# Mechanical Properties of Transcription

Stuart A. Sevier
*Department of Physics and Astronomy, Center for Theoretical Biological Physics, Rice University, Houston, Texas 77005, USA*

Herbert Levine
*Department of Bioengineering, Center for Theoretical Biological Physics, Rice University, Houston, Texas 77005, USA*

(Received 1 December 2016; published 27 June 2017)

The mechanical properties of transcription have recently been shown to play a central role in gene expression. However, a full physical characterization of this central biological process is lacking. In this Letter, we introduce a simple description of the basic physical elements of transcription where RNA elongation, RNA polymerase rotation, and DNA supercoiling are coupled. The resulting framework describes the relative amount of RNA polymerase rotation and DNA supercoiling that occurs during RNA elongation. Asymptotic behavior is derived and can be used to experimentally extract unknown mechanical parameters of transcription. Mechanical limits to transcription are incorporated through the addition of a DNA supercoiling-dependent RNA polymerase velocity. This addition can lead to transcriptional stalling and resulting implications for gene expression, chromatin structure and genome organization are discussed.

DOI: 10.1103/PhysRevLett.118.268101

The helical nature of DNA introduces a physical dimension to many important biological processes. Most notable is transcription, which is the first step in the conversion of genetic material into biological matter. Though the study of transcription has played a central role in modern molecular biology, much of its physical foundation and behavior is just now being appreciated [1]. Characterizing the physical aspects of transcription may offer insights into many open problems in gene expression and biology.

The physical nature of transcription is conceptualized in the twin-domain model [2], where it was first articulated that transcription and replication cause overtwisting and undertwisting of DNA. The overtwisting or undertwisting of DNA is referred to as supercoiling (SC), and a number of experimental observations have revealed its central role in transcription [3]. Recent results have pointed to SC and mechanical feedback as the sources of transcriptional bursting [4] and domain formation in bacteria [5]. However, a full description of these phenomena is still lacking.

In this Letter, we will introduce a simplified description of transcription with the three fundamental coordinates of DNA rotation, RNA polymerase (RNAP) rotation, and RNA elongation. Because of the helical nature of DNA, linear RNA elongation is coupled to the rotational motion of both RNAP and connected nascent RNA. We will refer to RNAP and nascent RNA collectively as the RNA complex (RNAC). Naturally, the basic coordinates are the RNAC position along a particular gene from the transcription start site (TSS) $x$ and the relative rotation of the RNAC $\theta(x)$ and the DNA $\phi(x)$. These quantities are tied together as

$$\omega_0 x = \phi(x) + \theta(x), \tag{1}$$

where $\omega_0 = 1.85\ \mathrm{nm}^{-1}$ encodes the natural linking number of DNA. The relative difficulty in twisting the DNA (because of opposing torque) or difficulty rotating the RNAC (because of drag) determines the form of the functions $\phi(x)$, $\theta(x)$. The relative amount of DNA twisting $\phi$ and RNAC rotation $\theta$ can be determined by the balance between DNA torque $\tau(\phi)$ and RNAC drag $\Gamma(x, \dot{\theta})$ as

$$\tau(\phi) = \Gamma(x, \dot{\theta}). \tag{2}$$

While many mechanical properties of DNA are well characterized, the mechanical nature of the rotating RNAC is largely unknown. From early studies, however, it is clear that RNA elongation plays a key role [6,7]. Even though this is a critical factor, the coefficient and functional dependence of transcript length on the rotational drag $\Gamma$ are not known at this time. We will posit an RNAC viscous rotational drag, which is linear in the rotation speed, with a power-law dependence on the transcript length as

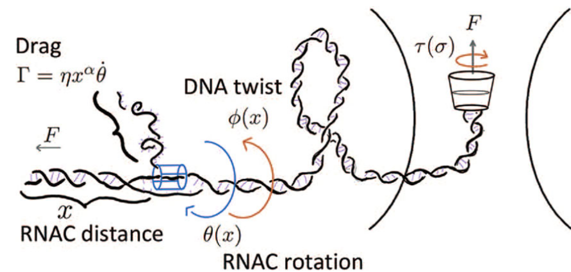


FIG. 1. A cartoon depicting RNA elongation $x$ through shared RNAP $\theta$ and DNA $\phi$ rotation. The DNA is attached to an optical bead for mechanical manipulation.

 © 2017 American Physical Society

$\Gamma = \eta x^{\alpha}\dot{\theta}$, where $\dot{\theta}$ is the angular speed of the RNAC and $\eta$ an unknown coefficient of friction. A length-independent frictional term for the drag associated with the rotation of the RNAP can be included but is not used here, considering the dominant role of elongation in generating SC [6,7]. Previous studies have also linked elongating RNA structure to RNAC activity [8,9]. Dotted and primed marks denote derivatives with respect to time and space, respectively.

The mechanical properties of RNA polymerase itself are well characterized, and it displays constant velocity [10] behavior over a wide range of torque (−20 to +12 pN nm) [11]. We will, therefore, start by assuming a constant elongation rate. Later, we will introduce ways for incorporating the mechanical limits of the RNA polymerase into the motion of the RNA complex.

With an application of the chain rule, we can turn the time derivative for the RNAC rotation into a spatial derivative as $\dot{\theta} = \partial_t\theta(x(t)) = v\theta'$, where $v = \dot{x}(t)$ is the linear velocity of the RNAC. Using this identity with Eqs. (1) and (2), we are left with an equation of motion for DNA twist as a function of RNAP translocation

$$x^{\alpha}\phi' + \frac{1}{\eta v}\tau(\phi) - \omega_0 x^{\alpha} = 0. \tag{3}$$

We now imagine that there is an impediment to the DNA twist a distance $L$ ahead of the TSS. This can be done explicitly in an *in vitro* experiment (shown in Fig. 1) or may occur naturally for DNA *in vivo* with obstructions or other active areas of transcription. Doing this turns the twist equation into an equation for supercoiling density $\sigma(x)$ (SCD) through the substitution $\sigma(x) = \phi/\omega_0(L-x)$; this expression assumes that twisting strain at the point of transcription immediately spreads throughout the specified DNA length (see below). At this time, we will imagine only one barrier to DNA rotation ahead of the RNAC. These assumptions yield the equation

$$\omega_0 x^{\alpha}(L-x)\sigma' - \omega_0 x^{\alpha}\sigma + \frac{1}{\eta v}\tau(\sigma) - \omega_0 x^{\alpha} = 0 \tag{4}$$

Under the additional assumption that the length of the gene is much smaller than the distance to the obstruction, we can drop the $L^{-2}$ terms to find the SC equation

$$x^{\alpha}\sigma' + \frac{1}{\eta v \omega_0 L}\tau(\sigma) - \frac{x^{\alpha}}{L} = 0. \tag{5}$$

To calculate the SCD $\sigma(x)$ as a function of translocation $x$ using Eq. (5), we must specify the torque response of DNA as a function of SC $\tau(\sigma)$. Supercoiling and DNA mechanical dynamics occur on a subsecond time scale [12,13], whereas typical speeds for transcription are 10–50 bp/s [10]. This means that for genes on the order of 1 kbp, transcriptional dynamics happen on the second and minute time scales. Additionally, RNAP operation is robust against subsecond torque fluctuations [11]. Subsequently, as stated above, we expect the locally produced supercoiling at the boundary to spread throughout the allowed DNA segment on a time scale faster than transcription occurs. More generally, we might expect to solve a supercoiling transport equation of the form

$$\frac{\partial\Phi(\tilde{x})}{\partial t} = D\frac{\partial^2\Phi(\tilde{x})}{\partial\tilde{x}^2}, \tag{6}$$

with the boundary conditions for the local twist angle $\Phi = 0$ at $\tilde{x} = L$ and $\Phi = \phi(x)$ at $\tilde{x} = x$. The aforementioned limit appears if the effective rate of strain relaxation $D$ is sufficiently fast, and this will be assumed in what follows. Finally, additional sources of SC dynamics will be ignored.

Thus, the torque $\tau(\sigma)$ response will be that of steady-state supercoiled DNA over a length $L$. In this framework, supercoiled DNA can exist in a purely twisted, purely plectonemic, or a mixed state. Following the phenomenological approach given by Marko [14], the torque in a given piece of DNA held at a constant force is specified by the SCD as

$$\tau(\sigma) = \begin{cases} S\sigma, & \sigma < \sigma_s^* \\ \tau_0, & \sigma_s^* < \sigma < \sigma_p^* \\ P\sigma, & \sigma_p^* < \sigma \end{cases}, \tag{7}$$

where the coefficients $S$, $\tau_o$, $P$, and SC transition values $\sigma_s^*$, $\sigma_p^*$ are given by DNA mechanical constants and are a function of applied force. It is worth noting that the introduction of a well-defined applied force is, at this time, cloudy from an *in vivo* perspective; its experimental implementation is straightforward.

This formulation of $\tau(\sigma)$ yields two types of equations for supercoiling during transcription. The first is for the constant torque response in the region $\sigma_s^* < \sigma_c < \sigma_p^*$,

$$x^{\alpha}\sigma_c' + \frac{\tilde{\tau}_0}{L} - \frac{x^{\alpha}}{L} = 0 \tag{8}$$

and the second for linear torque response in the two regimes $\sigma_l < \sigma_s^*$, $\sigma_p^* < \sigma_l$,

$$x^{\alpha}\sigma_l' + \frac{\tilde{w}}{L}\sigma_l - \frac{x^{\alpha}}{L} = 0, \tag{9}$$

where, in both equations, we have consolidated the constants as $\tilde{\tau}_0 = \tau_0/\omega_0\eta v$ and $\tilde{w} = S/\omega_0\eta v$; $P/\omega_0\eta v$ respectively.

Analytical solutions are obtainable for both the constant and linear SC equations [Eq. (8) and Eq. (9), respectively]. Numerical integration of the full nonlinear Eq. (4)

(including $L^{-2}$ terms) is possible but not considered here. For the constant torque equation [Eq. (8)], direct integration yields

$$\sigma_c(x) = -\frac{x}{L} + \frac{\tilde{\tau}_o}{L}\frac{x^{1-\alpha}}{1-\alpha} + K_1, \tag{10}$$

which describes the SC as a function of RNAC elongation in the hybrid DNA response region. The constant $K_1$ is used to match SCD levels at the boundary.

The linear torque equation [Eq. (9)] can be solved through the decomposition $\sigma(x) = f(x)g(x)$, with $g(x) = K_2 e^{-[\tilde{w}/L(1-\alpha)]x^{1-\alpha}}$. This leads to the solvable equation $f' = 1/g$, which can be used to obtain the full solution

$$\sigma_l(x) = \frac{K_2}{L} e^{-[\tilde{w}/L(1-\alpha)]x^{1-\alpha}} + \frac{1}{L} e^{-[\tilde{w}/L(1-\alpha)]x^{1-\alpha}} \int^x dy e^{+[\tilde{w}/L(1-\alpha)]y^{1-\alpha}}$$

for the SCD as a function of RNAC elongation in the both the pure twist and pure plectonemic regions of DNA response. Again, $K_2$ is a constant used to match boundary conditions. Though the solution is in the form of an integral for arbitrary $\alpha$, simple solutions exist for many rational values of $\alpha$. Typical SCD results $\sigma(x)$ and corresponding torques $\tau(\sigma)$ are shown in Fig. 2 for several barrier distances. For both the pure twisting or pure plectonemic regime, where the SC is governed by Eq. (9), there is an important asymptotic response for the SCD solution $\sigma_l$ in the limit $x^{\alpha-1} \gg \tilde{w}/L(1-\alpha)$,

$$\sigma_l \sim \frac{1}{\tilde{w}} x^{\alpha} \tag{11}$$

so that in the asymptotic limit, the SC and torque directly mirror the drag associated with RNAC rotation. This is an important result as it provides an easy method for measuring the unknown RNAC drag. Reaching the mixed $\sigma > \sigma_s^*$ and full plectonemic $\sigma > \sigma_p^*$ regimes occurs when transcription lengths exceed the values of $x_s^*$, $\sigma(x_s^*) = \sigma_s^*$ and $x_p^*$, $\sigma(x_p^*) = \sigma_p^*$, respectively. Though a closed analytical expression for $x_s^*$, $x_p^*$ is difficult, they can be numerically evaluated and for the case of $x_s^*$, estimated through the asymptotic expression as $x_s^* \sim (\tilde{w}\sigma_s^*)^{1/\alpha}$.

Observations of the most basic elements presented in this Letter can be obtained through laser trap experiments (shown schematically in Fig. 1). Independent determination of the drag coefficient $\eta$ and the functional exponent $\alpha$ is, in principle, possible. This is shown in Fig. 3, where the slopes of $\sigma(x)$ on a log-log plot are determined by the value of the exponent $\alpha$ and the $y$ intercepts through the value of the drag $\eta$.

We expect the functional dependence of the drag on the RNA length to reflect the effective cross-sectional area of the nascent RNA. For a random polymer, the effective cross-sectional area scales as the square root of the polymer length ($\alpha = \frac{1}{2}$); however, higher order RNA structure and electrostatic interactions may drive the drag away from this behavior. Additionally, it is possible that the cross-sectional

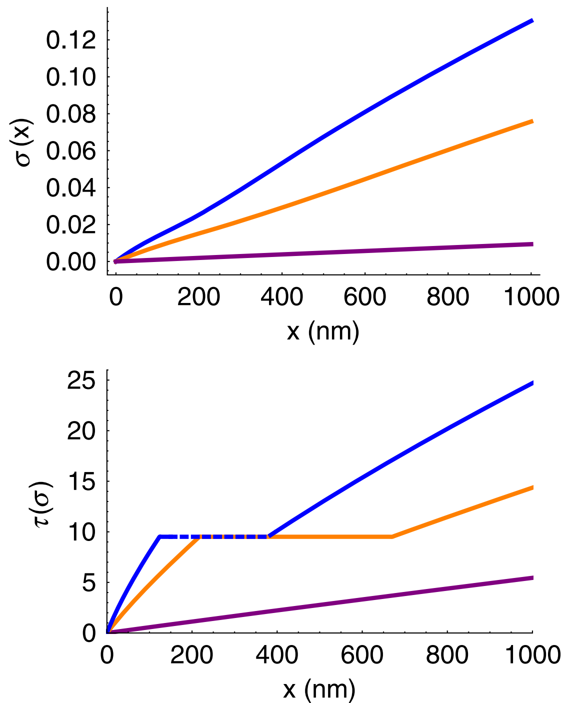


FIG. 2. Supercoiling $\sigma$ and torque $\tau$ as a function of RNA elongation $x$ for $L = 5$, 10, 100 $\mu$m (blue, orange, purple), with $\alpha = \frac{1}{2}$, $\eta v = 1$ at a fixed force of $F = 1/2$ pN nm ($S = 582.0$, $\tau_0 = 9.5$, $P = 189.6$ pN nm).

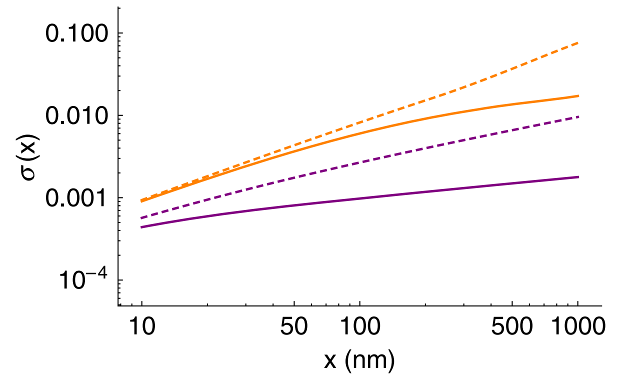


FIG. 3. Supercoiling $\sigma$ as a function of RNA elongation $x$ for $L = 10$ $\mu$m at a fixed force of $F = 1$ pN nm ($S = 622.6$, $\tau_0 = 14.5$, $P = 189.6$ pN nm). The solid and dashed curves are for $\alpha = \frac{1}{2}$ and $\alpha = \frac{1}{4}$, respectively, and the colors are such that $\eta_{\text{orange}}/\eta_{\text{purple}} = 10$. The $y$ intercepts are determined by the drag coefficient $\eta$ and the slopes by the functional exponent $\alpha$. The plot illustrates how the asymptotic behavior of SC and torque would allow for direct measurements of the unknown mechanical parameters of the RNAC.

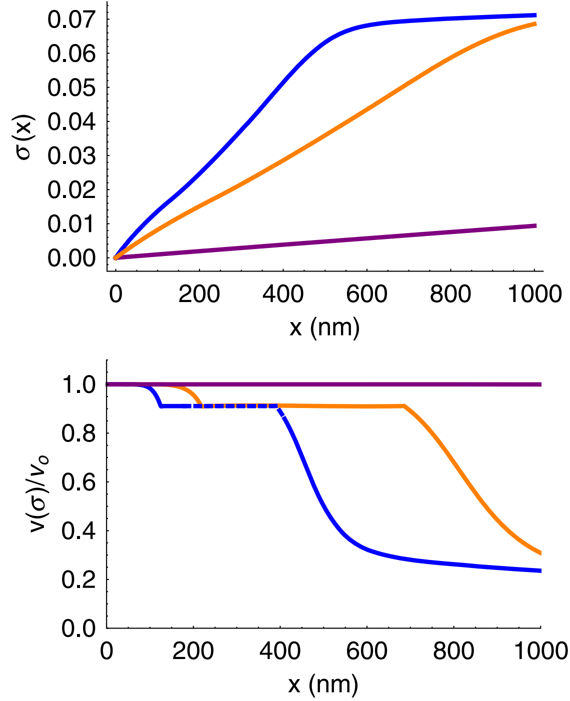


FIG. 4. Supercoiling $\sigma$ and velocity $v(\sigma)$ as a function of RNA elongation $x$ for $L = 5$, 10, 100 $\mu$m (blue, orange, purple), with $\alpha = \frac{1}{2}$, $\eta v = 1$ at a fixed force of $F = 1/2$ pN nm ($S = 582.0$, $\tau_0 = 9.5$, $P = 189.6$ pN nm) for a system with a torque-dependent velocity. For each length, the torque cutoff was set at $\tau_c = 12$ pN nm.

area depends on angular velocity, which would generate higher level terms in Eq. (5). While neglected here, such effects could be incorporated into future models.

So far, we have assumed that RNAP moves at a constant velocity for all SC levels. However, recent connections between the mechanical state of DNA, RNAC progress, and transcriptional bursting have been brought to light both theoretically and experimentally [4,15,16]. These effects are rooted in the stalling torque of RNAP, which we will call $\tau_c$ [11]. A natural way to implement this mechanical limit is through the incorporation of a torque-dependent velocity

$$v(\sigma) = \frac{v_0}{1 + (\frac{\tau(\sigma)}{\tau_c})^n}, \tag{12}$$

yielding a modified equation, Eq. (5), which now includes a stalling component

$$x^\alpha \sigma' + \frac{1}{\eta v_0 \omega_0 L}\left(\tau(\sigma) + \tau_c\left(\frac{\tau(\sigma)}{\tau_c}\right)^{n+1}\right) - \frac{x^\alpha}{L} = 0; \tag{13}$$

this new modified SCD equation can be numerically integrated to yield SC, torque, and velocity curves as a function of elongation (Fig. 4).

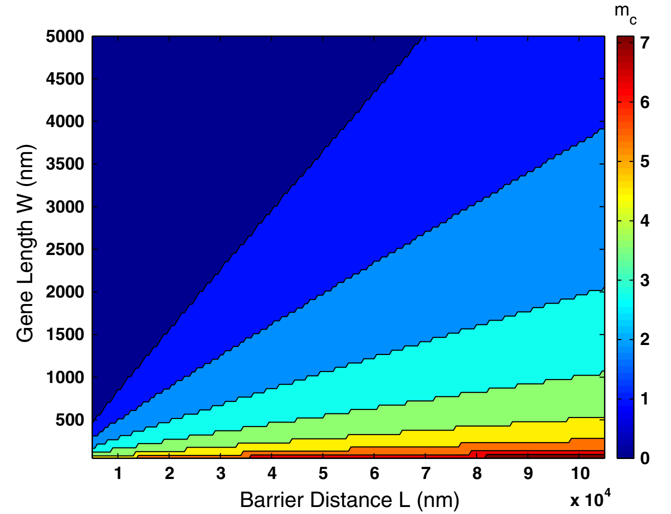


FIG. 5. Figure shows the number of transcripts $m_c$ that can be made for a gene of length $W$ against a barrier of length $L$ with no relaxation between transcription events. The mechanical values used are $\alpha = \frac{1}{2}$, $\eta v = 1$ at a fixed force of $F = 1/2$ pN nm ($S = 582.0$, $\tau_0 = 9.5$, $P = 189.6$ pN nm) for a system with a torque-dependent velocity.

Additionally, the number of transcripts, which can be made before $\tau(\sigma) > \tau_c$ for a gene of length $W$ given a barrier a distance $L$ away, is an important quantity for understanding the competition between mechanical frustration and relaxation in transcription [16]. To examine this, we have integrated Eq. (5) for various $W$, $L$ for a single RNAC operating on a gene and prohibited relaxation between independent transcriptional events. After a single RNAC has reached the end of the gene, the resulting SC level is used as an initial condition for the subsequent integration of Eq. (5). The number of transcripts that can be made before $\tau(\sigma) > \tau_c$ is called $m_c$, and it plays an important role in bounding transcriptional noise due to RNAP stalling [16]. The dependence of $m_c$ on $W$, $L$ is shown in Fig. 5.

It is thus clear that the physical location of a gene within the genome has a significant impact on its ability to be transcribed (Figs. 4 and 5). The values shown are for convenient choices of the RNAC drag parameters and will not necessarily match real values. Experimental measurements of RNAC drag parameters (as outlined earlier in this Letter) would allow for accurate predictions of the interplay between torque buildup and rotational drag. An enlarged framework, which includes a stochastic competition between mechanical frustration (through transcription) and mechanical relaxation (through topoisomerase action), is left for future work and will shed further light on the *in vivo* properties of gene expression.

The structural conformations realized by a particular piece of DNA is constrained by the SCD. The existence of domains in bacterial DNA has been linked to transcription and has been posited to be formed out of plectonemic DNA [5]. The framework presented here can be used to

quantitatively predict the existence of domains between actively transcribing genes. The 3D (as opposed to genomic) distance between the TSSs of two actively transcribing genes or between an actively transcribing gene and some sort of obstruction is captured by the DNA mechanical framework utilized in this Letter. In the absence of adequate relaxation, two actively transcribing adjacent genes with opposing orientation will generate increased SCD in the region of DNA between them. If, in this region, $|\sigma| > |\sigma_s^*|$, the TSSs and the area between them will have increased 3D proximity [14]. For a piece of DNA in the regime $|\sigma| > |\sigma_p^*|$, there will be very little 3D distance between the two TSSs. Therefore, if the value of the constant torque regime $\tau_o$ is below that of the operational limit of the RNAP $\tau_c$ and $x_g > x_p^*$, the 3D distance between the TSSs of two active genes (or obstacle) can be pointlike. This behavior is a possible explanation for the existence of DNA domains in bacteria [5] and eukaryotes [17], as well as a possible mechanism for the formation of transcription factories [15,18].

Thus, it is important to construct a framework which can accommodate simultaneously active RNACs for both the same gene, as well as for neighboring genes. It is straightforward to introduce additional torque constraints on a given piece of DNA, allowing for the incorporation of the action of multiple RNAC on the same gene, as well as multiple barriers to DNA rotation. A model for a supercoiling-dependent multigene system without mechanical arrest or explicit physical components of transcription has already been proposed [19]. Using elements constructed here, it is possible to imagine a system where the relative rotation of multiple RNACs is determined by the mechanical state of shared DNA. Each RNAC will operate using the basic coordinate system and torque balance outlined in Eqs. (1) and (2). For a system of more than two genes, dynamic simulations will likely be needed. A full description of a multigene system will benefit from measurements of the basic elements outlined in this Letter, and a detailed analysis of such a system is left for future work.

This work was supported by the National Science Foundation Center for Theoretical Biological Physics (Grant No. NSF PHY-1427654).

[1] C. Lavelle, Pack, unpack, bend, twist, pull, push: The physical side of gene expression, Curr. Opin. Genet. Dev. **25**, 74 (2014).

[2] L. F. Liu and J. C. Wang, Supercoiling of the DNA template during transcription, Proc. Natl. Acad. Sci. U.S.A. **84**, 7024 (1987).

[3] J. Ma, M. D. Wang, and M. D. Wang, DNA supercoiling during transcription, Biophys. Rev. **8**, 75 (2016).

[4] S. Chong, C. Chen, H. Ge, and X. S. Xie, Mechanism of Transcriptional Bursting in Bacteria, Cell **158**, 314 (2014).

[5] T. B. Le, M. V. Imakaev, L. A. Mirny, and M. T. Laub, High-resolution mapping of the spatial organization of a bacterial chromosome, Science **342**, 731 (2013).

[6] H. Y. Wu, S. Shyy, J. C. Wang, and L. F. Liu, Transcription generates positively and negatively supercoiled domains in the template, Cell **53**, 433 (1988).

[7] Y. P. Tsao, H. Y. Wu, and L. F. Liu, Transcription-driven supercoiling of DNA: direct biochemical evidence from *in vitro* studies, Cell **56**, 111 (1989).

[8] B. Zamft, L. Bintu, T. Ishibashi, and C. Bustamante, Nascent RNA structure modulates the transcriptional dynamics of RNA polymerases, Proc. Natl. Acad. Sci. U.S.A. **109**, 8948 (2012).

[9] P. P. Hein, K. E. Kolb, T. Windgassen, M. J. Bellecourt, S. A. Darst, R. A. Mooney, and R. Landick, RNA polymerase pausing and nascent-RNA structure formation are linked through clamp-domain movement, Nat. Struct. Mol. Biol. **21**, 794 (2014).

[10] K. Adelman, A. La Porta, T. J. Santangelo, J. T. Lis, J. W. Roberts, and M. D. Wang, Single molecule analysis of RNA polymerase elongation reveals uniform kinetic behavior, Proc. Natl. Acad. Sci. U.S.A. **99**, 13538 (2002).

[11] J. Ma, L. Bai, and M. D. Wang, Transcription under torsion, Science **340**, 1580 (2013).

[12] A. Crut, D. A. Koster, R. Seidel, C. H. Wiggins, and N. H. Dekker, Fast dynamics of supercoiled DNA revealed by single-molecule experiments, Proc. Natl. Acad. Sci. U.S.A. **104**, 11957 (2007).

[13] M. T. J. van Loenhout, M. V. de Grunt, and C. Dekker, Dynamics of DNA Supercoils, Science **338**, 94 (2012).

[14] J. F. Marko, Torque and dynamics of linking number relaxation in stretched supercoiled DNA, Phys. Rev. E **76**, 021926 (2007).

[15] X. Darzacq, Y. Shav-Tal, V. de Turris, Y. Brody, S. M. Shenoy, R. D. Phair, and R. H. Singer, In vivo dynamics of RNA polymerase II transcription, Nat. Struct. Mol. Biol. **14**, 796 (2007).

[16] S. A. Sevier, D. A. Kessler, and H. Levine, Mechanical Bounds to Transcriptional Noise, Proc. Natl. Acad. Sci. U.S.A. **113**, 13983 (2016).

[17] S. S. P. Rao, M. H. Huntley, N. C. Durand, E. K. Stamenova, I. D. Bochkov, J. T. Robinson, A. L. Sanborn, I. Machol, A. D. Omer, E. S. Lander, and E. L. Aiden, A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping, Cell **159**, 1665 (2014).

[18] C. S. Osborne, L. Chakalova, K. E. Brown, D. Carter, A. Horton, E. Debrand, B. Goyenechea, J. A. Mitchell, S. Lopes, W. Reik, and P. Fraser, Active genes dynamically colocalize to shared sites of ongoing transcription, Nat. Genet. **36**, 1065 (2004).

[19] C. A. Brackley, J. Johnson, A. Bentivoglio, S. Corless, N. Gilbert, G. Gonnella, and D. Marenduzzo, Stochastic Model of Supercoiling-Dependent Transcription, Phys. Rev. Lett. **117**, 018101 (2016).