the promoter OFF completely thereafter and prevents further RNAP loading. Lastly, we allow positive supercoils in front of RNAP_1 to diffuse downstream unhindered; however, the relaxation of

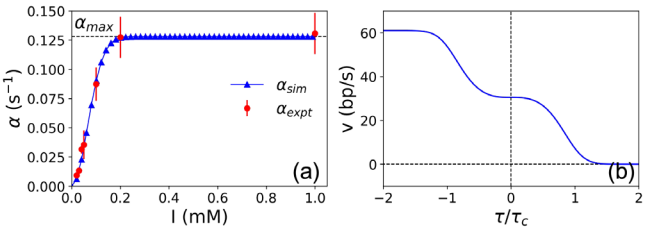

FIG. 2. Model parameters. (a) Initiation rates α used in the simulations (blue triangles) based on experimental data (red solid circles) in Ref. [10]. (b) Speed of an RNAP v as a function of torque τ/τ_c.

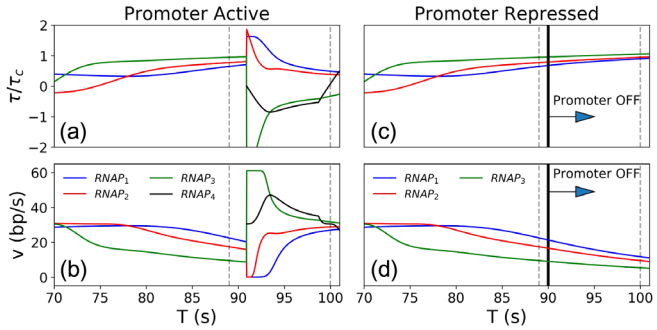

FIG. 3. Time series of τ/τ_c and v of the first 3–4 RNAPs for the case of active promoter (a)–(b) and promoter repression at $T = 90$ s (c)–(d) at the intermediate initiation rate $\alpha = 0.033$ s^{-1}. We only plot $\tau/\tau_c \in (-2, 2)$ for clarity. Gray dashed lines demarcate the time range for Supplemental Material, Fig. S2 [29].

this assumption does not change the main results of the model.

Model parameters and methods.—As a proof of concept, we apply our general model to the *lac* operon in *E. coli*, a paradigm of bacterial gene regulation, for which experimental results are available from Ref. [10]. We focus on the transcription of *lacZ*, the first gene in the *lac* operon, with length $L = 3072$ bp. LacI repressor is the TF.

The effective initiation rates $\alpha_{\rm sim}$ used in the simulations were obtained from a fit to the observed $\alpha_{\rm expt}$ as a function of the concentration of inducer (I) used in Ref. [10] [Fig. 2(a)]. $t_{\rm ON}$ is taken to be the inverse of $\alpha_{\rm max}$, the highest RNAP flux observed experimentally. We approximate the dependence of the restoring torque on RNAP density by a cubic polynomial of the form $f(n) = 1 + a(n-1) + b(n-1)^2 + c(n-3)^3$ (see Discussion), with $(a, b, c) = (0.78, 3.32, 0.38)$. Figure 2(b) shows the dependence of RNAP speed v on the restoring torque τ, where $v_0 = 30.5$ bp/s is the typical RNAP speed [10], $\tau_c = 11$ pN · nm is the stalling torque measured for *E. coli* RNAP [13], and $\tau_0 = 0.386$ pN · nm. A different value of stall torque [28] would only require a different choice of τ_0, without changing our main results.

To calculate the average elongation rate of the first RNAP ($RNAP_1$), we follow the prescription in Ref. [10]. When the promoter is active, the average elongation rate is $v_{\rm ON} = L/T_1$, where T_1 is the time taken by $RNAP_1$ to complete transcription. When the promoter is repressed at time $T_{\rm stop}$, the average elongation rate of $RNAP_1$ for the remaining portion of the gene after repression is $v_{\rm OFF} = (L - v_{\rm ON}T_{\rm stop})/(T_1 - T_{\rm stop})$. This definition of $v_{\rm OFF}$ assumes that the RNAPs move at a constant speed $v_{\rm ON}$ till the promoter is repressed at $T = T_{\rm stop}$. The assumption is not always valid (see Supplemental Material, Fig. S1 [29]), but we adhere to this definition of $v_{\rm OFF}$ for comparison with the experimental results of Ref. [10].

Simulation results.—Figures 3(a) and 3(b) show the time series of τ/τ_c and v for the first 3–4 RNAPs at the intermediate initiation rate $\alpha = 0.033$ s^{-1}, in which the promoter cycles between ON and OFF states [Fig. 1(c) second from top]. At $T = 70$ s, there are three RNAPs on the gene moving at v_0 bp/s. However, because the promoter is in the OFF state, DNA supercoil diffusion is blocked at the promoter by TF, leading to a sequential slowing down of RNAPs starting with $RNAP_3$, until the next TF dissociation event (at $T \approx 91$ s). At this time, the negative supercoils behind $RNAP_3$ dissipate, and all RNAPs on the gene can quickly equilibrate to the optimal speed v_0 by $T = 100$ s. This equilibration proceeds through the acceleration and deceleration of RNAPs reacting to the ambient torsional stress (detailed in the Supplemental Material, Fig. S2 [29]).

If the promoter is repressed at $T = 90$ s [Figs. 3(c) and 3(d)], TF remains bound ($RNAP_4$ does not load), and the speeds of $RNAP_1$, $RNAP_2$, and $RNAP_3$ continue to decrease. The speed of $RNAP_1$ is reduced to 11.68 bp/s at $T = 100$ s, and it decreases even further thereafter. Therefore, promoter repression in the model recapitulates RNAP slowdown observed in the experiments (For the dynamics at low and high initiation rates, see Supplemental Material, Figs. S1 and S3 [29]).

Figure 4 shows the average elongation rate $v_{\rm ON}$ vs initiation rate α. For low initiation rates ($\alpha_{\rm sim} = 0.006$ s^{-1}), there is only a single RNAP on the gene on average, and $v_{\rm ON}$ is less than the typical speed $v_0 = 30.5$ bp/s. However, for a large range of higher α values, $v_{\rm ON} \approx v_0$, independent of the initiation rate. The inset of Fig. 4 shows the average elongation rate $v_{\rm OFF}$ upon promoter repression for three different initiation rates tested in the experiments. At a low initiation rate ($\alpha_{\rm sim} = 0.006$ s^{-1}), only a single RNAP transcribe a gene at a time, and promoter repression at $T = 90$ s does not appreciably affect the RNAP speed. For intermediate ($\alpha_{\rm sim} = 0.033$ s^{-1}) and high ($\alpha_{\rm sim} = 0.127$ s^{-1}) initiation rates, promoter repression at $T = 90$ s causes speeds to drop to about a quarter of a single RNAP speed. Notably, the effect is smaller if the promoter is repressed earlier (e.g., at $T = 45$ vs $T = 90$ s for the intermediate initiation rate), consistent with the experimental observation [10].

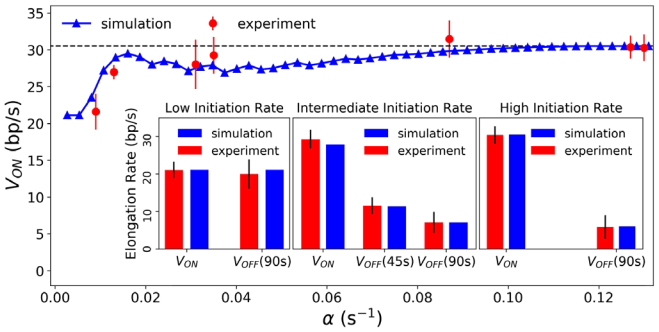

FIG. 4. Comparison of our model (blue) with the experimental results of Ref. [10] (red) on the speed of RNAPs v_{ON} as a function of initiation rates α. The inset shows the effect of promoter repression. Low initiation rate represents $\alpha_{\mathrm{sim}} = 0.006\ \mathrm{s}^{-1}$ ($\alpha_{\mathrm{expt}} = 0.009\ \mathrm{s}^{-1}$), yielding roughly a single RNAP on the gene at a given time. Intermediate and high initiation rates are from $\alpha_{\mathrm{sim}} = 0.033\ \mathrm{s}^{-1}$ ($\alpha_{\mathrm{expt}} = 0.035\ \mathrm{s}^{-1}$) and $\alpha_{\mathrm{sim}} = 0.127\ \mathrm{s}^{-1}$ ($\alpha_{\mathrm{expt}} = 0.127\ \mathrm{s}^{-1}$), respectively.

Discussion.—Our model generates three distinct modes of transcription based on RNAP flux and mechanical stress. For a continuous flux of RNAPs (active promoter), a collective cooperative mode is observed, where co-transcribing RNAPs can efficiently cancel each other's supercoils and move at optimal speeds, making many transcripts at a given time. When this flux is interrupted (promoter repression), we observe a collective antagonistic mode, characterized by a drastic reduction in the speed of co-transcribing RNAPs due to supercoil accumulation. The reduction is more pronounced for later repression, or if the repression occurs when the RNAPs are closer to the transcription termination. The slowdown may yield a sudden brake in transcription, especially when the slowdown results in pre-mature dissociation of RNAPs [10]. Lastly, for very low fluxes (a single RNAP), we observe a torsionally stressed mode of transcription elongation with lower than optimal speeds, but it is not altered by promoter repression. This likely corresponds to a basal level of transcription that is not actively regulated.

The total inertial resistance to DNA twisting is a product of the total mass of RNAPs on the gene as well as that of the increasing lengths of nascent mRNAs. While the former increases linearly with n, the total length of mRNA synthesized increases as $R \propto n(n-1)/2 \approx n^2$. Therefore, we approximate the dependence of the restoring torque on RNAP density, $f(n)$, with a cubic polynomial. However, the exact functional form for $f(n)$ and the coefficients are likely tied to the exact experimental conditions under consideration since they determine the twist modulus of DNA. Additionally, we choose a constant supercoil injection rate $\gamma = 0.01$, of the order of the approximate supercoiling density (number of supercoils per unit length of DNA transcribed) reported in the literature [11,22,23,30] because the actual supercoil injection rate *in vivo* remains unknown. It can be affected by various factors (e.g., frictional drag, topoisomerase activity, downstream topological barriers, supercoil diffusion, and DNA stretching forces), and it may change with the position of RNAP on the gene [11,22,23,30].

Our hypothesis that TFs can regulate the diffusion of DNA supercoils is supported by the observations that LacI functions as a topological barrier to constrain DNA supercoils [25,26]. This hypothesis allows for coupling the RNAP flux with the dissipation of DNA supercoils. We note that other molecules, such as RNAPs poised at the promoter, may have a similar effect as TFs [10,31,32]. Another novel implication of this hypothesis is for modeling transcription dynamics in the genomic context, where DNA supercoiling produced from neighboring genes should be considered [33,34]. For example, the diffusion of DNA supercoils or its lack thereof likely has important consequences for divergently transcribed genes commonly found in the genome [35]. It was shown in Ref. [10] that the divergent expression of another gene, positioned upstream of *lacZ*, reduces the transcription elongation rate of *lacZ* in the case of the high initiation rate $\alpha = 0.127\ \mathrm{s}^{-1}$. Moreover, this antagonistic effect was observed even when the two promoters are separated by as much as 2400 bp. This is entirely consistent with our model's prediction. At high initiation rates, the promoter is almost always ON [Fig. 1(c) highest α]. As a result, negative DNA supercoils generated by transcription of a neighboring gene can diffuse in and reduce the speed of RNAPs transcribing *lacZ*.

Our model is different from existing theories of transcription based on DNA supercoiling [5,8,23,36,37] because it considers not only transcription-induced DNA supercoils (local effect) but also the role of TFs as a barrier for the DNA supercoil diffusion and the dependence of the DNA restoring torque on the number of RNAPs on the gene (global effects). We find that these model features are critical in producing the two contrasting RNAP group dynamics between the active and repressed states of the promoter. Without them, it is not possible to consistently explain the experimental findings *even qualitatively* within our framework (details in Supplemental Material, Figs. S4–S7 [29]). For example, Fig. 5 demonstrates the consequences of relaxing both of the central assumptions of our theory. Here, we consider two scenarios in which the restoring torque does not depend on RNAP density [$f(n) = 1$]. In one case, TF does not constrain DNA supercoils when bound (similar to linear DNA whose ends are free to rotate and dissipate torsional stress). In this case, high speeds are observed for all initiation rates, even for a single RNAP, and moreover, the switch to antagonistic dynamics upon promoter repression cannot be recovered. In the other case, DNA supercoils are always constrained independent of TF binding (similar to DNA with clamped ends). Here, the collective cooperative mode is recovered,

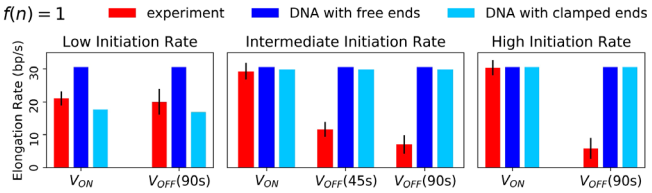

FIG. 5. Consequences of relaxing the central assumptions of our model. Red and blue bars represent experimental data and simulation results from $f(n) = 1$, respectively. (Dark blue) Free DNA: TF does not constrain supercoils even when bound near the promoter. (Light blue) Clamped DNA: supercoils are always constrained near the promoter, even when TF is unbound.

with multiple co-transcribing RNAPs translocating faster than a single RNAP when the promoter is active. However, their mutual antagonism upon promoter repression cannot be captured. We have also investigated the effect of imposed bursty initiation and found that contrary to existing literature [38], our model predicts that DNA supercoiling hinders the formation of convoys of RNAPs traveling at the same speed (see Supplemental Material, Fig. S8 [29]), even with bursty loading.

This work was supported by the National Science Foundation Center for Physics of Living Cells (Grant No. NSF PHY-1430124). P. C. acknowledges the Drickamer Research Fellowship, 2020. S. K. acknowledges support from the Searle Scholars Program and NIH (Grant No. R35GM143203).

Note added.—Recently, we received a preprint from Tripathi *et al*. [30] which overlaps with some of the results we report.

*Corresponding author.
sangjin@illinois.edu
†Present address: Department of Physics, University of Pennsylvania, Philadelphia, PA 19104, USA.
‡Present address: Department of Physics, University of California, San Diego, La Jolla, CA 92093, USA.

[1] V. Epshtein and E. Nudler, Science **300**, 801 (2003).
[2] H. Saeki and J. Q. Svejstrup, Mol. Cell **35**, 191 (2009).
[3] J. Jin, L. Bai, D. S. Johnson, R. M. Fulbright, M. L. Kireeva, M. Kashlev, and M. D. Wang, Nat. Struct. Mol. Biol. **17**, 745 (2010).
[4] O. I. Kulaeva, F.-K. Hsieh, and V. M. Studitsky, Proc. Natl. Acad. Sci. U.S.A. **107**, 11325 (2010).
[5] T. Heberling, L. Davis, J. Gedeon, C. Morgan, and T. Gedeon, PLoS Comput. Biol. **12**, e1005069 (2016).
[6] E. A. Galburt, J. M. Parrondo, and S. W. Grill, Biophys. Chem. **157**, 43 (2011).
[7] P. R. Costa, M. L. Acencio, and N. Lemke, PLoS One **8**, 1 (2013).
[8] A. Lesne, J.-M. Victor, E. Bertrand, E. Basyuk, and M. Barbi, in *Molecular Motors: Methods and Protocols*, edited by C. Lavelle (Springer, New York, NY, 2018), pp. 215–232.
[9] V. Belitsky and G. Schütz, J. Theor. Biol. **462**, 370 (2019).
[10] S. Kim, B. Beltran, I. Irnov, and C. Jacobs-Wagner, Cell **179**, 106 (2019).
[11] L. F. Liu and J. C. Wang, Proc. Natl. Acad. Sci. U.S.A. **84**, 7024 (1987).
[12] N. Rovinskiy, A. A. Agbleke, O. Chesnokova, Z. Pang, and N. P. Higgins, PLoS Genet. **8**, e1002845 (2012).
[13] J. Ma, L. Bai, and M. D. Wang, Science **340**, 1580 (2013).
[14] S. Chong, C. Chen, H. Ge, and X. S. Xie, Cell **158**, 314 (2014).
[15] P. Guptasarma, BioEssays **18**, 325 (1996).
[16] D. A. Koster, A. Crut, S. Shuman, M.-A. Bjornsti, and N. H. Dekker, Cell **142**, 519 (2010).
[17] M. H. Larson, R. A. Mooney, J. M. Peters, T. Windgassen, D. Nayak, C. A. Gross, S. M. Block, W. J. Greenleaf, R. Landick, and J. S. Weissman, Science **344**, 1042 (2014).
[18] A. Mayer, J. Di Iulio, S. Maleri, U. Eser, J. Vierstra, A. Reynolds, R. Sandstrom, J. A. Stamatoyannopoulos, and L. S. Churchman, Cell **161**, 541 (2015).
[19] I. M. Min, J. J. Waterfall, L. J. Core, R. J. Munroe, J. Schimenti, and J. T. Lis, Genes Dev. **25**, 742 (2011).
[20] K. Struhl, Cell **98**, 1 (1999).
[21] F. Spitz and E. E. Furlong, Nat. Rev. Genet. **13**, 613 (2012).
[22] J. F. Marko, Phys. Rev. E **76**, 021926 (2007).
[23] S. A. Sevier and H. Levine, Phys. Rev. Lett. **118**, 268101 (2017).
[24] H. Chen, K. Shiroguchi, H. Ge, and X. S. Xie, Mol. Syst. Biol. **11**, 781 (2015).
[25] F. Leng and R. McMacken, Proc. Natl. Acad. Sci. U.S.A. **99**, 9139 (2002).
[26] G. Fulcrand, S. Dages, X. Zhi, P. Chapagain, B. S. Gerstman, D. Dunlap, and F. Leng, Sci. Rep. **6**, 19243 (2016).
[27] M. T. J. van Loenhout, M. V. de Grunt, and C. Dekker, Science **338**, 94 (2012).
[28] J. Ma, C. Tan, X. Gao, R. M. Fulbright, J. W. Roberts, and M. D. Wang, Proc. Natl. Acad. Sci. U.S.A. **116**, 2583 (2019).
[29] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.127.218101 for details of RNAP dynamics at different initiation rates, and the consequences of relaxing the central assumptions of this theory.
[30] S. Tripathi, S. Brahmachari, J. N. Onuchic, and H. Levine, bioRxiv, https://doi.org/10.1101/2021.03.04.433986 (2021).
[31] S. Deng, R. A. Stein, and N. P. Higgins, Proc. Natl. Acad. Sci. U.S.A. **101**, 3398 (2004).
[32] D. H. Price, Mol. Cell **30**, 7 (2008).
[33] S. Meyer and G. Beslon, PLoS Comput. Biol. **10**, e1003785 (2014).
[34] P. Sobetzko, Nucl. Acids Res. **44**, 1514 (2016).
[35] W. Wei, V. Pelechano, A. I. Jrvelin, and L. M. Steinmetz, Trends Genet. **27**, 267 (2011).
[36] C. A. Brackley, J. Johnson, A. Bentivoglio, S. Corless, N. Gilbert, G. Gonnella, and D. Marenduzzo, Phys. Rev. Lett. **117**, 018101 (2016).
[37] A. Klindziuk and A. B. Kolomeisky, J. Phys. Chem. B **125**, 4692 (2021).
[38] K. Tantale, F. Mueller, A. Kozulic-Pirher, A. Lesne, J.-M. Victor, M.-C. Robert, S. Capozi, R. Chouaib, V. Bcker, J. Mateos-Langerak, X. Darzacq, C. Zimmer, E. Basyuk, and E. Bertrand, Nat. Commun. **7**, 12248 (2016).

Supplementary Material for:
DNA Supercoiling Drives a Transition between Collective Modes of Gene Synthesis

Purba Chatterjee, Nigel Goldenfeld, and Sangjin Kim
Department of Physics and Center for the Physics of Living Cells, University of Illinois at Urbana-Champaign, Loomis Laboratory of Physics, 1110 West Green Street, Urbana, Illinois 61801, USA
Carl R. Woese Institute for Genomic Biology, University of Illinois at Urbana-Champaign, 1206 West Gregory Drive, Urbana, Illinois 61801, USA

In this supplementary material, we provide the details of RNAP dynamics for different initiation rates and promoter activities, i.e., active and repressed promoter (Sec. S1). Furthermore, we try to relax the central hypotheses of our model and compare the results with experimental data (Sec. S2 and Sec. S3). We also investigate the effect of imposed bursty initiation in our model (Sec. S4). Throughout the supplementary materials, we will refer to negative supercoils as NS.

S1. COMPARISON OF DYNAMICS FOR ACTIVE AND REPRESSED PROMOTERS

Here, we compare the torsional stress and speeds of the first few RNAPs for different promoter strength (low, intermediate, and high initiation rates) and activities (active and repressed).

A. Low Initiation Rate

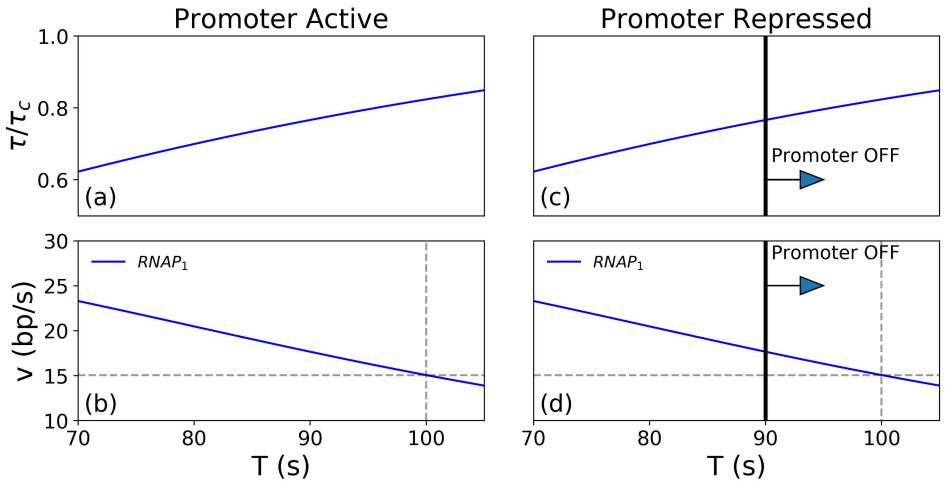

FIG. S1. (color online) Time series of τ/τ_c and v of an RNAP at the low initiation rate $\alpha = 0.006$ s^{-1} when the promoter is active (a-b) and when the promoter is repressed at $T = 90$ s (c-d). The speed of the single RNAP continues to decrease even when the promoter is active (b), and the same RNAP dynamics is seen even when the promoter is repressed $T = 90$ s (d).

At the low initiation rate ($\alpha = 0.006$ s^{-1}), there is only a single RNAP on the gene on average. Fig. S1 shows the time series of restoring torque τ/τ_c and speed v of the single RNAP for the active promoter (Fig. S1(a,b)) and for the promoter repressed at $T = 90$ s (Fig. S1(c,d)). The speed decreases continuously, such that at $T = 100$ s we have $v = 15.05$ bp/s for both active *and* repressed promoters. The time taken by the single RNAP to complete transcription, T_1, is the same in both cases. Due to low α, there is no upstream RNAP that can assist through DNA supercoil cancellation. Thus, the dynamics of a single RNAP is subject to similar levels of torsional stress irrespective of the promoter state.

Because the RNAP speed decreases continuously at low initiation rates, the definition of the average elongation rate after repression in [13], $v_{OFF} = (L - v_{ON}T_{stop})/(T_1 - T_{stop})$, underestimates the position of the single RNAP

upon repression (i.e., when $T = T_{stop}$). As a result, it predicts a higher v_{OFF} than if we were to consider the actual position $r_1(T_{stop})$. This is a limitation of the experiment, which does not track the position of the RNAPs. One has to assume a constant speed v_{ON} till T_{stop} in order to calculate the speed after repression, v_{OFF}.

Whenever the time taken to complete transcription with the active promoter (T^{ON}_{end}) is equal to that with the promoter repressed at T_{stop} (T^{OFF}_{end}), this definition of v_{OFF} always predicts $v_{OFF} = v_{ON}$. That is, with $L = v_{ON}T^{ON}_{end}$, we have

$$\begin{aligned} v_{OFF} &= \frac{L - v_{ON}T_{stop}}{T^{OFF}_{end} - T_{stop}}, \\ &= \frac{v_{ON}(T^{ON}_{end} - T_{stop})}{T^{OFF}_{end} - T_{stop}}, \\ \implies v_{OFF} &= v_{ON}, \end{aligned} \tag{S1}$$

for $T^{OFF}_{end} = T^{ON}_{end}$. The actual elongation rate after repression would be lower than that calculated by this prescription, on average. However, we chose to adhere to this definition of v_{OFF} for accurate comparison with the experimental results of [13]. What is important to note is that the dynamics of the RNAP remains unaffected by the promoter state and that the time taken to complete transcription is the same for both active and repressed promoters (regardless of T_{stop}) when there is a single RNAP on the DNA. Thus, one should read the result $v_{ON} = v_{OFF}$ for the single RNAP case as $T^{OFF}_{end} = T^{ON}_{end}$, that is the time of transcription completion for a single RNAP is unaffected by active or repressed conditions of the promoter.

B. Intermediate Initiation Rate

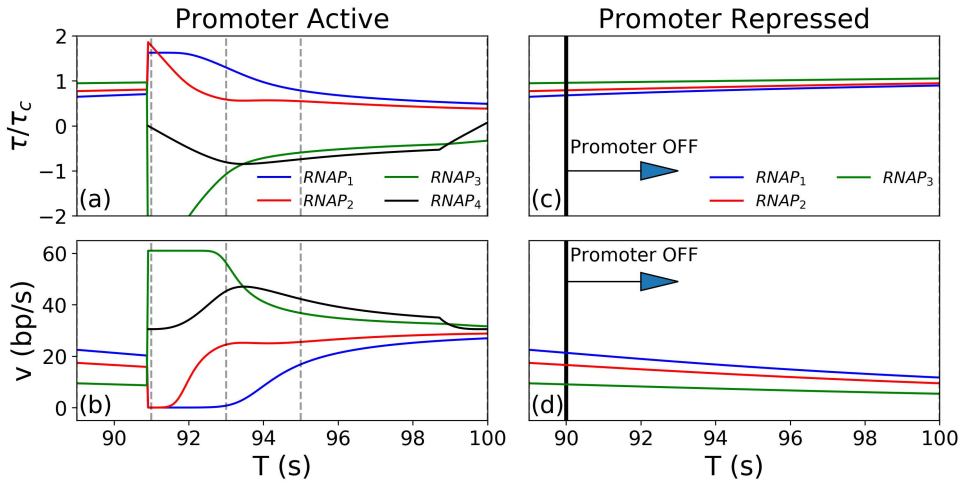

FIG. S2. (color online) Time series of τ/τ_c and v of the first 3 - 4 RNAPs from a promoter with the intermediate initiation rate $\alpha = 0.033$ s^{-1}. (a-b) is when the promoter is continuously active, and (c-d) is when the promoter is repressed at $T = 90$ s. This is a zoom-in version of Fig. 3 of the main text. The dashed gray lines mark the time points discussed below. When the promoter is active, the first three RNAPs start at the typical speed v_0, but their speeds reduce due to NS accumulation. However, $RNAP_4$ loads at $T \approx 91$ s, and the resulting dynamics allows the speed of all four RNAPs to equilibrate to the typical speed v_0 by $T = 100$ s. In contrast, with promoter repression at $T = 90$ s, speeds of the three loaded RNAPs continue decreasing beyond $T = 91$ s.

Fig. S2 shows the dynamics of RNAPs for the intermediate initiation rate $\alpha = 0.033$ s^{-1}. We plotted the time series of restoring torque τ/τ_c and speed v for the first 3 - 4 RNAPs within the time range demarcated by gray dashed lines in Fig. 3 of the main text. At $T = 89$ s, there are three RNAPs on the gene moving at speeds less than v_0, with $RNAP_3$ the slowest and $RNAP_1$ the fastest. $RNAP_3$ is slow because NS accumulates behind it while TF remains bound. There is a sequential slowing down of all downstream RNAPs starting from the promoter region due to insufficient cancellation of their NS by their slow upstream neighbor RNAPs.

When the promoter stays active, TF dissociates for the next RNAP loading. In Fig. S2(a,b), TF dissociates at $T \approx 91$ s for the loading of $RNAP_4$, and there are a few consequences. In our model, when TF dissociates, the NS

behind $RNAP_3$ diffuses out first prior to $RNAP_4$ loading, and hence, the remaining NS *in front* of the $RNAP_3$ causes its speed to increase (green). As $RNAP_4$ loads, the speeds of $RNAP_1$ (blue) and $RNAP_2$ (red) fall, owing to the increased difficulty of translocating by overtwisting the DNA with an additional RNAP on the gene. The fast-moving $RNAP_3$ can cancel supercoils ahead more efficiently, so it speeds up $RNAP_2$ (see $T \approx 92$ s). $RNAP_4$ (black) initially speeds up right after loading because it has NS ahead, owing to the high speed of $RNAP_3$. Thus, at $T = 93$ s, we see both $RNAP_2$ and $RNAP_4$ accelerating. Soon after, at $T = 95$ s, $RNAP_3$ and $RNAP_4$ start slowing down due to NS accumulation behind them, and at the same time, $RNAP_1$ and $RNAP_2$ are speeding up due to better cancellation of their NS. Eventually, at around $T = 100$ s, all four RNAPs have once again settled to the typical speed v_0.

TF rebinds at some point after $RNAP_4$ loading and blocks supercoil diffusion. As more NS accumulate behind $RNAP_4$, it starts to slow down, re-initiating a sequential decrease in the speeds of downstream RNAPs. However, as long as loading is uninterrupted (active promoter), RNAPs can always equilibrate to the optimal speed. In contrast, when the promoter is repressed at $T = 90$ s (Fig. S2(c,d)), $RNAP_4$ does not load, and the speeds of $RNAP_1$, $RNAP_2$, and $RNAP_3$ continue to decrease. At $T = 100$ s, the speed of $RNAP_1$ reduces to 11.68 bp/s, as compared to 26.95 bp/s when the promoter remains active. The RNAP slows down even further after $T = 100$ s to finally record an average elongation rate $v_{OFF} = 7.02$ bp/s, as shown in Fig. 4 of the main text.

C. High Initiation Rate

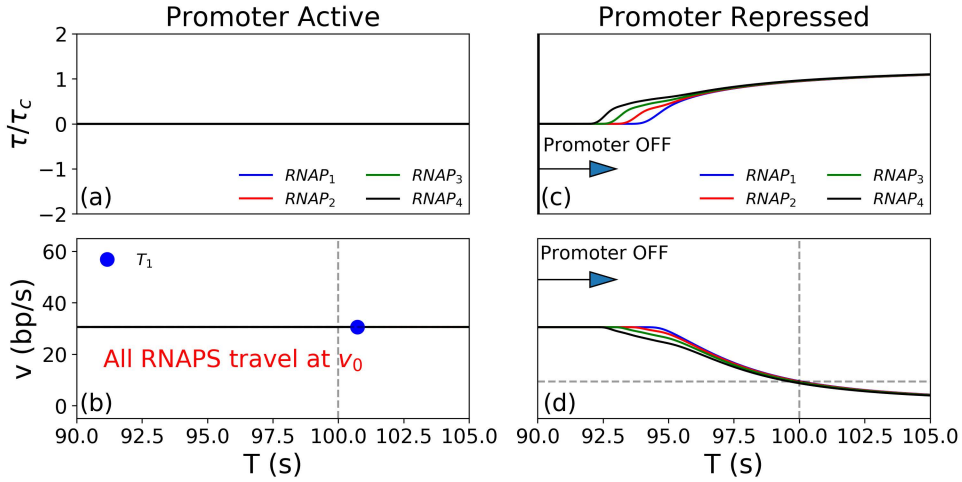

FIG. S3. (color online) Time series of τ/τ_c and v of the first four RNAPs in the case of uninterrupted loading (a-b) and promoter repression at $T = 90$ s (c-d) for the high initiation rate $\alpha = 0.127$ s^{-1}. When the promoter remains active, all RNAPs travel at the typical speed v_0. In contrast, after promoter repression at $T = 90$ s, speeds of the first four RNAPs suddenly reduce drastically over a short period of time as the torsional stress crosses threshold values.

At a high initiation rate ($\alpha = 0.127$ s^{-1}) of the active promoter, TF dissociates frequently, and the promoter is almost always ON (see Fig. 1 for the highest α). As such, NS do not accumulate behind the last loaded RNAP as long as the promoter is active. Fig. S3 shows the time series of τ/τ_c and v of the first four RNAPs for the active promoter (Fig. S3(a,b)) and for the promoter repressed at $T = 90$ s (Fig. S3(c,d)). It is clear that for the active promoter, there is negligible torsional stress throughout, and all RNAPs travel at the typical speed v_0. In Fig. S3(b), we have marked $T_1 = 100.73$ s, the time taken for the first RNAP to complete transcription. In contrast, when the promoter is repressed at $T = 90$ s, we see a drastic reduction of RNAP speeds over a very short period of time. For example, the speed of the first RNAP reduces to $v = 9.48$ bp/s at $T = 100$ s. This reduction is caused by TF binding at T_{stop}, which prevents both further initiation as well as the diffusion of NS produced by the last loaded RNAP. Moreover, because there are approximately $n = 12$ RNAPs on the gene when the promoter is repressed, the accumulation of a very small amount of supercoiling is sufficient to increase the torsional stress beyond threshold values. Thus, promoter repression in the model recapitulates the drastic reduction in RNAP speeds observed in the experiments [13]. A greater reduction is expected for larger RNAP densities on the gene, i.e for higher initiation rates, suggesting that this antagonistic effect is another group effect of RNAPs.

S2. TORSIONAL STRESS INDEPENDENT OF RNAP DENSITY

In our model, we hypothesize that the presence of many RNAPs on the gene exacerbates the torsional stress by making the DNA more difficult to overtwist. The dependence of the restoring torque τ on the RNAP density n is encoded by the function $f(n)$. To relax this assumption in our model, we consider the situation where τ is independent of n, i.e. $f(n) = 1$. In the following subsections, we examine $f(n) = 1$ under three different scenarios related to supercoil diffusion at the promoter.

A. TF Blocks NS Diffusion in its Bound State

Fig. S4 shows elongation rates of various conditions assuming that DNA-bound TF blocks NS diffusion (as in our main model) but $f(n) = 1$. In the active state of the promoter, we once again see high elongation rates independent of initiation rates for intermediate to high α, whereas the single RNAP case at low initiation rates (e.g., $\alpha = 0.006$ s^{-1}) has a lower speed. However, for all initiation rates, promoter repression at neither $T = 45$ s nor $T = 90$ s shows any change in elongation rates v_{OFF} from their v_{ON} values (Fig. S4 (inset)). This is in contrast to the experimental observation of [13] that promoter repression causes a large reduction in RNAP speeds. Thus, without the dependence of torsional stress on RNAP density, even with TF blocking NS diffusion when bound, we cannot reproduce the observed switch from cooperative to antagonistic collective dynamics of RNAPs upon promoter repression.

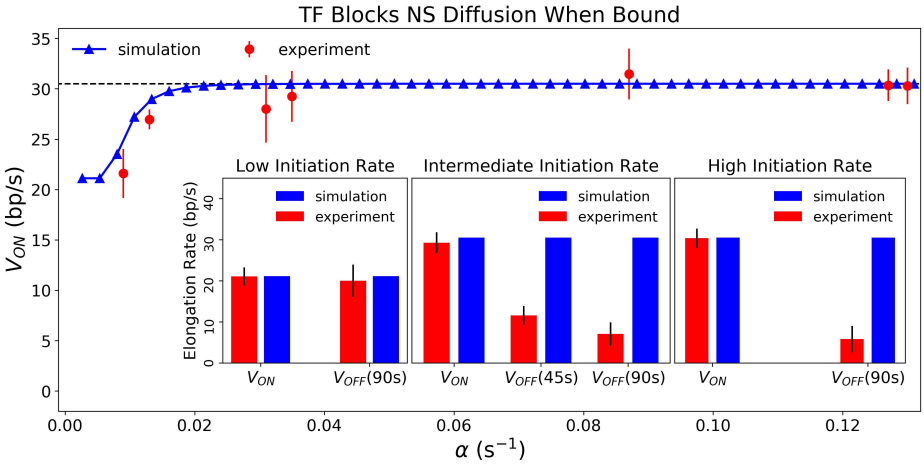

FIG. S4. (color online) Scenario with $f(n) = 1$ and TF blocking NS diffusion in its bound state. v_{ON} is low for a single RNAP ($\alpha_{sim} = 0.006$ s^{-1}), but it remains high independent of initiation rates for a large range of α. The inset shows the effect of promoter repression. Promoter repression at $T = 45$ s or $T = 90$ s does not appreciable change v_{OFF} from v_{ON} for low ($\alpha_{sim} = 0.006$ s^{-1}, $\alpha_{expt} = 0.009$ s^{-1}), intermediate ($\alpha_{sim} = 0.033$ s^{-1}, $\alpha_{expt} = 0.035$ s^{-1}), and high ($\alpha_{sim} = 0.127$ s^{-1}, $\alpha_{expt} = 0.127$ s^{-1}) initiation rates.

B. DNA with Free Ends: TF Never Blocks NS Diffusion

Fig. S5 shows elongation rates for the case where $f(n) = 1$ and TF never blocks NS diffusion. In other words, NS always diffuse out. This situation could arise in the case of linear DNA that always allows supercoil dissipation through its free ends or in the case where TF is a comparatively smaller molecule and cannot constrain supercoils. In the active promoter, we see high elongation rates independent of initiation rates for *all* α. Even a single RNAP ($\alpha = 0.006$ s^{-1}) transcribes at the optimal speed. This is contradictory to the experimental observation that co-transcribing RNAPs can collectively increase their elongation rates in comparison to a single RNAP. Additionally, like in Sec. S2 A, promoter repression at $T = 45$ s or at $T = 90$ s does not show any change in elongation rates for any initiation rate (inset in Fig. S5). Thus, with the torsional stress independent of RNAP density ($f(n) = 1$) and with TF unable to block NS diffusion even when bound, we cannot reproduce either the collective or the antagonistic dynamics of RNAPs observed in [13].

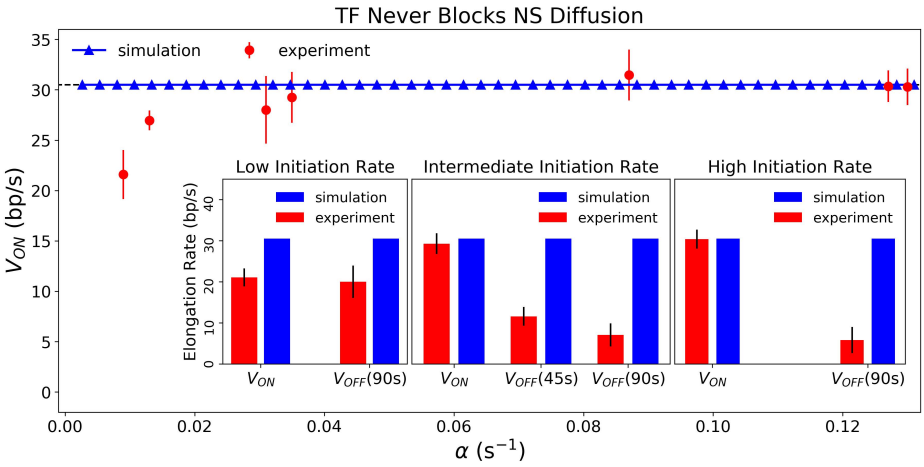

FIG. S5. (color online) Scenario with $f(n) = 1$ and free DNA with TF never blocking NS diffusion. v_{ON} is high and independent of initiation rates for all α, even for a single RNAP ($\alpha_{sim} = 0.006$ s^{-1}). The inset shows the effect of promoter repression. Promoter repression at $T = 45$ s or $T = 90$ s does not appreciable change v_{OFF} from v_{ON} for low ($\alpha_{sim} = 0.006$ s^{-1}, $\alpha_{expt} = 0.009$ s^{-1}), intermediate ($\alpha_{sim} = 0.033$ s^{-1}, $\alpha_{expt} = 0.035$ s^{-1}), and high ($\alpha_{sim} = 0.127$ s^{-1}, $\alpha_{expt} = 0.127$ s^{-1}) initiation rates.

C. DNA with Clamped Ends: No NS Diffusion

Fig. S6 shows elongation rates for $f(n) = 1$ and for DNA clamped at its ends, with no NS diffusion. This case explores the scenario where TF binding or unbinding only affects RNAP loading but is irrelevant to the torsional stress. This can also be considered as the general case with a bulky molecule always bound to the DNA upstream of the promoter, which does not affect RNAP loading but blocks NS diffusion. Fig. S6 shows that elongation rates

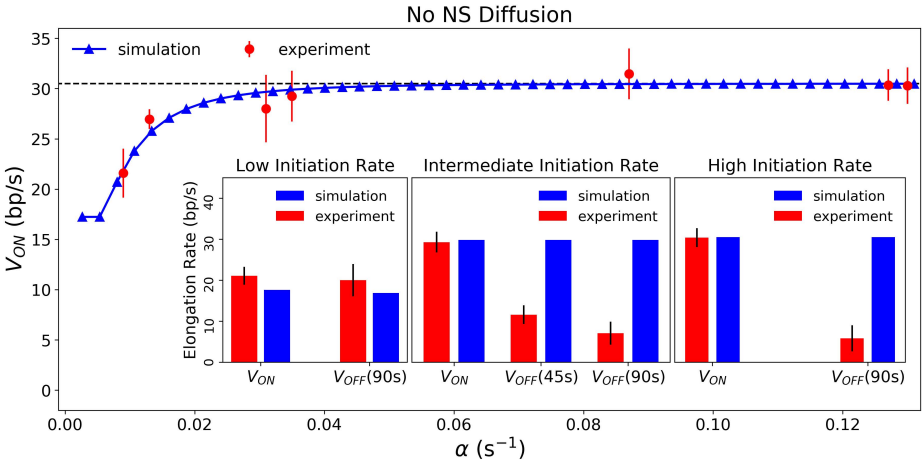

FIG. S6. (color online) Scenario with $f(n) = 1$ and clamped DNA with no NS diffusion. The main panel shows that v_{ON} is low for low initiation rates and high and independent of the initiation rates for intermediate and high α. The inset shows the effect of promoter repression. Promoter repression at $T = 45$ s or $T = 90$ s does not appreciable change v_{OFF} from v_{ON} for intermediate ($\alpha_{sim} = 0.033$ s^{-1}, $\alpha_{expt} = 0.035$ s^{-1}) and high ($\alpha_{sim} = 0.127$ s^{-1}, $\alpha_{expt} = 0.127$ s^{-1}) initiation rates.

are low for low initiation rates but high and independent of α for intermediate to high initiation rates of active promoters. This agrees with the observations of [13]. However, promoter repression at $T = 45$ s or at $T = 90$ s does not show any change in the elongation rates for intermediate and high initiation rate (Fig. S6 (inset)), contrary to [13]. Thus, without n dependence and NS diffusion, we cannot capture the negative effect of promoter repression on the co-transcribing RNAPs.

S3. SUPERCOIL DIFFUSION NOT BLOCKED BY TF

In our main model, we hypothesize that the presence and absence of TF on the DNA imposes different conditions of torsional stress on the transcription elongation dynamics. When TF is bound, it blocks NS diffusion and constrains them between itself and the last loaded RNAP. Unbinding of TF immediately results in the dissipation of this torsional stress, and we say that the NS behind the last loaded RNAP can diffuse out and not affect its speed. Keeping the n dependence of the torsional stress identical to the main text, we tried to relax this assumption in two ways - first by looking at the scenario where NS always diffuse out (free DNA as in Sec. S2 B) and second by considering the case where NS never diffuse out (clamped DNA as in Sec. S2 C).

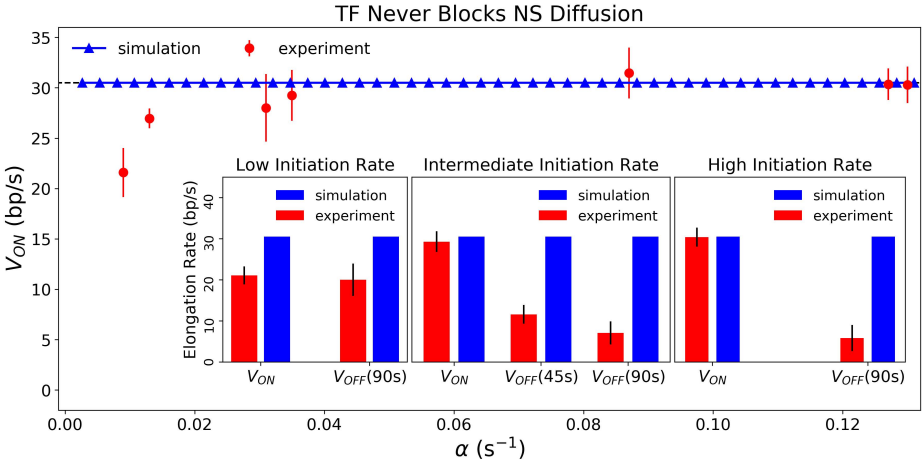

FIG. S7. (color online) Scenario with $f(n)$ as in original model, and free DNA, i.e. TF never blocking NS diffusion. v_{ON} is high and independent of initiation rates for all α, even for a single RNAP ($\alpha_{sim} = 0.006$ s^{-1}). The inset shows the effect of promoter repression. Promoter repression at $T = 45$ s or $T = 90$ s does not appreciable change v_{OFF} from v_{ON} for low ($\alpha_{sim} = 0.006$ s^{-1}, $\alpha_{expt} = 0.009$ s^{-1}), intermediate ($\alpha_{sim} = 0.033$ s^{-1}, $\alpha_{expt} = 0.035$ s^{-1}), and high ($\alpha_{sim} = 0.127$ s^{-1}, $\alpha_{expt} = 0.127$ s^{-1}) initiation rates.

Fig. S7 explores the first case, where TF never blocks NS diffusion but the torsional stress depends on RNAP density. The results are identical to the $f(n) = 1$ case shown in Fig. S5: a fast elongation rate is maintained for *all* initiation rates for both active *and* repressed promoters. This result suggests that without TF blocking NS diffusion (as in linear DNA or less massive TF), increasing torsional stress with RNAP density $f(n)$ is not sufficient to capture the observed cooperative dynamics of RNAPs for the active promoter nor the antagonistic dynamics upon promoter repression.

In the second case, where there is no NS diffusion but with the original $f(n)$, elongation rates drop to zero for all initiation rates and promoter states. This does not exclude the possibility that we can find another $f(n)$ that results in qualitative agreement with experimental observations, even with no NS diffusion. However, our efforts to find such a set of parameters revealed that under no conditions can we simultaneously capture two different phenomena: (i) high elongation rates independent of initiation rates (and hence of RNAP density) in the active state of the promoter and (ii) drastic slow-down upon promoter repression, with lower v_{OFF} for a larger number of RNAPs on the gene. This result implies that LacI, the TF used in the experimental study [13], controls not only RNAP loading events but also the torsional stress of the elongation complexes depending on the ON and OFF states of the promoter. It remains to be tested whether this new role of TF can be found in other TFs or DNA-binding proteins, such as histones in eukaryotic cells.

S4. BURSTY INITIATION

In our model, RNAP loading is assumed to be at regular intervals, motivated by the absence of convoy formation (bursty transcription) under the experimental conditions of [13]. However, [6] proposes that RNAPs loaded close to each other translocate at the same speed and travel as a convoy during elongation. To test this scenario, we explicitly modeled bursty initiation, with 5 RNAPs loading in quick succession within a single burst and a longer duration between bursts. The results are shown in Fig. S8(b) in comparison to nonbursty loading in Fig. S8(a). We find that

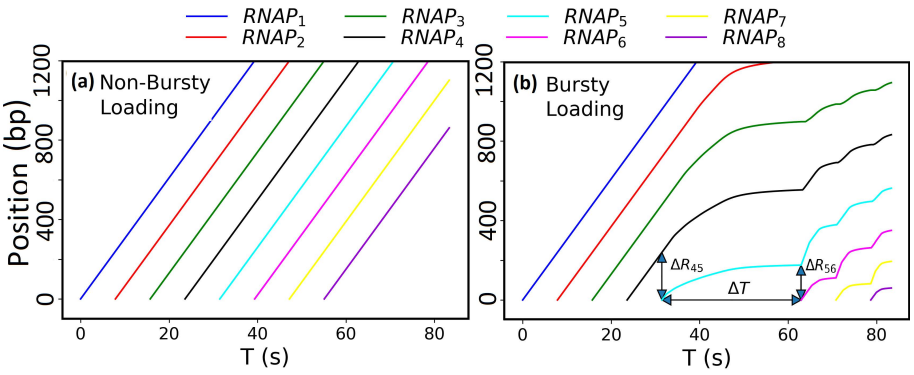

FIG. S8. (color online) Translocation dynamics of RNAPs from nonbursty and bursty initiation. (a) Trajectories of RNAPs from nonbursty initiation at $\alpha = 0.127$ s^{-1}. (b) Trajectories of RNAPs from bursty initiation, where 5 RNAPs are loaded in a burst and $\Delta T > \alpha^{-1}$ is the duration between bursts.

even though the loading of the last RNAP in a burst ($RNAP_5$) and the first RNAP in the next burst ($RNAP_6$) is separated by a longer time duration (ΔT) than the loading of RNAPs within a burst(e.g., $RNAP_4$ and $RNAP_5$), the physical separation between them (ΔR_{56}) is comparable to the separation between RNAPs in a particular burst (ΔR_{45}). Here, the accumulation of NS behind the last RNAP in a burst leads to reducing its speed and thus the physical distance to the RNAP in the next burst. This prediction of our model suggests that convoy formation is hindered even with imposed bursty initiation. Additional mechanisms removing DNA supercoils may help maintain convoys after bursty transcription initiation.