



The Role of Supercoiling in the Motor Activity of RNA Polymerases

Annick Lesne, Jean-Marc Victor, Edouard Bertrand, Eugenia Basyuk, and Maria Barbi

Abstract

RNA polymerase (RNAP) is, in its elongation phase, an emblematic example of a molecular motor whose activity is highly sensitive to DNA supercoiling. After a review of DNA supercoiling basic features, we discuss how supercoiling controls polymerase velocity, while being itself modified by polymerase activity. This coupling is supported by single-molecule measurements. Physical modeling allows us to describe quantitatively how supercoiling and torsional constraints mediate a mechanical coupling between adjacent polymerases. On this basis, we obtain a description that may explain the existence and functioning of RNAP convoys.

Key words Molecular motor, RNA polymerase, Supercoiling, Convoy, Torsional constraints, Physical modeling, Force, Torque, DNA

1 Introduction

Increasing evidence has been obtained for a long time of the *in vivo* occurrence of *DNA supercoiling*, i.e., the overwinding (or underwinding) of the DNA superhelix, and its functional role [1–4]. DNA supercoiling can be measured both *in vitro* and *in vivo*, and it is observed in a number of biological processes, for instance DNA compaction and transcriptional regulation [5]. Supercoiling may rescue transcription by generating a negative torque, and it affects the elongation rate [6].

Supercoiling regulatory role is presumably due to the mechanical coupling between the DNA-bound molecular motor and its DNA substrate. Because of this expected mechanical coupling, the description of supercoiling-based regulation of molecular motor activity involves a physical view on molecular motors, with a focus on the role of topological constraints and torques [7–9]. We will focus in this chapter on the elongating phase of

RNA-polymerase (RNAP, [10]), for which a lot of data are available.

Remarkably, the interplay between RNAP progression and supercoiling appears to be reciprocal. On the one hand, the progression of the elongating enzyme can be hindered, or facilitated, by the topological state of its DNA substrate. On the other hand, a transcribing RNAP exerts a mechanical torque on the DNA that will modify its supercoiling. Both the oriented motion and torque generation by the RNAP are achieved by the coupling of RNAP conformational transitions with the far-from-equilibrium NTP hydrolysis reaction. The DNA structure itself plays a crucial role in this coupling, due to the intrinsically chiral structure of the double helix. The joint presence of an oriented substrate and nonequilibrium fluctuations is indeed mandatory to produce a mechanical work [11].

As an illustration of the importance of a physical approach in this context, we will give a detailed account of a scenario that centrally involves this reciprocal coupling. When two or more transcribing RNAPs follow each other, the modification of DNA supercoiling by polymerase activity and its feedback on polymerase velocities induce a torsional coupling between adjacent polymerases, enforcing the constancy of their separation along DNA. As a result, polymerases progress in the form of a “polymerase convoy” [12]. We will quantitatively describe this phenomenon and derive some important consequences on the RNAP convoy efficiency.

2 DNA Supercoiling

2.1 *Constrained DNA Supercoiling in Bacteria, Archaea and Eukaryotes*

DNA supercoiling can be described in mathematical terms, starting with the notion of *linking number*: the number of turns that one strand makes around the other. In a circular DNA (plasmid) or a DNA segment with fixed ends, the linking number is a topological invariant. Either the variation of the linking number with respect to a relaxed configuration, or the relative variation termed the *superhelical density*, can be used to measure the amount of supercoiling. The linking number can be decomposed in two contributions: the *writhe* (corresponding to the spatial torsion of the double-helix axis) and the *twist* (corresponding to the winding of the strands around this axis). When DNA is stretched (e.g., by a pulling force in micromanipulations), the writhe vanishes and the linking number is entirely stored in the form of twist [13]. On physical grounds, a change in the twist induces a change in the DNA torsional elastic energy, while writhe is associated with bending energy. Since this stored elastic energy may be released and converted, DNA supercoiling will have direct mechanical consequences.

DNA supercoiling can be observed and measured experimentally. In vitro, negative supercoils have been detected in bulk DNA through preferential binding of psoralen to undercoiled regions. In vivo, in experiments using psoralen photobinding to probe DNA topology, Kouzine et al. [14] were able to measure the spreading of supercoiling (about 1.5 kilo-basepairs (kb) upstream the start sites of active genes) locally induced by a transcribing RNA polymerase in human Burkitt lymphoma cells.

Part of DNA supercoiling can be termed *constrained*, in the sense that it is maintained by the presence of architectural proteins. In bacteria, the action of nucleoid-associated proteins (NAPs) may constrain and stabilize DNA supercoiling by locally bending or wrapping DNA [15, 16]. In particular, the mechanical interplay between DNA supercoiling and the architectural properties of histone-like nucleoid structuring proteins (H-NS) bound to DNA has been investigated by magnetic tweezers [17]. The NAP Factor for Inversion Stimulation (FIS) also modulates the topology of bacterial DNA in a growth-phase dependent manner both by reducing the expression of DNA gyrases and by a self-regulation mechanism that depends on the level of negative superhelical density [18].

Such a constrained form of supercoiling is the main compaction strategy adopted in archaea and especially in eukaryotes, with the DNA quite regularly wrapped on specific octameric *histone* complexes and forming *nucleosomes* [19]. Wrapping of DNA around the histone core introduces in DNA an average supercoiling of -1 turn per nucleosome, with some variation depending on the details of the nucleosome architecture and relative position of the linkers. The supranucleosomal organization can then induce an additional supercoiling [20, 21].

2.2 Free DNA Supercoiling in Bacteria, Archaea, and Eukaryotes

In addition to the constrained supercoiling that characterizes DNA packaging in equilibrium situations in all three domains of life (bacteria, archaea, and eukaryotes), additional nonconstrained, or *free* supercoiling can be actively generated in nonequilibrium situations, typically by the activity of DNA-bound motor proteins [4]. In bacteria, the (generally) circular DNA is observed to be negatively supercoiled, and displays topologically independent domains having an average size of 10 kb [22]. DNA supercoiling spreads freely inside each of these domains. The maintenance of such a high level of negative supercoiling is possible thanks to the action of a specific bacterial enzyme, the DNA gyrase, a special topoisomerase II enzyme that is able to introduce negative supercoils in DNA [23]. DNA supercoiling can also be removed by topoisomerase I activity, which relaxes excess supercoiling by creating a single-strand break (also termed a DNA “nick”). In practice, the time scale of topoisomerase activity has to be compared to the time scale of the considered functional events to appreciate its

impact. In *E. coli*, a steady-state level of supercoiling is thus actively maintained by the balanced activities of gyrase and topo I. Supercoiling in turn regulates the expression level of a significant number of genes, functionally diverse and widely dispersed throughout the chromosome. This global effect has been evidenced using novobiocin, a competitive inhibitor of the ATPase reaction catalyzed by the gyrase subunit *gyrB* [24].

Interestingly, hyperthermophile microorganisms, which live optimally at temperatures above 80 °C, including some bacteria but mostly archaea, have their DNA (generally) *positively* supercoiled thanks to reverse gyrase, an ATP-dependent topoisomerase I. This atypical behavior may be understood in the following way: in usual microorganisms, which live at lower temperatures (and therefore called mesophilic), most DNA transaction processes, such as transcription or replication, require DNA melting which is favored by negative supercoiling. On the contrary, at temperatures above 80 °C, DNA is prone to melting, and hence positive supercoiling prevents DNA denaturation [25].

Active supercoiling generation is also present in eukaryotes. Supercoiling domains, sharing distributed free supercoiling excess, have been brought out recently [26]. They were shown to largely overlap with *topologically associating domains*, i.e., linear units of chromatin that fold as discrete three-dimensional (3D) structures tending to favor internal interactions, as evidenced in genome-wide measurements of chromosomal contacts [27–29]. Supercoiling is confined in these topological domains, which are separated by the domain borders, highly enriched in binding sites for an important architectural protein, CTCF (*CCCTC-binding factor*).

2.3 Supercoiling Confinement and Spreading

Free supercoiling may therefore be locally generated, hence distributed over a given genome domain. Physics modeling may help to determine the characteristic time and space scales of its spreading, hence to identify the mechanisms underlying its confinement.

The propagation of DNA torsional stress has been studied theoretically in [8] at the chromatin fiber scale. We recall here the basic results: For a straight fiber of length L , with a linking number Lk , the twist rate is $Tw = Lk/L$, with $Tw^0 = Lk^0/L$ in relaxed conditions. Hence the superhelical density is:

$$\sigma(x, t) = (Lk - Lk^0)/Lk^0 = (Tw - Tw^0)/Tw^0.$$

Because inertial effects are negligible at this scale, $\sigma(x, t)$ is solution of a plain diffusion equation:

$$\partial\sigma/\partial t = D\partial^2\sigma/\partial t^2.$$

where

$$D = k_B T L_{PT}^f / \eta R^2.$$

with k_B the Boltzmann constant, L_{PT}^f the twist persistence length of the chromatin fiber, η the viscosity coefficient of the surrounding medium, and R the radius of the fiber. Consequently, the time for the torsional stress to spread all over a chromatin topological domain of size L is:

$$T_{\text{spread}} \approx L^2 / D.$$

For $L_{PT}^f = 30$ nm, $R \approx 15$ nm and $\eta \approx 10^{-2}$ Pa·s (never more than ten times the viscosity of pure water), the diffusion coefficient is $D \approx 5 \times 10^{-11}$ m²/s. In mammals, the length L of a topological domain corresponds to hundreds of kb, hence the order of T_{spread} is tens of ms at most in compact chromatin (for which a kb amounts to a length of 2–5 nm). It is even shorter within subdomains called chromatin loops [29]. Propagation of torsional stress is thus much faster than the progression of the RNAP along the DNA [8].

These estimations are made for rather compact chromatin. One may therefore wonder how the result would change in the case of either a less compact nucleosome array, or for free DNA. In this case, the persistence length will be approximately of the same order of magnitude but generally slightly larger (about 50 nm for naked DNA), while the radius R of the assembly is reduced by a factor 3 for a bead-on-a-string fiber, and up to a factor 15 if one considers the crystallographic radius of naked DNA, $r_{\text{DNA}} = 1$ nm. As a consequence, the time T_{spread} will be further reduced by up to two orders of magnitude. Therefore, a torsional stress induced locally may in any case be considered to propagate instantaneously, hence to be in quasi-static equilibrium between the next boundary elements upstream and downstream. Supercoiling confinement can only be achieved by insulator elements at the boundaries.

It should be noted, however, that the previous derivation does not account for the role of three-dimensional structures that may capture torsional constraints in the form of writhe, as *plectonemes*. In vitro, these extended DNA braids will grow in length when additional supercoiling is added, by converting more and more torsion into writhe. This process generally goes with a rotation of the whole plectonemes in space, so that a much larger effective radius would intervene, in principle. In vivo, however, steric hindrance will presumably prevent such large structures to rotate, if they exist. Hence, we expect torsional constraints to only spread as twist along DNA or a chromatin fiber, instead of inducing large conformational rearrangements.

2.4 How Supercoiling May Be Inherited

The fact that topologically associating domains are associated with given supercoiling suggests a physiological function for supercoiling. The question then arises of how this supercoiling may be inherited through cell division. DNA supercoiling in a given topological domain is a function of (a) *topoisomerase* activity, acting on free supercoiling (twist) and (b) *remodeling* activity, mainly acting on DNA architecture hence on constrained supercoiling (writhe).

The localized action of topoisomerases and structural proteins such as *condensins* participates in determining some domain structuration, and its associated topological states. On the other hand, active remodeling processes achieved by ATP-consuming remodeling factors—and crucially through active nucleosome removal—have been shown to be essential for driving biologically relevant nucleosome positioning [30]. Indeed the fractions of free and constrained supercoiling in a domain are evidently tuned by the total number of nucleosomes in the domain, hence by the average distance (along DNA) between neighboring nucleosomes, called the *nucleosome repeat length* (NRL). Active nucleosome removal thus tunes the amount of free supercoiling by fixing the average NRL.

Importantly both these ATP-consuming mechanisms are under active control of the cell metabolism. Since both remodeling and topoisomerase activities determine the overall architecture of a given domain, they also tune its supercoiling state. If this architecture can be transmitted through cell division, then supercoiling may be inherited by the same mechanism. In this sense, the average NRL of a genomic domain appears to be a physical epigenetic mark of this domain [31]. It has been shown, indeed, that the distribution of topoisomerases and condensins may help in transmitting some domain structuration and topological states through cell division [32, 33]. Interestingly, the recently observed spreading mode of histone *posttranslational modifications* (PTMs) over transcription cycles [34] might explain the spreading and maintenance of the NRL on epigenetic domains. Challenging genome-wide studies are needed to further correlate supercoiling maps to cell differentiation states.

3 Coupling Polymerase Activity with Supercoiling In Vivo

The complex coupling between transcription and supercoiling in the context of eukaryote chromatin raises many questions and has inspired experimental and theoretical work, reviewed in [35]. One prototypical nonequilibrium situation in which the action of a DNA-bound molecular motor is coupled with DNA supercoiling is transcription by RNAP. It is to note that both free and constrained supercoiling may be coupled with RNAP transcriptional activity. Importantly, the coupling is reciprocal: a variety of experimental results indeed demonstrate that transcription both depends on and controls transcriptional activity [14].

3.1 Transcriptional Regulation of Supercoiling: Torque Exerted by a Polymerase

In bacteria, a model in which topoisomerases act near highly transcribed operons in order to equilibrate the supercoil flux generated by transcription suggests that matched rates of gyrase turnover and transcription elongation speed determine the average supercoil density [36]. Transcription-dependent supercoiling has also been observed genome-wide in *Drosophila* and human cells ([14, 37]. Supercoiling generation by a processing RNAP is due to the requirement of a relative rotation of the enzyme (and its nascent RNA chain) with respect to DNA. In the *twin supercoiled-domain model*, DNA screws into the polymerase, and experiences therefore supercoiling downward and undercoiling in the wake of the transcriptional activity [38]. The opposite view of RNAP rotating around DNA in its processing activity is generally excluded because of the high friction force that RNAP would experience (compared to that experienced by DNA during its screwing motion) and the topological entanglement of the nascent RNA around DNA. The high molecular mass of the transcription machinery and some experimental evidence further support this idea [39]. The resulting negative supercoiling experienced upstream induces DNA denaturation, whereas positive supercoiling induces either plectoneme formation or nucleosome conformational transitions. All these transitions (denaturation, plectoneme formation, nucleosome conformational transitions) absorb each in their own way part of the induced supercoiling, thus buffering the torsional stress.

The analysis of the supercoiling generated by an active polymerase and its consequences has been extended to heterochromatin in [8, 9]. On the basis of single molecule experiments with magnetic tweezers [40], it has been argued that nucleosomes can be turned into *reversomes*. In these metastable particles (stabilized precisely by the positive torque applied on the DNA), the nucleosomal DNA has a reverse chirality, as well as the histone core. The transition of one nucleosome into one reversome has been shown to absorb two turns of positive supercoiling (i.e., the linking number decreases by 2 units). A prediction of this model is a wavefront of reversome propagating ten times faster than the RNAP. Strikingly such a wavefront has been evidenced during the transcription of the *Drosophila* Hsp70 locus [41]. Most importantly, the wavefront was confined inside the Hsp70 locus, strongly indicating that supercoiling stops at the locus boundary.

3.2 Transcriptional Regulation by Supercoiling: Control of the Transcription Rate

Dependence of patterns of transcriptional activity on supercoiling has been observed for long in bacteria [42]. By combining the effects of mutations in the *fis* and *hns* genes with experimentally induced changes of global superhelicity [43], argue that a global coordination of transcription during bacterial growth is achieved by the regulation of the supercoiling of the circular genome, obtained by topoisomerase activity or binding of architectural proteins onto DNA. Torsional stress may also modulate protein affinity

[44]. Torsional constraints affect the elongation rate and even rescue transcription by applying negative turns (thus reducing the torque experienced by the enzyme) [6]. It has been observed in vitro that positive supercoiling accumulation during viral (T7) and bacterial polymerase processing may result in a decrease of transcription elongation and eventually stop transcription initiation; transcription can be then resumed by the action of gyrases releasing the supercoiling [39]. In view of these results, Travers and Muskhelishvili [5] proposed that transcription depends on an effective superhelical density resulting from the balance between the action of topoisomerases, transcription machinery and chromatin proteins.

4 Polymerase Force and Torque Measurements in Single Molecule Manipulations

To unravel the reciprocal coupling between polymerase activity and DNA supercoiling in the mechanistic point of view adopted here, single-molecule experimental data and modeling based on these data turned to be essential. They allowed one to measure the force exerted by a processing polymerase, or equivalently the force required to stop a processing polymerase (*stall force*), the torque exerted by a processing polymerase, and the propagation of the ensuing torsional constraints.

The advent of single-molecule techniques permitted the elucidation of the successive structural rearrangements of the motor proteins during their motion and activity, and to measure the kinetics of these steps [45]. Seminal works have been the design of optical trapping assays [46] and tethered particle motion assay tracking down the motion of the motor with a spatial resolution down to a single base pair [47]. Other techniques include atomic force microscopy, fluorescence exchange transfer and particle tracking [48]. In particular, optical and magnetic tweezers or atomic force microscopy (AFM) approaches consist in attaching one end of the molecule to a surface and the other end to a probe (either an optically trapped bead, a magnetic bead, or an AFM tip), through which force (and, in some cases, torque) is applied. Elongation activity of RNA polymerase have been investigated in depth using these techniques, providing invaluable data on the parameters influencing the elongation speed and the pausing mechanism [47, 49–52].

4.1 Force Measurements

Initially, the effect of a stretching force opposing to transcription has been investigated [52, 53]. In these experiments, a viral (T7) RNAP, tethered on a glass plate, is transcribing a DNA molecule tethered to a micrometric bead that is manipulated by *optical tweezers*. While the DNA template screws through the polymerase, a stretching force can be applied to the bead thus opposing the

motion. Force-elongation curves show that the elongation is independent of the applied force provided it is smaller than a limiting stall force. It is however highly fluctuating, ranging between 50 and 130 bp/s according to the individual enzyme and the run. These variations in the elongation velocity are presumably due to differences in the protein structure, especially when the protein is composed of many subunits. Recent *in vivo* measurements confirm this order of magnitude for the elongation velocity [12, 54]. Pauses are also observed, and attributed to obstacles, amounting to locally apply a force larger than the stall force. The measurement of the stall force allows one to give an upper limit to the force exerted by RNAP. For the viral T7 RNAP, this force is in the range 20–25 pN [52], and can be larger than 25 pN for *E. coli* RNAP [55].

4.2 Torque Measurements

Twist degree of freedom and torsional constraints are ignored in force experiments, since optical tweezers do not allow to constraint the DNA molecule in rotation. In contrast, twist effects have been considered in *magnetic tweezers* experiments, allowing one to apply a torsional constraint to DNA and thus investigate transcription under torsion [56, 57]. In particular, single-molecule experiments have been designed to measure the torque generated by a polymerase, and shown it to be larger than 5 pN·nm, i.e., $1.25 k_B T$ [6, 58].

A major result of these experiments is the invariance of the torque generated, which happens to be independent of the RNAP processing velocity. This can be understood as follows. Let K be the torque generated by a processing RNAP. The energy supplied by the polymerase during one complete turn of the DNA double helix inside the RNAP is $2\pi K$. It can be estimated from the characteristics of the polymerization reaction, namely it is fixed by the *number* of dNTP converted into PP_i (pyrophosphate) during one complete turn of the DNA into the polymerase. This number is simply 10.6, because one dNTP is converted into one PP_i for each transcribed base pair. In contrast, the angular velocity ω appears to be strongly dependent on the dNTP concentration [52]. Indeed the dNTP concentration modulates the kinetics of the chemical reactions involved in the polymerization cycle, hence the period T_{turn} of a one-turn screwing of DNA inside the polymerase and eventually the angular velocity $\omega = 2\pi/T_{\text{turn}}$.

By simultaneously measuring rotation, torque, displacement and force for a *E. coli* RNAP with a specially designed technique, it was shown that positive supercoiling make the pause-free velocity decrease, while the pause density and the duration of pauses are both increased [56, 59]. In this experiment, stalling is observed at a torque of approximately 10 pN·nm ($2.5 k_B T$). Similarly, the enzyme could work against a negative supercoiling of the order of 10 pN·nm.

5 From One RNAP to RNAP Convoys

Multiple RNAPs can act simultaneously on a single DNA template. The distance between adjacent polymerases (a few hundreds of bp in the electron-microscopic observation of [60]) prevents any direct interaction. However, recent *in vivo* observations suggest that the array of active polymerases elongate at the same velocity, forming a polymerase convoy progressing as a whole [12]. A first mechanical argument supporting the existence of polymerase convoys is the fact that supercoiling upward (resp. downward) a polymerase convoy is the same as the supercoiling upward (resp. downward) a single polymerase: the management of torsional constraints generated by transcription is the same for a convoy and a single unit [61].

Interestingly, RNAP cooperation has been shown to represent an important advantage in enhancing transcription efficiency. Indeed various DNA binding proteins can act as roadblocks, hindering or even interrupting the elongation of a single RNAP. However, several RNAPs transcribing in sequence have been observed to proceed through roadblock proteins either *in vitro* or *in vivo* [62]. We will quantify this effect below, by estimating the stall force necessary to stop a RNAP convoy, and showing how multiple polymerases enhance transcription processivity.

We here go further and analyze quantitatively how a local desynchronization (pausing or slowing down of a polymerase affecting the spacing to neighboring polymerases) reflects in DNA supercoiling and local increase of torsional energy. We expect that such energetic cost will penalize RNAP desynchronization and result in an effective force, computed below, tending to keep constant the distance between polymerases and maintain convoy cohesion.

5.1 Effective Forces Generated by Torsional Coupling

We consider a convoy of N elongating RNAPs. We neglect here the role of chromatin, because RNAP convoys happen in highly active transcriptional conditions for which nucleosomes are mostly disrupted. We moreover neglect end constraints. We label polymerases, $RNAP_i$, in the order of their loading onto DNA. We assume that polymerases actively perform a translation along the DNA axis, due to a strong viscous force preventing the rotation of the transcription complex, while DNA screws into the polymerases, according to the twin supercoiled-domain model described above. The viscous force against DNA rotation is assumed to be negligible.

We note $\Delta_i = x_i - x_{i+1}$ the distance between $RNAP_i$ and $RNAP_{i+1}$ at a given time (Fig. 1). If the two polymerases move forward at the same constant velocity, this distance remains equal to its value Δ_i^0 right after the loading of $RNAP_{i+1}$ (fixed by the promoter activity). Any change in the distance Δ_i reflects in a

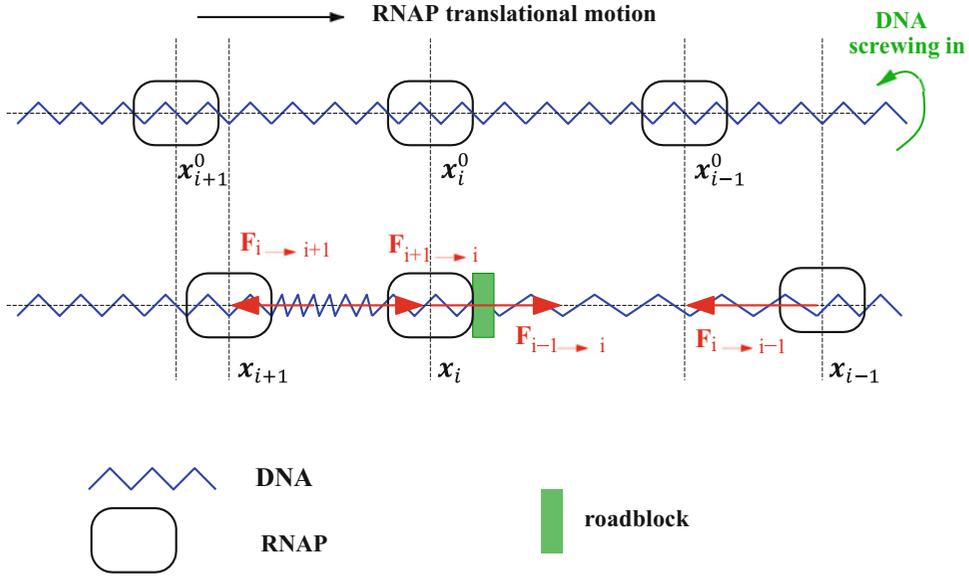


Fig. 1 Physical modeling of an elongating polymerase convoy. Top: DNA screws into an elongating polymerase convoy. Bottom: a polymerase stops at a roadblock (in green) and generates supercoiling constraints with the preceding and succeeding polymerases. $F(i + 1 \rightarrow i)$ indicates the force exerted on the i^{th} polymerase (RNAP) by the $(i + 1)^{\text{th}}$ one (RNAP $_{i+1}$)

supercoiling of the DNA stretch separating RNAP $_i$ and RNAP $_{i+1}$, since the total linking number Lk^0_i trapped between them remains constant (note that Lk^0_i is an integer number of turns because all the RNAPs have the same rotational position). Considering that the DNA stretch has a negligible writhe, the twist rate is $Tw_i = Lk^0_i/\Delta_i$, with $Tw^0_i = Lk^0_i/\Delta^0_i$ in relaxed conditions. The corresponding increase is

$$\Delta Tw_i = Tw_i - Tw^0_i = Tw^0_i [(\Delta^0_i - \Delta_i)/\Delta_i]$$

The relaxed twist rate accounts for the number of turns per unit length for unconstrained DNA, hence it can be simply written $Tw^0_i = 1/h$, where h the pitch of the relaxed DNA double helix (generally $h = 3.4$ nm corresponding to one turn per 10.6 bp, but this parameter can depend on global cellular constraints for instance in the case of bacteria). Denoting L_{PT} the twist persistence length of DNA, the potential energy of torsion stored in the DNA stretch (between RNAP $_i$ and RNAP $_{i+1}$) can be written as.

$$E_{Pi} = \frac{1}{2}k_B TL_{PT}[2\pi\Delta Tw_i]^2\Delta_i = \frac{1}{2}k_B TL_{PT}(2\pi/h)^2(\Delta^0_i - \Delta_i)^2/\Delta_i \approx \frac{1}{2}k_B TL_{PT}(2\pi/h)^2(\Delta^0_i - \Delta_i)^2/\Delta^0_i.$$

where the last approximation is valid in the limit of small distortions, when $(\Delta^0_i - \Delta_i)^2 \ll \Delta_i$. In this approximation, the effective interaction between RNAP $_i$ and RNAP $_{i+1}$ is harmonic,

i.e., the DNA stretch between them acts as a torsional spring. The RNAP convoy is therefore analogous to a system of N harmonically coupled particles in collective motion.

The effective force $F(i+1 \rightarrow i)$ exerted on $RNAP_i$ by $RNAP_{i+1}$ (see Fig. 1) is obtained by deriving the potential energy:

$$\begin{aligned} F(i+1 \rightarrow i) &= -\partial E_P / \partial x_i \approx k_B T L_{PT} (2\pi/h)^2 (\Delta_i^0 - \Delta_i) / \Delta_i^0 \\ &\approx k_B T L_{PT} (2\pi/h)^2 \sigma_i. \end{aligned}$$

In the second line of this equation, we have expressed the force in terms of the local DNA supercoiling σ_i , which in our case also accounts for the relative length variation since

$$\sigma_i = (Tw_i - Tw_i^0) / Tw_i^0 = (\Delta_i^0 - \Delta_i) / \Delta_i \approx (\Delta_i^0 - \Delta_i) / \Delta_i^0.$$

The numerical value of the prefactor $k_B T L_{PT} (2\pi/h)^2$ is about 1300 pN ($k_B = 1.38 \times 10^{-23}$ JK⁻¹, $h = 3.4$ nm, $T = 300$ K, and $L_{PT} = 90$ nm). For a relative decrease of 1% of the distance between the two adjacent polymerases, i.e., for $\sigma_i = 0.01$ (compatible with the small distortion assumption), the force $F(i+1 \rightarrow i)$ pushing $RNAP_i$ forward is already as large as 13 pN (recall that the maximal force that can be exerted by a RNAP is around 25–30 pN). Note that if, on the contrary, the lagging polymerase is $RNAP_{i+1}$, the force $F(i+1 \rightarrow i)$ pulls $RNAP_i$ backward and slows it down.

This push-pull mechanism strongly stabilizes the distance between successive polymerases. With Δ_i^0 of the order of a few hundreds of bp (as confirmed by recent *in vivo* experiments by Tantale et al. [12], measuring an average distance of about 200 bp), a relative variation of 1% of this distance amounts to a few bp: this means that the distance between neighboring RNAPs is exquisitely regulated by supercoiling constraints, at the base pair scale.

5.2 Collective Effects and Convoy Propagation as a Whole

We now address the complete problem of the motion of N harmonically coupled polymerases. This problem is reminiscent of some issues in traffic theory, e.g., the motion of an array of harmonically coupled vehicles [63]. More precisely, let us now consider any opposing action exerting a maximal resistive torque of magnitude M onto *one* polymerase $RNAP_i$. This torque M may be due for instance to an experimental setup, a topological constraint acting on the system, or a protein roadblock (Fig. 1). We consider the (most likely) case when the external torque M acts on the first polymerase, $RNAP_1$, and ask in which conditions the RNAP convoy can overcome this external obstacle.

We denote K the torque exerted by any active polymerase. As discussed above, K is fixed by the number of dNTP converted into PP_i (pyrophosphate) during one complete turn of the DNA into the polymerase. The angular velocity ω of DNA screwing motion generally depends on the dNTP *concentration* [52]. However the

dNTP concentration is fixed in vivo, hence the angular (and linear) velocity too. To obtain an estimate of K , we note that this torque should relate to the stall force F_S , as RNAP progresses along DNA according to a relative screwing motion. In a general manner, the force f and the torque Γ for a motor screwing with respect to DNA can be related by comparing the power that both stresses deliver for a given displacement along DNA, namely $fV = \Gamma\omega$, with V and ω the linear and angular velocities, respectively. By taking into account the coupling between linear and angular displacements due to the helical constraint, this leads to write $fh = \Gamma 2\pi$, or $\Gamma = f(h/2\pi)$. The force f and the torque Γ are said to form a “wrench.” By using this rule, we can therefore identify K to the effective stall torque $\Gamma_S = F_S (h/2\pi)$. Accordingly, the measure $F_S \approx 17$ pN for yeast RNAP [47] yields the estimate $K \approx 9$ pN·nm. Interestingly, recent measures using an angular optical trap, allowing for a direct measurement of the torque [56], give a stall torque of 11 pN·nm, in fair agreement with our theoretical calculation.

If the external torque M opposing the motion of $RNAP_1$ is small enough, i.e., if $M < K$, then $RNAP_1$ will keep on moving at the same velocity V through the roadblock. On the contrary, if $M > K$, the roadblock eventually blocks the polymerase activity and stops $RNAP_1$. But the next polymerase $RNAP_2$ is still moving at velocity V , hence the distance Δ_1 decreases with time. Consequently, a force $F(2 \rightarrow 1)$ acts on $RNAP_1$ as previously discussed. At the same time, $RNAP_2$ experiences an opposite force $-F(2 \rightarrow 1)$, namely it is pushed backward by the reaction force. By using again the force-torque conversion rule, the force $F(2 \rightarrow 1)$ results in an effective torque $\Gamma(2 \rightarrow 1) = F(2 \rightarrow 1) (h/2\pi)$. This pushing effect will make the blocked polymerase resume its motion as soon as $K + \Gamma(2 \rightarrow 1) > M$. Of course, this is only possible if the required torque $\Gamma(2 \rightarrow 1) = M - K$ is smaller than the maximum RNAP torque K , i.e., if $M < 2K$. Otherwise the second RNAP would also stop before the first one has resumed its motion. When $M > 2K$, once both $RNAP_1$ and $RNAP_2$ had stop, the disturbance propagates and the next part of the convoy enters the scene: $RNAP_3$ starts pushing $RNAP_2$, with a force of increasing magnitude as it proceeds. The above scenario applies and the constraints keep on accumulating on $RNAP_1$ until either $RNAP_1$ resumes its motion (if $M < 3K$) or $RNAP_3$ stops (if $M > 3K$). The iteration of this reasoning leads us to conclude that the RNAP stall torque within a convoy of N polymerases may be as large as NK .

Importantly, the same result is obtained if the external constraint applies not to the first enzyme $RNAP_1$ but on any other $RNAP_i$ within the convoy. In this case, both flanking polymerases $RNAP_{i-1}$ and $RNAP_{i+1}$ are initially moving so that Δ_{i-1} and Δ_i respectively increases and decreases with time. Hence $RNAP_{i-1}$ and $RNAP_{i+1}$ exert respectively an increasing force $F(i-1 \rightarrow i)$ and $F(i+1 \rightarrow i)$ onto $RNAP_i$, of the same magnitude

and acting in the same direction: $RNAP_{i-1}$ pulling and $RNAP_{i+1}$ pushing $RNAP_i$. Therefore the resulting torque acting on $RNAP_i$ is twice $\Gamma(2 \rightarrow 1)$, and the blocked polymerase will resume its motion as soon as $K + 2\Gamma(2 \rightarrow 1) > M$. This requires $M < 3K$. Following the same reasoning, three RNAPs will stop if M exceeds this threshold, but then the next two flanking RNAPs come into play, shifting the limiting torque to $5K$, and so on until all RNAPs come into play, up to the same overall stall torque NK .

5.3 Transcription Against Supercoiling

It is interesting to remark that, while the previous calculations assume the presence of a roadblock acting as an obstacle for the transcribing RNAPs, the same results apply in the case where the RNAP convoy proceeds toward a region of positive supercoiling, generated in particular by the convoy activity itself. A twisted region of positive supercoiling σ acts indeed on the incoming RNAP as an effective torque $M = k_B T L_{PT} \sigma / h$, to be compared to the effective stall torque NK . The same reasoning as before leads to the conclusion that, whereas a single RNAP will be stopped once the supercoiling is increased above a critical value $\sigma_1 = K h / (k_B T) L_{PT}$, a convoy containing N RNAPs may proceed against supercoiling up to a threshold $\sigma_N = NK h / (k_B T) L_{PT}$.

6 Conclusion and Further Remarks

A physical and often underrated feature of transcription is the reciprocal coupling between polymerase activity and DNA supercoiling: an active polymerase locally modifies DNA supercoiling, which in turn affects its motion and activity. A major prediction of our physical analysis is the existence of polymerase convoys, due to the torsional coupling between adjacent polymerases: any local desynchronization modifies DNA supercoiling, and the associated increase of local torsional energy generates an apparent force sufficiently strong to restore the initial distance between the polymerases and ensures the cohesion of the convoy. Moreover the ensuing collective behavior increases the value of the stall force required to stop polymerase activity. Such polymerase convoys should not be confused with possible polymerase complexes, as suggested for instance in [61]. In a convoy, there is no direct interaction (molecular contact) between the polymerases. The cohesiveness of the train is ensured only by torsional constraints.

Note that in contrast, polymerases Pol I (transcribing ribosomal RNAs) display a far higher density on DNA (tenfold variation between Pol II and Pol I densities) and are presumably in direct physical contact, as supported by recent experimental observation in yeast [64]. However, our proposed scenario would still apply: any disruption of a direct molecular contact and appearance of a DNA spacing between active adjacent polymerases would also be associated with the appearance of torsional constraints, tending to

restore the close proximity. In both cases, the distance between polymerases is determined by the activity of the promoter (controlling the loading rate of polymerases) and not by direct or indirect molecular interactions between the polymerases.

In principle, other processive motors as myosin or kinesin may be considered from the point of view developed here. In contrast to DNA, however, actin filaments or microtubules are very rigid and a molecular motor could not generate any significant twist constraint (DNA, actin filaments, and microtubules have persistence lengths of about 50 nm, 20 μ m, and 1 mm, respectively). Myosin on actin, and kinesin or dynein on microtubules are just translating, with no relative rotation of these motors with respect to their substrate. Accordingly synchronization of molecular motors on such rigid filaments relies on different mechanisms [65, 66].

We also draw attention on a major difference between our modeling of supercoiling-coupled trafficking and the basic model of out-of-equilibrium systems, *asymmetric simple exclusion process* (ASEP), which is currently used to describe trafficking onto biological filaments. In this model, the main effect arises from steric interactions between neighboring particles. The model was besides introduced to account for the motion of ribosomes onto mRNA [67]. It is important to notice that applying this model to polymerases onto DNA would be incorrect, although some examples can be found in theoretical literature. Indeed, there is a basic difference between polymerases and ribosomes. In transcription elongation, torsional constraints are responsible for the effective coupling between RNAPs, which acts at a distance. Steric interactions are therefore useless for the case of RNAP convoys. In contrast, torsional constraints do not affect the single-strand mRNA processed by ribosomes, that can consequently come into contact and interact sterically. While ASEP is a valid model for ribosome processivity, it is meaningless for polymerases.

Overall, our approach centered on forces and torques underlines the importance of modeling physical mechanisms to quantitatively interpret biological observations, and eventually understand the interplay between the physical properties of DNA and the action of specific biological actors centrally involved in genomic functions, such as transcription.

Acknowledgments

We acknowledge our team “Multiscale Modeling of Living Matter” at LPTMC, Thierry Forné, Christophe Lavelle, and Marc Nadal for stimulating discussions. This work was funded by the French Institut National du Cancer, grant INCa_5960, and the French Agence Nationale de la Recherche, grant ANR-13-BSV5-0010-03.

References

- Wang JC, Peck LJ, Becherer K (1983) DNA supercoiling and its effects on DNA structure and function. In: Cold Spring Harbor symposia on quantitative biology, vol 47. Cold Spring Harbor Laboratory Press, New York, pp 85–91
- Champoux JJ (2001) DNA topoisomerases: structure, function, and mechanism. *Annu Rev Biochem* 70(1):369–413
- Lavelle C (2008) DNA torsional stress propagates through chromatin fiber and participates in transcriptional regulation. *Nat Struct Mol Biol* 15(2):123–125
- Lavelle C (2014) Pack, unpack, bend, twist, pull, push: the physical side of gene expression. *Curr Opin Genet Dev* 25:74–84
- Travers A, Muskhelishvili G (2005a) DNA supercoiling—a global transcriptional regulator for enterobacterial growth? *Nat Rev Microbiol* 3(2):157–169
- Edelstein AD (2009) The effect of torque on RNA polymerase. University of California, Berkeley
- Bustamante C, Keller D, Oster G (2001) The physics of molecular motors. *Acc Chem Res* 34(6):412–420
- Bécavin C, Barbi M, Victor JM, Lesne A (2010) Transcription within condensed chromatin: steric hindrance facilitates elongation. *Biophys J* 98(5):824–833
- Lesne A, Bécavin C, Victor JM (2012) The condensed chromatin fiber: an allosteric chemo-mechanical machine for signal transduction and genome processing. *Phys Biol* 9(1):013001
- Svejstrup JQ (ed) (2013) Special issue on RNA polymerase II transcript elongation. *Biochim Biophys Acta* 1829:1–186
- Vicsek T (ed) (2001) Fluctuations and scaling in biology. Oxford University Press, New York
- Tantale K, Mueller F, Kozulic-Pirher A, Lesne A, Victor JM, Robert MC, Capozzi S, Cheia R, Bäcker V, Mateos-Langerak J, Darzacq X, Zimmer C, Basyuk E, Bertrand E (2016) A single-molecule view of transcription reveals convoys of RNA polymerases and multi-scale bursting. *Nat Commun* 7:12248
- White JH, Bauer WR (1986) Calculation of the twist and the writhe for representative models of DNA. *J Mol Biol* 189(2):329–341
- Kouzine F, Gupta A, Baranello L, Wojtowicz D, Ben-Aissa K, Liu J, Przytycka TM, Levens D (2013) Transcription-dependent dynamic supercoiling is a short-range genomic force. *Nat Struct Mol Biol* 20(3):396–403
- Dame RT (2005) The role of nucleoid-associated proteins in the organization and compaction of bacterial chromatin. *Mol Microbiol* 56(4):858–870
- Travers A, Muskhelishvili G (2007) A common topology for bacterial and eukaryotic transcription initiation? *EMBO Rep* 8(2):147–151
- Lim CJ, Kenney LJ, Yan J (2014) Single-molecule studies on the mechanical interplay between DNA supercoiling and H-NS DNA architectural properties. *Nucleic Acids Res* 42(13):8369–8378
- Travers A, Schneider R, Muskhelishvili G (2001) DNA supercoiling and transcription in *Escherichia coli*: the FIS connection. *Biochimie* 83(2):213–217
- Van Holde KE (2012) Chromatin. Springer, New York
- Worcel A, Strogatz S, Riley D (1981) Structure of chromatin and the linking number of DNA. *Proc Natl Acad Sci U S A* 78(3):1461–1465
- Barbi M, Mozziconacci J, Victor JM, Wong H, Lavelle C (2012) On the topology of chromatin fibres. *Interface Focus* 2(5):546–554
- Travers A, Muskhelishvili G (2005b) Bacterial chromatin. *Curr Opin Genet Dev* 15(5):507–514
- Gellert M, Mizuuchi K, O’Dea MH, Nash HA (1976) DNA gyrase: an enzyme that introduces superhelical turns into DNA. *Proc Natl Acad Sci U S A* 73(11):3872–3876
- Peter BJ, Arsuaga J, Breier AM, Khodursky AB, Brown PO, Cozzarelli NR (2004) Genomic transcriptional response to loss of chromosomal supercoiling in *Escherichia coli*. *Genome Biol* 5:R87
- Nadal M (2007) Reverse gyrase: an insight into the role of DNA-topoisomerases. *Biochimie* 89(4):447–455
- Naughton C, Avlonitis N, Corless S, Prendergast JG, Mati IK, Eijk PP, Cockcroft SL, Bradley M, Ylstra B, Gilbert N (2013) Transcription forms and remodels supercoiling domains unfolding large-scale chromatin structures. *Nat Struct Mol Biol* 20(3):387–395
- Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, Hu M, Liu JS, Ren B (2012) Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* 485(7398):376–380
- Nora EP, Lajoie BR, Schulz EG, Giorgetti L, Okamoto I, Servant N, Piolot T, van Berkum NL, Meisig J, Sedat J, Gribnau J, Barillot E, Blüthgen N, Dekker J, Heard E (2012) Spatial partitioning of the regulatory landscape of the

- X-inactivation center. *Nature* 485 (7398):381–385
29. Ea V, Baudement MO, Lesne A, Forné T (2015) Contribution of topological domains and loop formation to 3D chromatin organization. *Genes* 6(3):734–750
 30. Padinhateeri R, Marko JF (2011) Nucleosome positioning in a model of active chromatin remodeling enzymes. *Proc Natl Acad Sci U S A* 108(19):7799–7803
 31. Cortini R, Barbi M, Caré BR, Lavelle C, Lesne A, Mozziconacci J, Victor JM (2016) The physics of epigenetics. *Rev Modern Phys* 88:025002
 32. Aragon L, Martinez-Perez E, Merckenschlager M (2013) Condensin, cohesin and the control of chromatin states. *Curr Opin Genet Dev* 23 (2):204–211
 33. Hirano T (2014) Condensins and the evolution of torsion-mediated genome organization. *Trends Cell Biol* 24(12):727–733
 34. Terweij M, van Leeuwen F (2013) Histone exchange: sculpting the epigenome. *Front Life Sci* 7(1–2):63–79
 35. Lavelle C (2009) Forces and torques in the nucleus: chromatin under mechanical constraints. *Biochem Cell Biol* 87(1):307–322
 36. Rovinskiy N, Agbleke AA, Chesnokova O, Pang Z, Higgins NP (2012) Rates of gyrase supercoiling and transcription elongation control supercoil density in a bacterial chromosome. *PLoS Genet* 8(8):e1002845
 37. Matsumoto K, Hirose S (2004) Visualization of unconstrained negative supercoils of DNA on polytene chromosomes of *Drosophila*. *J Cell Sci* 117(17):3797–3805
 38. Liu LF, Wang JC (1987) Supercoiling of the DNA template during transcription. *Proc Natl Acad Sci U S A* 84(20):7024–7027
 39. Chong S, Chen C, Ge H, Xie XS (2014) Mechanism of transcriptional bursting in bacteria. *Cell* 158(2):314–326
 40. Bancaud A, Wagner G, e Silva NC, Lavelle C, Wong H, Mozziconacci J, Barbi M, Sivolob A, Le Cam E, Mouawad L, Viovy JL, Victor JM, Prunell A (2007) Nucleosome chiral transition under positive torsional stress in single chromatin fibers. *Mol Cell* 27(1):135–147
 41. Petesch SJ, Lis JT (2008) Rapid, transcription-independent loss of nucleosomes over a large chromatin domain at Hsp70 loci. *Cell* 134 (1):74–84
 42. Jeong KS, Ahn J, Khodursky AB (2004) Spatial patterns of transcriptional activity in the chromosome of *Escherichia coli*. *Genome Biol* 5 (11):R86
 43. Blot N, Mavathur R, Geertz M, Travers A, Muskhelishvili G (2006) Homeostatic regulation of supercoiling sensitivity coordinates transcription of the bacterial genome. *EMBO Rep* 7(7):710–715
 44. Sarkar A, Marko JF (2001) Removal of DNA-bound proteins by DNA twisting. *Phys Rev E* 64(6):061909
 45. Bustamante C (2008) In singulo biochemistry: when less is more. *Annu Rev Biochem* 77:45–50
 46. Abbondanzieri EA, Greenleaf WJ, Shaevitz JW, Landick R, Block SM (2005) Direct observation of base-pair stepping by RNA polymerase. *Nature* 438(7067):460–465
 47. Larson MH, Landick R, Block SM (2011) Single-molecule studies of RNA polymerase: one singular sensation, every little step it takes. *Mol Cell* 41(3):249–262
 48. Greenleaf WJ, Woodside MT, Block SM (2007) High-resolution, single-molecule measurements of biomolecular motion. *Annu Rev Biophys Biomol Struct* 36:171
 49. Adelman K, La Porta A, Santangelo TJ, Lis JT, Roberts JW, Wang MD (2002) Single molecule analysis of RNA polymerase elongation reveals uniform kinetic behavior. *Proc Natl Acad Sci U S A* 99(21):13538–13543
 50. Zhou J, Schweikhard V, Block SM (2013) Single-molecule studies of RNAPII elongation. *Biochim Biophys Acta* 1829(1):29–38
 51. Herbert KM, Greenleaf WJ, Block SM (2008) Single-molecule studies of RNA polymerase: motoring along. *Annu Rev Biochem* 77:149–176
 52. Thomen P, Lopez PJ, Bockelmann U, Guillerez J, Dreyfus M, Heslot F (2008) T7 RNA polymerase studied by force measurements varying cofactor concentration. *Biophys J* 95(5):2423–2433
 53. Wuite GJ, Smith SB, Young M, Keller D, Bustamante C (2000) Single-molecule studies of the effect of template tension on T7 DNA polymerase activity. *Nature* 404 (6773):103–106
 54. Danko CG, Hah N, Luo X, Martins AL, Core L, Lis JT, Siepel A, Kraus WL (2013) Signaling pathways differentially affect RNA polymerase II initiation, pausing, and elongation rate in cells. *Mol Cell* 50(2):212–222
 55. Wang MD, Schnitzer MJ, Yin H, Landick R, Gelles J, Block SM (1998) Force and velocity measured for single molecules of RNA polymerase. *Science* 282:902–907
 56. Ma J, Bai L, Wang MD (2013) Transcription under torsion. *Science* 340(6140):1580–1583

57. Dangkulwanich M, Ishibashi T, Bintu L, Bustamante C (2014) Molecular mechanisms of transcription through single-molecule experiments. *Chem Rev* 114(6):3203–3223
58. Harada Y, Ohara O, Takatsuki A, Itoh H, Shimamoto N, Kinoshita K (2001) Direct observation of DNA rotation during transcription by *Escherichia coli* RNA polymerase. *Nature* 409(6816):113–115
59. Deufel C, Forth S, Simmons CR, Dejongha S, Wang MD (2007) Nanofabricated quartz cylinders for angular trapping: DNA supercoiling torque detection. *Nat Methods* 4(3):223–225
60. Miller OJ, McKnight S (1979) Post-replicative nonribosomal transcription units in *D. melanogaster* embryos. *Cell* 17(3):551–563
61. Wong H, Winn PJ, Mozziconacci J (2009) A molecular model of chromatin organisation and transcription: how a multi-RNA polymerase II machine transcribes and remodels the β -globin locus during development. *BioEssays* 31:1357–1366
62. Epshtein V, Toulmé F, Rahmouni AR, Borukhov S, Nudler E (2003) Transcription through the roadblocks: the role of RNA polymerase cooperation. *EMBO J* 22(18):4719–4727
63. Wagner P (2011) A time-discrete harmonic oscillator model of human car-following. *Eur Phys J B* 84(4):713–718
64. Albert B, Léger-Silvestre I, Normand C, Ostermaier MK, Pérez-Fernández J, Panov KI, Zomerdijk JCBM, Schultz P, Gadal O (2011) RNA polymerase I-specific subunits promote polymerase clustering to enhance the rRNA gene transcription cycle. *J Cell Biol* 192:277–293
65. Berger F, Keller C, Klumpp S, Lipowsky R (2012) Distinct transport regimes for two elastically coupled molecular motors. *Phys Rev Lett* 108(20):208101
66. Kohler F, Rohrbach A (2015) Synchronization of elastically coupled processive molecular motors and regulation of cargo transport. *Phys Rev E* 91(1):012701
67. MacDonald CT, Gibbs JH, Pipkin AC (1968) Kinetics of biopolymerization on nucleic acid templates. *Biopolymers* 6(1):1–25