

A looping-based model for quenching repression

Y. Pollak, S. Goldberg, and R. Amit

November 5, 2016

Abstract

We model the regulatory role of proteins bound to looped DNA using a simulation in which dsDNA is represented as a self-avoiding chain, and proteins as spherical protrusions. We simulate long self-avoiding chains using a sequential importance sampling Monte-Carlo algorithm, and compute the probabilities for chain looping with and without a protrusion. We find that a protrusion near one of the chain's termini reduces the probability of looping, even for chains much longer than the protrusion-chain-terminus distance. This effect increases with protrusion size, and decreases with protrusion-terminus distance. The reduced probability of looping can be explained via an eclipse-like model, which provides a novel inhibitory mechanism. We test the eclipse model on two possible transcription-factor occupancy states of the *eve* 3/7 enhancer, and show that it provides a possible explanation for the experimentally-observed *eve* stripe 3 and 7 expression patterns. **Keywords:** DNA looping, self-avoiding wormlike chain, enhancers, quenching repression. The authors declare no conflict of interests

INTRODUCTION

Polymer looping is a phenomenon that is critical for the understanding of many chemical and biological processes. In particular, DNA looping has been implicated in transcriptional regulation across many organisms, and as a result plays a crucial role in how organisms develop and respond to their environments. While DNA looping has been studied extensively over the last several decades both experimentally (1–5) and theoretically (6–11), many aspects of looping-based transcriptional regulation remain poorly understood.

In vivo, the simplest looping-based regulatory architecture is comprised of a protein that interacts with a distal site via DNA looping. Such simple architectures can be found in bacteria, where, for example, σ^{54} (σ^N) promoters are activated via such a mechanism (12–14). In eukaryotes, DNA-looping-based regulation is associated with the interaction between the core promoter and distal regulatory regions called enhancers. These ~500 bp enhancers typically contain clusters of transcription factor (TF) binding sites, and may be located from 1 Kbp to several Mbps away from their regulated promoters. Detailed studies of enhancers from several organisms (15) have revealed that TFs can upregulate, inhibit, or both upregulate and inhibit gene expression via a variety of mechanisms. For example, repressors like *D. melanogaster* Giant, Knirps, Krüppel, and Snail inhibit expression either by partial overlap of their binding sites with that of an activator, or via a short-range repression mechanism termed “quenching”, whereby TFs positioned several tens of bps either upstream or downstream from the nearest activator inhibit gene expression (16–19).

In quenching, a bound protein inhibits gene expression, apparently without any direct interaction between the protein and either the promoter or the nearest-bound activator. In the prevailing model, quenching is assumed to be a result of histone deacetylation, which is facilitated by the formation of a ~450 kDa DNA-bound chromatin remodelling complex made of a DNA binding protein such as Knirps, C-terminus binding protein (CtBP) (20–22), and the histone deacetylases (HDACs) Rpd3 (23) and Sin3 (24). However, this model falls short of providing a full description of quenching. In particular, short-range repressors such as Knirps have been shown to retain a regulatory function even without the CtBP binding domain (22, 25, 26), and tests for HDAC activity using either Rpd3 knock-outs (27) or the HDAC inhibitor trichostatin A failed to alleviate the observed repression effects (26). Consequently, the mechanistic underpinnings of quenching are still poorly understood.

In this report, we show that DNA looping can provide a mechanistic model for quenching repression. We use a modified worm-like chain model that takes excluded-volume considerations into account (28–30) to show that volume effects can substantially affect loop formation for chains that are arbitrarily long. Using this approach, we show that the excluded volumes of DNA and a protein bound within ~1 Kuhn length (~300 bp) either upstream or downstream of one of the loop termini can block the “line-of-sight” of the other terminus, generating an eclipse-like effect, which leads to a reduction in the probability of looping.

THEORY

DNA in the absence of bound proteins

We model the DNA as a discrete semi-flexible chain made of individual links of length l . A chain is described by the locations \mathbf{r}_i of its link ends, and a local coordinate system defined by three orthonormal vectors $\hat{u}_i, \hat{v}_i, \hat{t}_i$ at each link, where \hat{t}_i points along the direction of the i th link. We use the following notations for a specific chain configuration: θ_i, ϕ_i are the zenith and azimuthal angles of \hat{t}_i in local spherical coordinates of link $i - 1$, respectively. $\{\theta, \phi\}_n \equiv \{\theta_1, \dots, \theta_n, \phi_1, \dots, \phi_n\}$ denotes all the angles until link n . Joint i is the end-point of link i and joint 0 is the beginning terminus of the chain. w is the effective cross-section of the polymer. Each chain joint is engulfed by a “hard-wall” spherical shell of diameter w . The total elastic energy associated with the polymer chain can be written as follows (28):

$$E(\{\theta, \phi\}_N) = \sum_{i=2}^N E^{\text{bend}}(\theta_i, \phi_i) + \sum_{i=2}^N E_i^{\text{hw}}(\{\theta, \phi\}_i), \quad (1)$$

where the elastic contribution to the energy is given by:

$$\beta E^{\text{bend}}(\theta_i, \phi_i) = a(1 - \cos \theta_i), \quad (2)$$

a is the bending constant of the polymer chain, and we have assumed azimuthal symmetry. The hard-wall contribution is given by:

$$\beta E_i^{\text{hw}}(\{\theta, \phi\}_i) = \begin{cases} \infty & i \text{ overlaps with one or more} \\ & \text{joints } 0 \dots (i - \Delta i) \\ 0 & \text{otherwise} \end{cases}. \quad (3)$$

Here $\beta = (k_b T)^{-1}$, k_b is the Boltzmann factor and T is the temperature. In case $l \geq w$, $\Delta i = 1$. In case $l < w$, two or more consecutive spheres overlap and Δi ensures that links j and k interact only if $|j - k| \geq \Delta i \geq \frac{w}{l}$. For simplicity, we disregard the twist degree of freedom in this work (see discussion in Supplementary Information Section 1.1).

DNA in the presence of bound proteins

We model the bound proteins as hard-wall spherical protrusions positioned adjacent to the polymer chain, with radius R_o representative of the protein's volume (Fig. 1). Since we neglect torsion effects in our present model, there is no intrinsic rotation of \hat{u}_i around the polymer axis. Thus, we define the orientation of the bound proteins by rotating \hat{u}_i around \hat{t}_i . Therefore, the center of a protrusion bound to chain link k and rotated around the chain axis by an angle of γ_k is given by:

$$\mathbf{r}_{\text{object}} = \mathbf{r}_k + \left(\frac{w}{2} + R_o \right) \mathbf{R}(\gamma_k, \hat{t}_k) \hat{u}_k, \quad (4)$$

where $\mathbf{R}(\gamma_k, \hat{t}_k)$ is a rotation matrix by an angle γ_k around \hat{t}_k (see Fig. 1). Addition of a protrusion at link k slightly alters Eq. (3), requiring to test whether joint i overlaps with one or more joints $0 \dots (i - \Delta i)$ and with the protrusion at $\mathbf{r}_{\text{object}}$, if $i > k$.

Definition of looping probability ratio

We consider a chain looped if \mathbf{r}_N is confined to a volume $\delta \mathbf{r}$ around \mathbf{r}_0 (see Fig. 1), defined by:

1. $d_{\min} \leq |\mathbf{r}_N - \mathbf{r}_0| \leq d_{\min} + \varepsilon$.
2. $(\mathbf{r}_N - \mathbf{r}_0)$ is collinear with \hat{u}_0 within $\delta\omega'$.

For a specific choice of $\delta \mathbf{r}$, we define the probability of a polymer chain of length L to form a loop as (28):

$$P_{\text{looped}}(L) \equiv \int_{\delta \mathbf{r}} \mathbf{C}(\mathbf{r}) \, d\mathbf{r}, \quad (5)$$

where $\mathbf{C}(\mathbf{r})$ is the probability density function of the end-to-end vector $\mathbf{r} \equiv \mathbf{r}_N - \mathbf{r}_0$. In this work we study the effect of a bound object on the probability of the polymer to form a loop, with the looping criteria defined above. We quantify this effect by the looping probability ratio:

$$F(L, \text{object}) \equiv \frac{P_{\text{looped}}^{\text{object}}(L)}{P_{\text{looped}}^{\text{baseline}}(L)}, \quad (6)$$

where $P_{\text{looped}}^{\text{baseline}}(L)$ is the looping probability of the bare polymer chain and $P_{\text{looped}}^{\text{object}}(L)$ is the looping probability of the polymer chain with a protrusion bound to it.

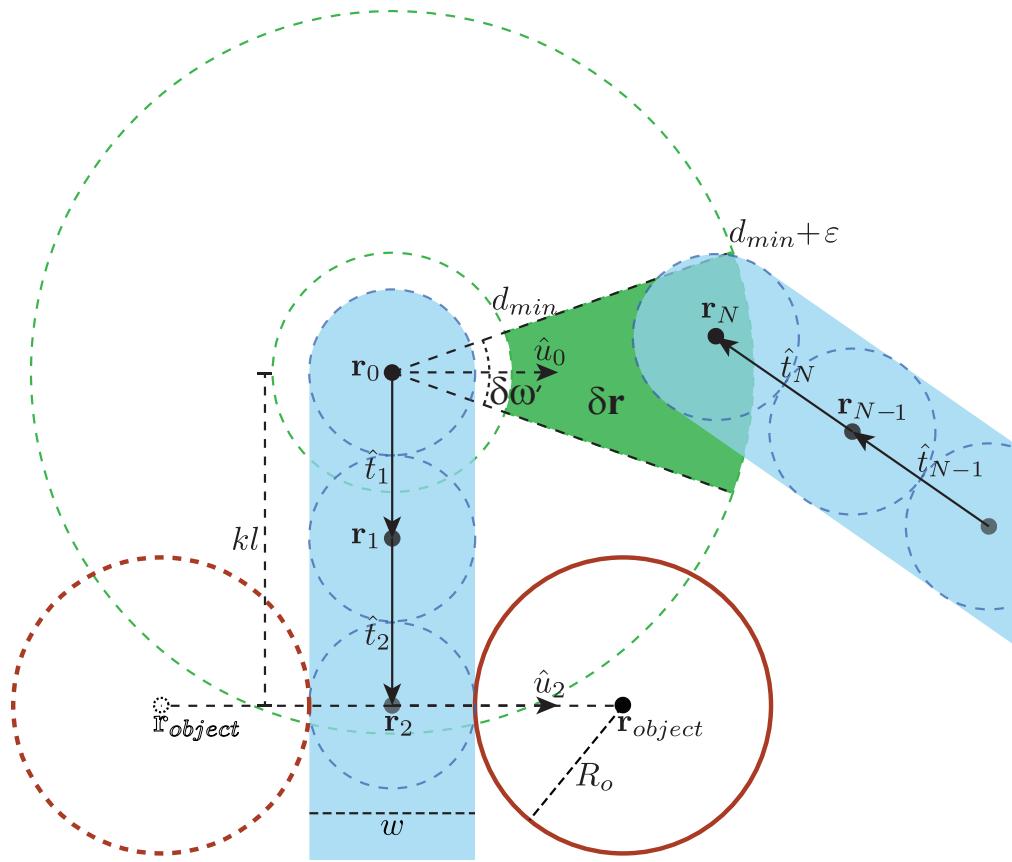


Figure 1: Loop and protrusion geometry. The chain (shaded in cyan) is modeled by spheres (blue dashed circles), here with link length l equal to diameter w . Looping volume δr is represented by a green wedge. The protrusion is positioned on the second link ($k = 2$). In-phase-with- δr ($\gamma_2 = 0^\circ$) and out-of-phase ($\gamma_2 = 180^\circ$) positions are illustrated by the solid and dashed red circles, respectively.

MATERIALS AND METHODS

Linear polymers have been studied via simulations using a variety of methods (31, 32). To model DNA fully it is necessary to take into account the experimentally observed non-entangled chromosomal DNA structure (33), and to include details regarding bound proteins. Recently, polymer rings have been studied as a model system for topologically constrained polymer melts, such as chromosomal DNA (34, 35). In this work, we address looping of a linear polymer, neglecting the additional topological constraints, but including bound protrusions. We chose to simulate a linear DNA chain with a bound protein using a sequential importance sampling Monte-Carlo approach that we used previously to simulate the configurational space of bare DNA (28). To adapt our algorithm to the case of protein-bound DNA, we take into account not only the growing chain but also the location of the protrusion (see Eq. (4)). During chain generation, upon reaching link k , the simulation adds a hard-wall spherical protrusion with radius R_o at the location r_{object} . If the protrusion overlaps any of the previously-generated chain links or protrusions, the chain is discarded. After generating the configurational ensemble, we identify the subset of “looped” chains. We provide the

essential details of the simulation in this Section. Additional details can be found in the extended Materials and Methods online.

Sequential importance sampling

Our Monte-Carlo algorithm is an off-lattice sequential importance sampling algorithm adapted from a method developed by Rosenbluth and Rosenbluth (36) to generate self-avoiding walks on a lattice (as described in (28)). We generate faithful statistical ensembles consisting of $N_c \approx 10^9$ self-avoiding DNA chains with bound proteins modeled as hard-wall spheres. A chain j in the ensemble is assigned with a Rosenbluth factor w_j in order to negate the biasing effect introduced by sampling of self-avoiding chains. The partition function of the ensemble takes the form of

$$Z = \sum_{j=1}^{N_c} w_j, \quad (7)$$

and any physical observable can then be computed from the generated ensemble by

$$\langle f \rangle = \frac{\sum_{j=1}^{N_c} f(\mathbf{x}_j) w_j}{Z}, \quad (8)$$

where \mathbf{x}_j are the generalized coordinates of configuration j . The probability of looping is computed by

$$P_{\text{looped}} = \langle f_{\text{looped}} \rangle, \quad (9)$$

where

$$f_{\text{looped}}(\mathbf{x}_j) = \begin{cases} 1 & \text{configuration } j \text{ is looped} \\ 0 & \text{otherwise} \end{cases}. \quad (10)$$

Simulation parameters

In our simulations, $d_{\min} = w$, $\varepsilon = 2w$ and $\delta\omega' = 2\pi \times 0.1$, unless stated otherwise. Changing these parameters did not alter the results significantly, and the relatively large ε chosen minimized noise. We simulated the DNA chain with diameter $w = 4.6$ nm and Kuhn length (7) $b = 106$ nm, where a is computed using (30):

$$\frac{b}{l} = \frac{a - 1 + a \coth a}{a + 1 - a \coth a}. \quad (11)$$

Chains were simulated in two stages as a compromise between resolution and running time. The first N_1 links of the chain were simulated with link length $l_1 = 0.34$ nm, corresponding to the length of a base-pair in dsDNA. We used $\Delta i \approx \frac{4}{3} \frac{w}{l}$ for this stage (37). All bound objects were positioned at links $k < N_1$. The remaining N_2 links of the chain were simulated with link length $l_2 = w$. We denote the overall length of the chain by $L = N_1 l_1 + N_2 l_2$, and the distance along the chain of an object binding link k from the chain origin by $K = k l_1$. For further details see Extended Materials and Methods.

Natural system simulation

We modeled the structure of the 3/7 enhancer in *D. melanogaster* based on (38). The dStat (86 kDa), Zld (146 kDa), bare Knirps (46 kDa), Knirps bound to CtBP dimers (130 kDa), and a full putative 450 kDa complex as reported by (23) were modeled as hard-wall spheres with sizes corresponding to globular proteins with radii of 3.04 nm, 3.75 nm, 2.38 nm, 3.59 nm and 5.83 nm, respectively, based on the work by (39). We used $2\pi/10.5$ as the DNA native twist. The chain was assumed to be rigid with respect to the twist degree of freedom. The RNA polymerase-and-cofactors complex was not modeled. Instead, for a loop to form, the chain terminus distant from the enhancer was required to be in close proximity to one of the three activators. A chain was considered looped with respect to a specific activator if it fulfilled the following conditions:

1. $d_{\min} \leq |\mathbf{r}_N - \mathbf{r}_{\text{activator}}| \leq d_{\min} + \varepsilon$.
2. $(\mathbf{r}_N - \mathbf{r}_{\text{activator}})$ is collinear with $\hat{u}_{\text{activator}}$ within $\delta\omega'$.

Here $\mathbf{r}_{\text{activator}}$ is the center of the activator sphere ($\mathbf{r}_{\text{object}}$ in Fig. 1) and $\hat{u}_{\text{activator}}$ is the vector pointing from the chain axis to the center of the activator (\hat{u}_2 in Fig. 1). We simulated DNA chains of length 3900 bp, which corresponds to the native *eve* 3/7 enhancer geometry and its separation from the *eve* promoter. The total probability of looping was computed as the sum of the looping probabilities with respect to the individual activators. In this simulation we used $d_{\min} = R_{\text{chain}} + R_{\text{activator}}$ and $\varepsilon = 3$ nm. The simulation was relatively insensitive to ε . For $\delta\omega'$, we chose several values: $\delta\omega' = 2\pi \times 0.1$ corresponding to a narrow cone directed away from the chain, $\delta\omega' = 2\pi$ corresponding to a full spherical shell around the activator sphere, and $\delta\omega' = -2\pi \times 0.5$ corresponding to a relatively wide cone directed towards the chain.

Computation of F_∞ and its error estimation

See Extended Materials and Methods.

RESULTS

Long-range down-regulatory effect

To model the quenching effect of a bound repressor on DNA looping, we generated configurational ensembles for DNA with a spherical protrusion of size $R_o = 9.2$ nm or $R_o = 18.4$ nm located a distance of $K = 95$ bp or $K = 135$ bp from the chain origin along the chain, oriented either in the same direction as the looping volume $\delta\mathbf{r}$ or 180° from it (see Fig. 1). We plot $F(L)$ for the various configurations of R_o and K in Fig. 2A. The data show that in the elastic regime ($L \leq b$), protrusions bound in-phase with $\delta\mathbf{r}$ (solid lines, \uparrow) strongly reduce the looping probability relative to that of the bare DNA, while protrusions positioned out-of-phase to $\delta\mathbf{r}$ (dashed lines, \downarrow) increase the looping probability, as we showed previously (37). However, in the entropic regime ($L \gg b$), all chains converge to values of $F \xrightarrow{L \gg b} F_\infty \leq 1$. For protrusions that are bound in-phase (i.e. $\gamma_k = 0$), F_∞ is distinctly smaller than one, and strongly depends on both the distance to the nearest terminus and

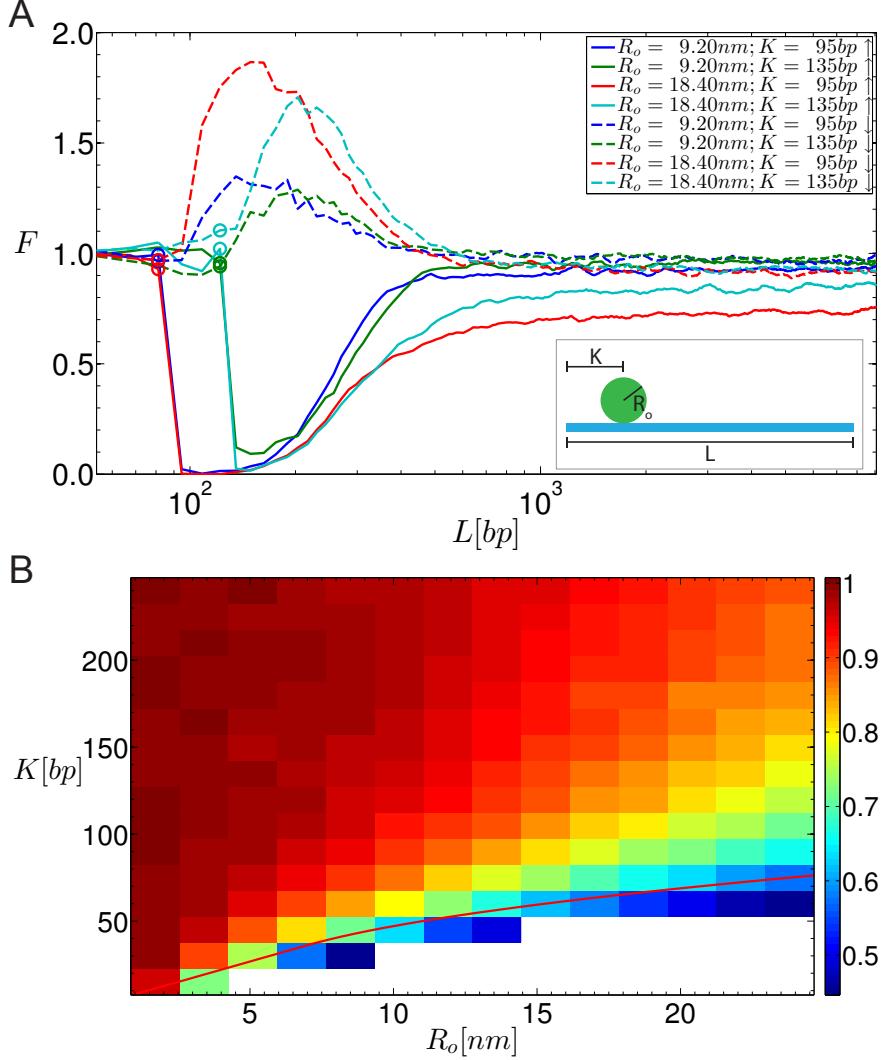


Figure 2: Simulating looping probability ratio F . A) F plotted as a function of chain length L for several values of R_o and K , and $\gamma_k = 0^\circ$ (solid lines) or 180° (dashed lines). The locations of the protrusions are denoted by circles on the corresponding curves. B) F_∞ plotted as a function of R_o and K , for $\gamma_k = 0$. F_∞ is calculated numerically as the average of $F(L)$ over the range of L values where $F(L) \approx \text{const}$. The solid red curve is a visual aid: if the segment of the chain between the origin and the protrusion location was straight, points on the red curve would result in the protrusion touching the looping volume δr .

the size of the protrusion. Conversely, for protrusions bound out-of-phase ($\gamma_k = 180^\circ$), F_∞ is only slightly smaller than 1, with weak dependence on both protrusion size and position.

To further explore the extent of the quenching effect in the entropic or long-chain-length regime, we plot in Fig. 2B the value of F_∞ as a function of a wide-range of R_o and K , for $\gamma_k = 0$. The heatmap shows both a non-linear decrease in F_∞ as a function of protrusion size and a non-linear increase as a function of protrusion distance from the chain origin. The red line in the figure demarcates the closest possible location at which a protrusion can be bound without physically penetrating part of the looping volume δr . Thus, for volumes and protrusion positions that fall

below this line, a second excluded volume effect contributes to the reduction in the probability of looping, leading to a sharp increase in the overall effect. Together, the panels in Fig. 2 show that a sufficiently-large protrusion can substantially reduce the probability of looping, independent of loop length, provided that its binding site is within a small distance from either of the loop termini.

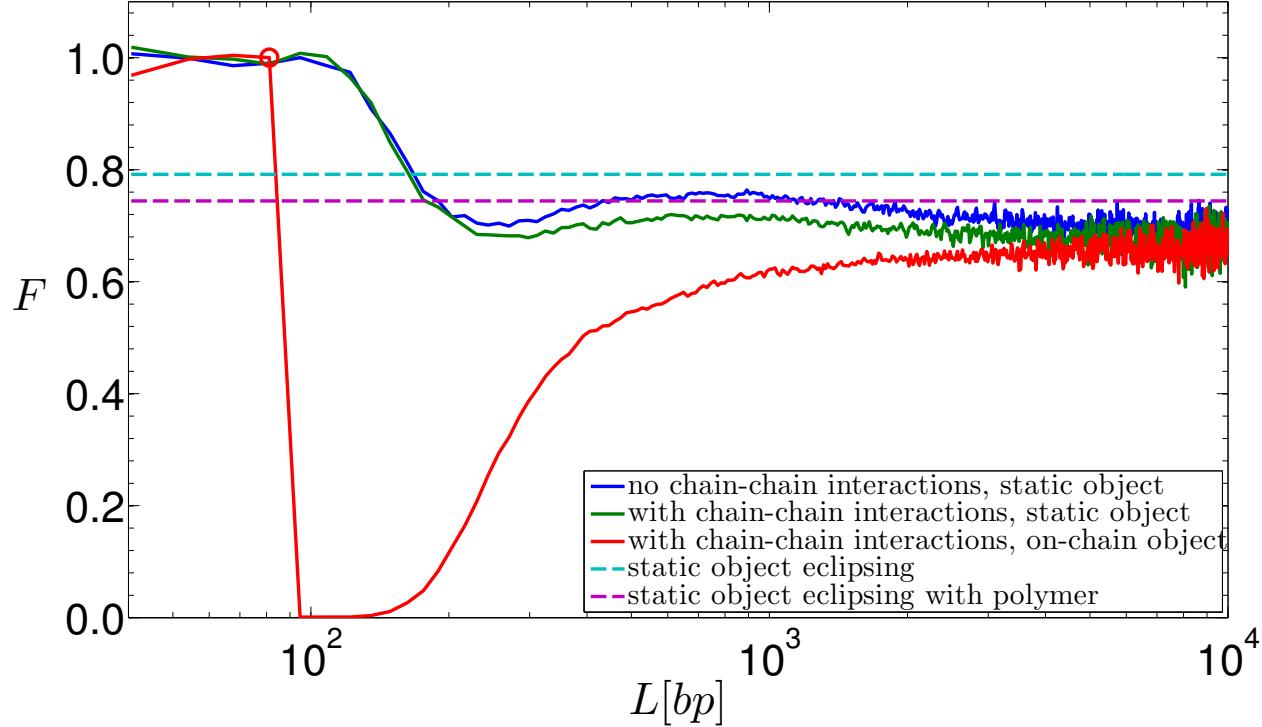


Figure 3: Simplified eclipse models. Simulation results for a chain without chain-chain interactions and a static object (solid blue line) are compared to estimates for F_∞ of the “rod” (dashed cyan line) and “terminating-segments” (dashed magenta line) models, also for a chain without chain-chain interactions. The static object is located at $d\hat{u}_0$, where $d = 41.45$ nm, and $R_o = 23$ nm. For comparison, we plot simulation results for a chain with chain-chain interactions and the same static object parameters (solid green line), and a chain with chain-chain interaction and an on-chain object (solid red line). The on-chain object is located at $K = 94$ bp (denoted by a circle).

“Eclipsing” approximation

To understand the long-range, length-independent effect shown in Fig. 2, we examine the chains’ terminating segments of length $T \ll L$. If $T \sim \max_{\mathbf{r} \in \delta\mathbf{r}} |\mathbf{r} - \mathbf{r}_{\text{object}}| \ll b$, these segments resemble stiff rods, and the object obstructs the line-of-sight of one chain terminus from the other. This eclipse-like phenomenon is manifested by a reduction in the number of polymer chains that are able to reach $\delta\mathbf{r}$. This, in turn, results in a smaller P_{looped} as compared with the case in which no protrusion is present. In the entropic regime and in the absence of protrusions, the generated “rods” approach the looping volume $\delta\mathbf{r}$ from all directions that are unobscured by the volume of the polymer in a homogeneous fashion (7). Due to this isotropy in the distribution of the chain

termini orientations within $\delta\mathbf{r}$, the reduction in P_{looped} can be approximated by the solid angle that the eclipsing object subtends at $\delta\mathbf{r}$. Consequently, F (Eq. (6)) can be approximated for this “rod model” by:

$$\begin{aligned} F_\infty(\tilde{\mathbf{r}}', R_o) &= \frac{P_{\text{looped}}^{\text{object}}(L, \{\tilde{\mathbf{r}}', R_o\})}{P_{\text{looped}}^{\text{baseline}}(L)} \Big|_{L \gg b} \\ &\approx \frac{4\pi\delta\mathbf{r} - \mathcal{I}_{\text{chain}} - \mathcal{I}_{\text{object}(\tilde{\mathbf{r}}', R_o)} + \mathcal{I}_{\text{chain}\cap\text{object}(\tilde{\mathbf{r}}', R_o)}}{4\pi\delta\mathbf{r} - \mathcal{I}_{\text{chain}}} \\ &= 1 - \frac{\mathcal{I}_{\text{object}(\tilde{\mathbf{r}}', R_o)}}{4\pi\delta\mathbf{r} - \mathcal{I}_{\text{chain}}} + \frac{\mathcal{I}_{\text{chain}\cap\text{object}(\tilde{\mathbf{r}}', R_o)}}{4\pi\delta\mathbf{r} - \mathcal{I}_{\text{chain}}}, \end{aligned} \quad (12)$$

where R_o is the radius of the spherical object, $\tilde{\mathbf{r}}'$ is the location of the object, which could be located statically at point \mathbf{r}' (in which case $\tilde{\mathbf{r}}' \equiv \mathbf{r}'$), or located on the chain a distance K from the chain origin (in which case we use the terminology $\tilde{\mathbf{r}}' \equiv \tilde{\mathbf{r}}'(K)$ figuratively to specify the progression of the protrusion along the chain). $\mathcal{I}_{\text{chain}} \equiv \int_{\delta\mathbf{r}} \Omega_{\text{chain}}(\mathbf{r}) d^3\mathbf{r}$, where $\Omega_{\text{chain}}(\mathbf{r})$ is the solid angle subtended at \mathbf{r} by the polymer chain links. $\mathcal{I}_{\text{object}(\tilde{\mathbf{r}}', R_o)} \equiv \int_{\delta\mathbf{r}} \Omega_{\text{object}(\tilde{\mathbf{r}}', R_o)}(\mathbf{r}) d^3\mathbf{r}$, where $\Omega_{\text{object}(\tilde{\mathbf{r}}', R_o)}(\mathbf{r})$ is the solid angle subtended at \mathbf{r} by the object, and $\mathcal{I}_{\text{chain}\cap\text{object}(\tilde{\mathbf{r}}', R_o)}$ corresponds to the solid angle contained in both $\mathcal{I}_{\text{chain}}$ and $\mathcal{I}_{\text{object}}$.

In order to test the eclipsing hypothesis, we first computed F_∞ for the case of an object statically positioned at an off-chain location $\mathbf{r}' = d\hat{\mathbf{u}}_0$, and without chain-chain interactions. In this simplified case, F_∞ in Eq. (12) can be approximated by the following eclipsing expression:

$$\begin{aligned} F_\infty(\mathbf{r}', R_o) &\approx 1 - \frac{\mathcal{I}_{\text{object}(\mathbf{r}', R_o)}}{4\pi\delta\mathbf{r}} \\ &= 1 - \frac{\int_{\delta\mathbf{r}} 2\pi \left[1 - \sqrt{1 - \left(\frac{R_o + w/2}{|\mathbf{r} - \mathbf{r}'|} \right)^2} \right] d^3\mathbf{r}}{4\pi\delta\mathbf{r}} \equiv F_\infty^{\text{static}}, \end{aligned} \quad (13)$$

where we substituted

$$\frac{\Omega_{\text{object}}}{2\pi} = \int_0^1 d\cos\theta = 1 - \sqrt{1 - \frac{(R_o + w/2)^2}{|\mathbf{r} - \mathbf{r}'|^2}}. \quad (14)$$

In Fig. 3, we compare the value computed from Eq. (13) (dashed cyan line) to $F(L)$ computed by our sequential importance sampling algorithm for the same conditions (solid blue line). The data show that the eclipsing approximation F_∞^{static} overestimates F_∞ . We reasoned that the main cause for this estimation error is that Eq. (13) disregards the flexible polymer nature of the chain. We ran an additional Monte-Carlo simulation to quantify the correction resulting from polymer flexibility. Here, we generated pairs consisting of an end-terminus point in $\delta\mathbf{r}$ and a direction vector of the terminal link, both distributed uniformly. Short polymer chains of length T originating at the chosen points were grown with their first links oriented in the chosen directions. These chains can be thought of as the terminating segments of long chains that have a uniform distribution of their end-termini in $\delta\mathbf{r}$. We found that the probability of a flexible-polymer chain to overlap the object increased relative to the probability within the “rod model”, resulting in a decrease in the

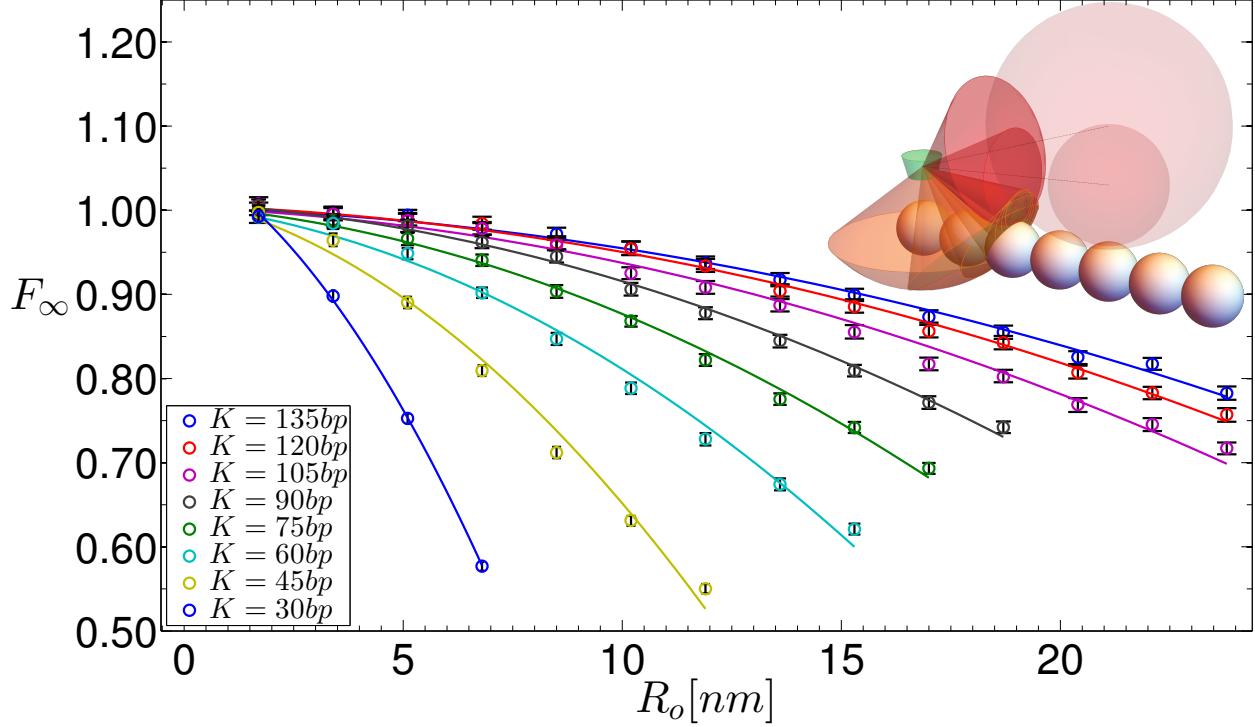


Figure 4: Dependence of F_∞ on protrusion size R_o . Simulation data (circles) are fit (solid curves) by $f_K(R_o)$ (Eq. 15). Data points are the mean of the last 480 points of the simulated F . Error bars are ± 1.96 times the standard error of these points. Inset illustrates a doubling of R_o . $\delta\mathbf{r}$ is shown by the green volume. Chain links are shown by white spheres. Protrusions are shown by transparent red spheres. Solid angles subtended by the chain links and protrusion are shown by orange and red cones, respectively. Ω_{chain} is the area on the unit sphere around the center of $\delta\mathbf{r}$ intersecting the orange cones. $\Omega_{\text{object}(\tilde{\mathbf{r}}', R_o)}$ is the area on the unit sphere intersecting the red cone. $\Omega_{\text{chain} \cap \text{object}(\tilde{\mathbf{r}}', R_o)}$ is the area on the unit sphere intersecting both orange and red cones.

probability of the chain to form a loop (magenta dashed line in Fig. 3). Using this ‘‘terminating-segments’’ correction, the discrepancy between F_∞ from the simulation and F_∞^{static} from Eq. 13 is partially accounted for. We attribute the additional reduction in the simulated F_∞ to interactions between the object and the remaining $L - T$ length of the chain.

In Fig. 4 we plot F_∞ as a function of an on-chain object of radius R_o , for several values of K . To compare the results of the numerical simulation to the full eclipsing model (Eq. 12), we first note that when K is kept constant, $\mathcal{I}_{\text{chain} \cap \text{object}(\tilde{\mathbf{r}}', R_o)}|_{K=\text{const}} \approx \text{const}$, as can be seen from the inset in Fig. 4: the overlap between Ω_{chain} (orange cones) and Ω_{object} (red cones) changes only slightly when the object grows by a factor of two. Furthermore, $|\tilde{\mathbf{r}}'(K) - \mathbf{r}||_{K=\text{const}}$ is approximately independent of R_o if $(R_o + w/2) \ll |\tilde{\mathbf{r}}'(K) - \mathbf{r}|$ for all $\mathbf{r} \in \delta\mathbf{r}$. Thus, the dependence of F_∞ on the

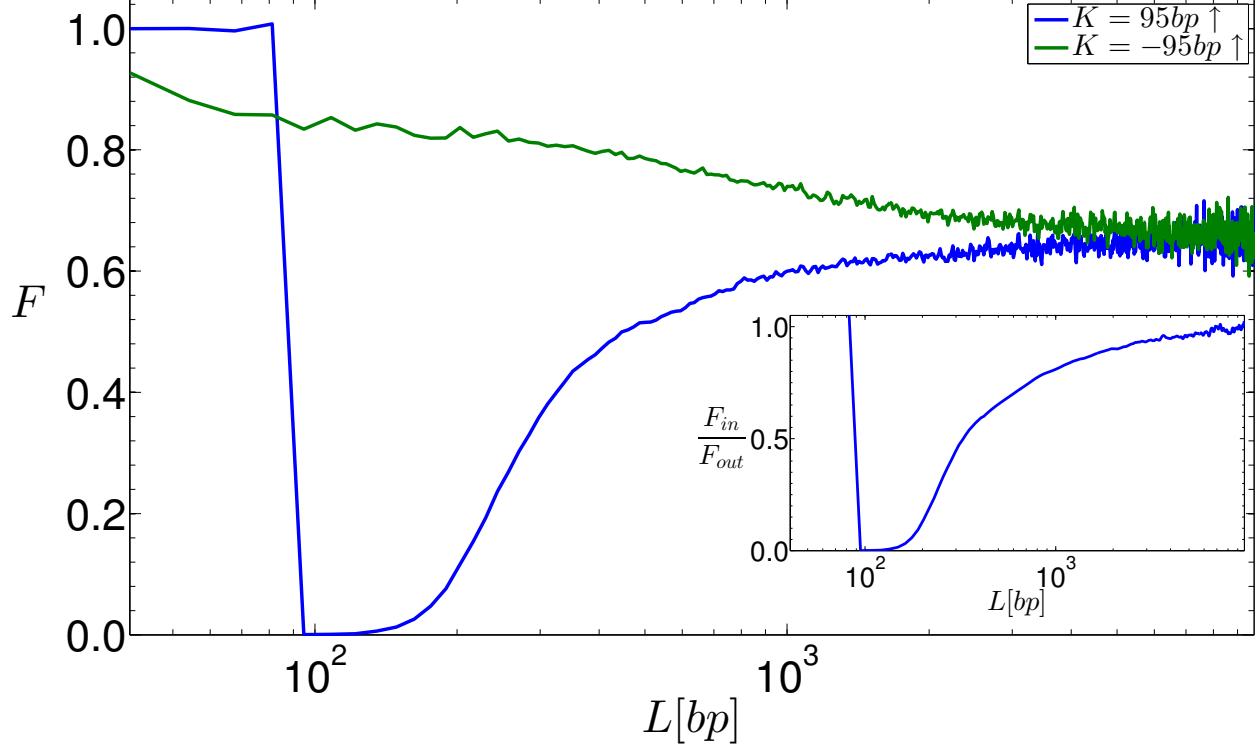


Figure 5: Dependence of $F(L)$ on location of the protrusion, either inside (K) or outside ($-K$) the chain segment between link 0 and the terminating link. $R_o = 23$ nm and $K = \pm 95$ bp. Inset: the ratio $F(L, K) / F(L, -K)$.

radius R_o of an on-chain object can be derived from Eq. (12):

$$\begin{aligned}
 F_\infty(\tilde{\mathbf{r}}', R_o) |_{K=const} &\approx \\
 &\approx 1 - \frac{\mathcal{I}_{\text{object}}(\tilde{\mathbf{r}}', R_o)}{4\pi\delta\mathbf{r} - \mathcal{I}_{\text{chain}}} + \frac{\mathcal{I}_{\text{chain}\cap\text{object}}(\tilde{\mathbf{r}}', R_o)}{4\pi\delta\mathbf{r} - \mathcal{I}_{\text{chain}}} \\
 &\approx 1 - A\mathcal{I}_{\text{object}}(\tilde{\mathbf{r}}', R_o) + B_K \\
 &\approx 1 - A_K(R_o + w/2)^2 + B_K \equiv f_K(R_o),
 \end{aligned} \tag{15}$$

where we approximated $\mathcal{I}_{\text{object}}(\tilde{\mathbf{r}}', R_o) \approx (R_o + w/2)^2 \pi \int \frac{d^3\mathbf{r}}{|\tilde{\mathbf{r}}' - \mathbf{r}|^2}$ using Eq. (14) and $(R_o + w/2) \ll |\tilde{\mathbf{r}}' - \mathbf{r}|_{\mathbf{r} \in \delta\mathbf{r}}$. In Fig. 4, we fit the numerical results for different values of K with functions of the form $f_K(R_o)$ (Eq. (15)). The fits are in excellent agreement ($R^2 \sim 0.99$) with the numerical data.

“Eclipsing” can occur both upstream and downstream of an activator

A salient feature of enhancers is that binding sites for TFs are positioned both upstream and downstream of the activators. The eclipse model predicts that $F_\infty(\tilde{\mathbf{r}}'(K), R_o) = F_\infty(\tilde{\mathbf{r}}'(-K), R_o)$, since for long chain lengths the correlation between the positions of both chain termini is completely abolished. Thus, from the perspective of one terminus, looping events can be initiated from

any possible direction. To see if our simulation captures this symmetry, we explored a geometry in which the protein-like protrusion was positioned at negative K values, corresponding to a location outside of the looping segment delimited by the looping-volume (activator) and distal-terminus (promoter) locations. To do so, we generated an additional chain segment of length Q in the direction opposite to \hat{t}_1 , starting from link 0, where $Q \gg K$. We plot the results in Fig. 5. The data show that in the elastic regime, the “outside” geometry (green line, $K < 0$) generates a significantly smaller effect on the probability of looping as compared with the “inside” architecture (blue line, $K > 0$). This lack of symmetry is probably due to the propensity of short loops to form a “tear-drop” shape, thereby reducing the quenching effect of any elements bound outside the looping segment (37). However, for sufficiently large L/b , $F(L)$ for both $\pm K$ enhancer geometries converge to the same value, as predicted by the eclipse model.

Looping at chain ends vs. looping in the middle of the chain

So far, we modeled an enhancer-promoter region by a chain of discrete semi-flexible links, with the activator and the promoter located at both ends of the chain. However, *in vivo*, the DNA chain extends far beyond the enhancer-promoter region in both directions, and is subject to confinement. We address the first issue in this section, and the issue of DNA confinement in the next section.

We computed the probability of looping for a generic loop of length L , located between two internal links of the chain. To do so, we extended the original length of the chain L by two flanking segments of 10 Kuhn lengths ($10b$) on both ends of the chain. We assigned the cumulative Rosenbluth weight of the extended chain to the central segment of length L containing the enhancer-promoter region. See Extended Materials and Methods, Section 1.3.

We ran a simulation with $K = 60$ bp and $R_0 = 11.9$ nm (Fig. 6, magenta line) and compared $F(L)$ to the same case without flanking segments (Fig. 6, blue line). The data show that there is a discrepancy in the long-range looping probability ratio between the two looping models. In particular, the down-regulatory effect of the protrusion is stronger when the enhancer-promoter region is modeled as part of a larger chain. In order to determine the individual contributions of the two flanking segments to the discrepancy, we ran additional simulations with different flanking-segment configurations. We found that only the trailing segment beyond the distal terminus of the looping segment contributes to this discrepancy. The addition of a single chain link of length $w = 4.6$ nm ($0.092b$) after the terminus of the looping segment (Fig. 6, green line) diminishes $F(L)$, accounting for approximately half of the discrepancy. The addition of only 4 links ($0.368b$) after the terminus of the looping segment (Fig. 6, aqua line) accounts for the entire discrepancy, fully agreeing with the results for the case with $10b$ flanking segments on both ends of the looping segment. While an addition of a leading segment before the looping volume diminishes the looping probability (data not shown), it does not alter the looping probability ratio $F(L)$. This can be seen from the comparison between $F(L)$ for the case with a single link after the terminus of the looping segment (Fig. 6, green line) and the case of a segment of length b before the looping volume and a single chain link after the terminus of the looping segment (Fig. 6, red line).

We believe that these results depend strongly on the looping conditions. For the conditions used here ($\delta\omega' = 0.1 \times 2\pi$, see Fig. 1) the leading flanking segment is inconsequential to the looping probability ratio. However, utilizing a uniform looping volume ($\delta\omega' = 4\pi$) would lead to a diminished looping probability ratio as a result of leading segment addition. Similarly, the trailing

flanking segment diminishes $F(L)$ due to the fact that chain configurations that approached the looping volume from a direction colliding with the chain are no longer possible when the chain is extended by a trailing segment. If, however, the looping conditions were such that the chain terminus is required to approach the looping volume from a direction perpendicular to the looping volume cone axis, the trailing segment would become inconsequential to the looping probability ratio.

These results suggest that while there is a discrepancy between the two simulation modes (with flanking regions and without them), it arises from regions that are immediately adjacent to the enhancer-promoter region, and further downstream or upstream segments of the chain do not contribute to the down-regulatory effect.

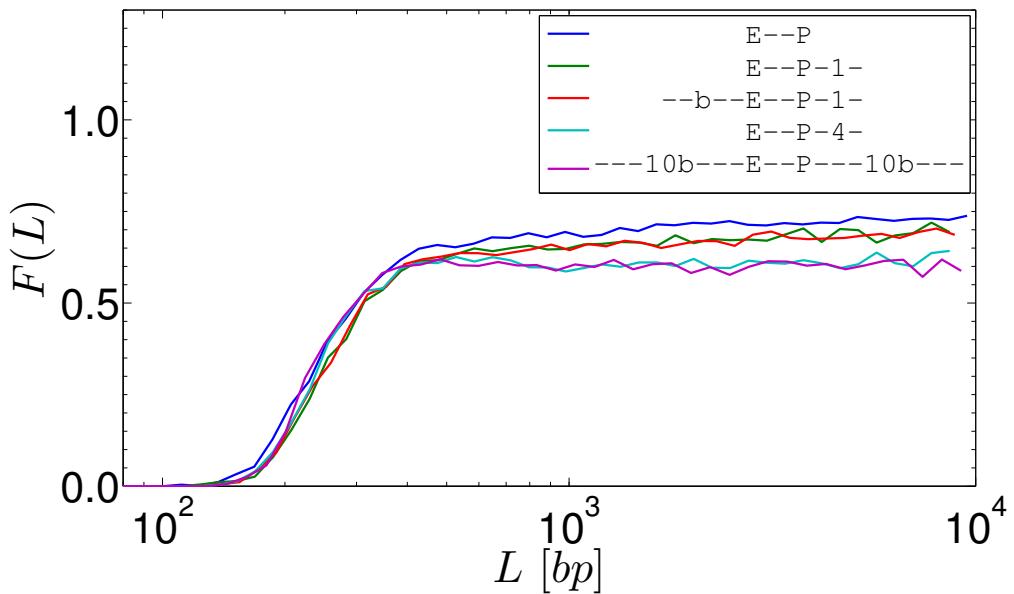


Figure 6: Comparison of $F(L)$ between enhancer-promoter regions with different-sized flanking segments. All simulations ran with $K = 60$ bp and $R_0 = 11.9$ nm. Blue: No flanking segments. Green: One link $w = 4.6$ nm ($0.092b$) after the chain terminus. Red: Medium-sized segment of length b before the looping volume and one link ($0.092b$) after the chain terminus. Aqua: 4 links ($0.368b$) after the chain terminus. Magenta: Long flanking segments of length $10b$ on each side of the enhancer-promoter region.

Looping probability ratio for confined polymer

Chromosomal DNA is fundamentally different from the idealized version assumed in the simulations presented thus far. Cellular DNA is confined, and as a result is compacted into a globular state. This state is characterized by its local nature of fluctuations, as opposed to the coiled state modeled in this work, in which the fluctuations are of the order of the size of the macromolecule itself (40). The nature of the globular state varies between organisms (equilibrated globule in yeast

vs. crumpled or fractal globule in eukaryotes (33, 41)) and spans a wide range of volume fractions (35). For lower volume fractions, DNA can be viewed as a semi-dilute polymer solution, in which coils are strongly overlapping but the volume fraction of the polymer in the solution is small. This semi-dilute solution may be pictured as a system of domains (blobs) of characteristic size ξ . Inside each blob, the chain behaves as an isolated macromolecule with excluded volume, and different blobs are statistically independent of one another (8, 40). If the volume fraction is sufficiently low, linearized chromosomal DNA within a blob can explore nearly the entire volume of the blob without interacting with other parts of the chromosome. To simulate this scenario we generated chains confined by spheres of various radii corresponding to a range of realistic blob sizes (35) (for full simulation details, see Extended Materials and Methods).

In Figure 7, we present the normalized looping probability and the looping probability ratio for the range of confining-sphere radii 125-625 nm. Note that the looping probability increases relative to that of the unconstrained polymer when the gyration radius of the chain becomes comparable to the size of the confining sphere (Fig. 7A). However, as can be seen in Fig. 7B, the introduction of a bound protrusion does not influence the looping probability ratio $F(L)$ up to a chain length of 10 kbp for confining spheres with radii ≥ 250 nm. For the smallest confining sphere (radius of 125 nm) we were only able to simulate chains of length ≤ 4.5 kbp (see Supplementary Information Section 3.4), and here too the looping probability ratio remained unaffected. This implies that our results are valid for ~ 10 kbp of linearized DNA localized within a blob of semidilute chromatin. For additional discussion see Supplementary Information.

We may think of the results of Fig. 7 in a different way. By scaling arguments the simulation also describes a polymer with width of 10 nm ($w \times 10/4.8$), length of ~ 6800 nm ($L \times 10/4.8$) and persistence length of ~ 110 nm ($l_p \times 10/4.8$), which is comparable to ~ 100 kbp of 10 nm chromatin fiber if we assume a density of 15 bp/nm for the fiber (42). The confining sphere forces a specific volume fraction on the confined polymer. The range of confining radii and chain lengths corresponds to polymer volume fractions of up to 0.2%. This interpretation implies that our previous results may be applicable to looping of chromatin fiber at low volume fractions.

“Eclipsing” effect for *eve* 3/7 stripe

Finally, we applied our model to a real enhancer-promoter system. To do so, we make the following assumptions (based on current experimental evidence). First, when the promoter is active, the enhancer and the promoter region are depleted of nucleosomes for up to 500-1000 bp both upstream and downstream from the center of each regulatory segment (43). This implies that if the enhancer and promoter are sufficiently close, the entire segment can be assumed to be mostly bare dsDNA that is decorated by some DNA binding proteins or protrusions. Second, chromatin in the nucleus is in a globular state. Due to locality of concentration fluctuation in the globular state, we assume that the linearized DNA segment, if not too long, remains trapped in a blob. In order to apply our model to a system with these assumptions, we had to choose an experimentally well-characterized enhancer-promoter system, with an enhancer-promoter separation < 10 kbp (consistent with the results presented in Fig. 7B). A likely candidate is the well-characterized *eve* 3/7 enhancer of *D. melanogaster*, which is separated from its promoter by about 4 kbp. In addition, the volume fraction for *D. melanogaster* is estimated between 0.1-1.5% (35). Together, this places the *eve* 3/7 enhancer within the applicability region of our model.

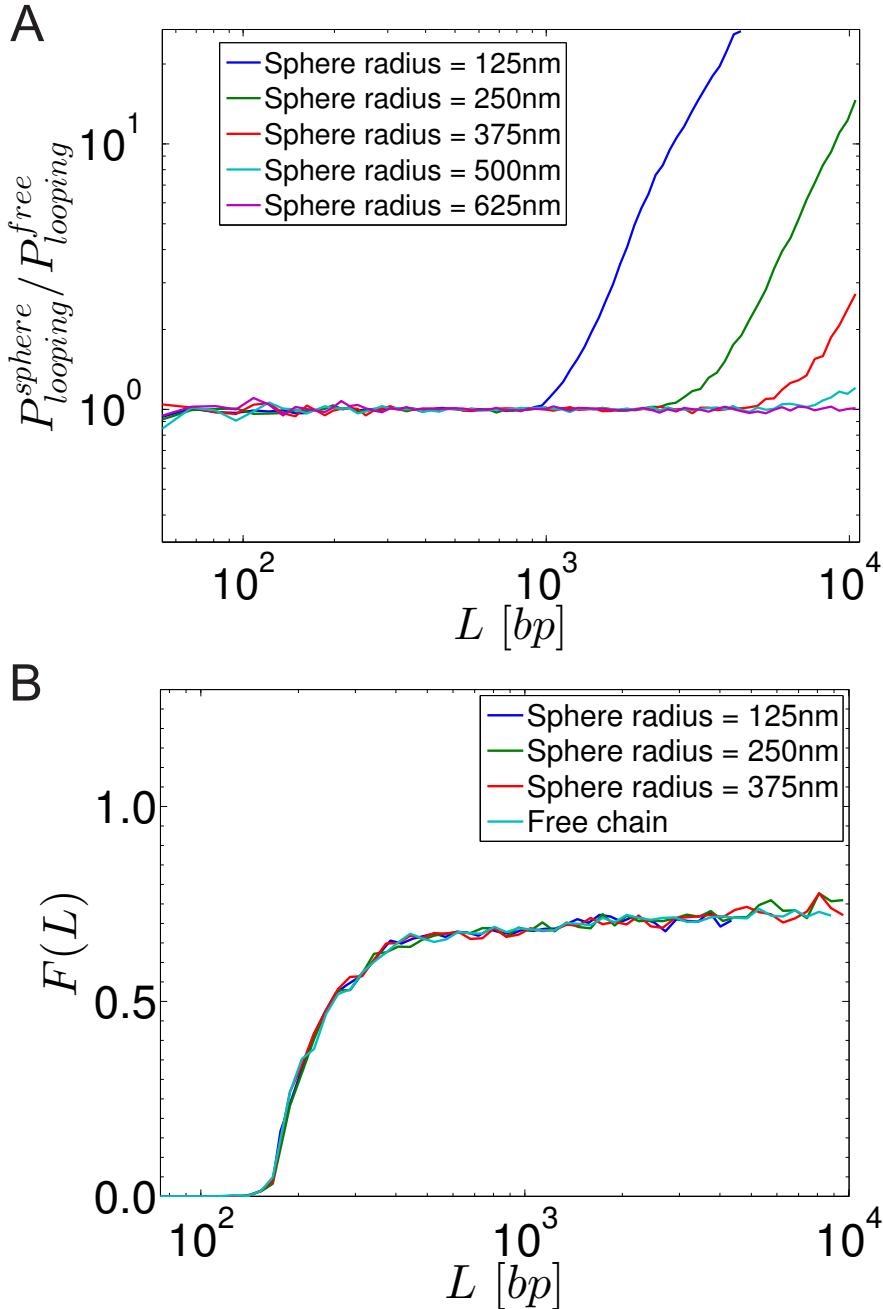


Figure 7: Effect of a confining sphere on the probability of looping. A) Values of $P_{looping}^{sphere} / P_{looping}^{free}$ for various sphere sizes. B) Values of $F(L)$ for three of the smallest spheres compared to $F(L)$ for the chain without a confining sphere.

We computed our model's predictions for the active *eve* 3/7 enhancer, i.e. in the early developmental stages (before gastrulation - stage 14), and in the anterior region of the *D. melanogaster* embryo. To check whether the resultant probability of looping could provide a mechanistic explanation for the observed expression pattern, we assumed two possible states: the enhancer is fully occupied by the activators dStat, Zld and the Knirps repressor–co-repressor complexes, and a sec-

ond state where the enhancer is entirely devoid of bound Knirps repressor complexes. We tested three possible repressor complex sizes: Knirps alone (46 kDa), Knirps bound to CtBP dimers (130 kDa), and a full putative 450 kDa complex, as reported by (23). Since the exact looping conditions of the enhancer looping are unknown, we considered three choices of $\delta\omega'$ (see Materials and Methods) that represent different extreme cases.

The results, plotted in Fig. 8, show that for all chosen looping conditions, there is a decrease in the probability of enhancer looping when the enhancer is fully bound by Knirps. However, only the largest Knirps–co-repressor complex (containing dCtBP and the HDACs Rpd3 (23) and Sin3 (24)) is large enough to generate a significant reduction in looping.

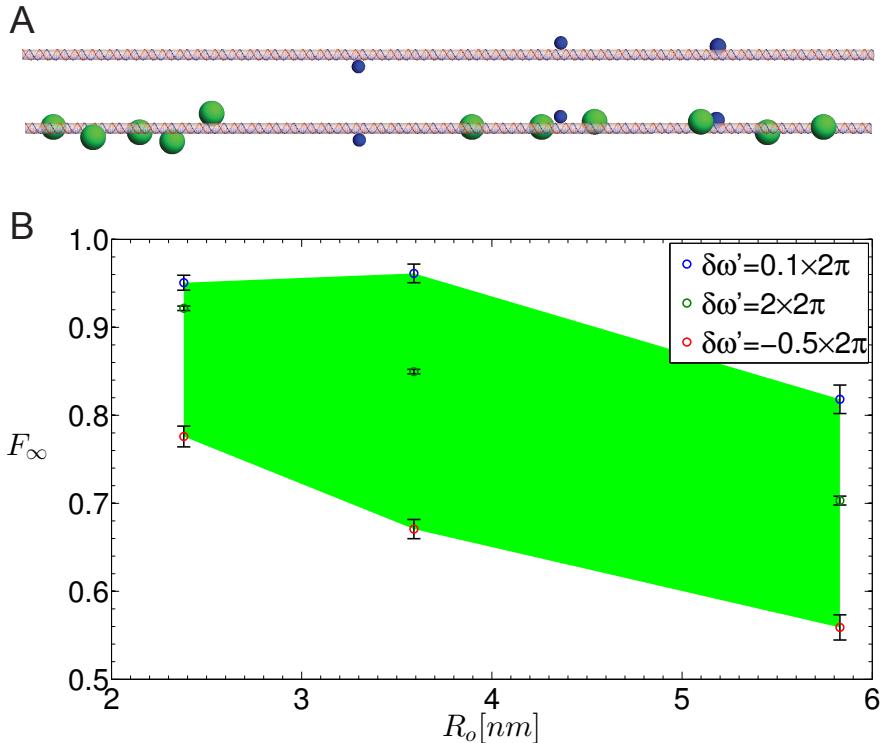


Figure 8: Simulation of the *eve* 3/7 enhancer in *D. melanogaster*. A) Illustration of the *eve* 3/7 enhancer geometry with bound activators (top) vs. with both activators and full putative Knirps+dCtBP+Rpd3+Sin3 complexes (450 kDa) bound (bottom). B) Dependence of F_∞ on protrusion size R_o and the looping conditions $\delta\omega'$. Negative $\delta\omega'$ corresponds to anti-collinear orientation of $(\mathbf{r}_N - \mathbf{r}_0)$ with \hat{u}_0 . Here $d_{min} = R_{chain} + R_{activator}$ and $\varepsilon = 3$ nm. F_∞ was computed from 150 simulated data points corresponding to DNA chains of lengths 3751–3900 bp. Error bars are ± 1.96 times the standard error of these points. The area shaded in green depicts the range of possible F_∞ values, depending on the chosen looping criteria.

DISCUSSION

We previously established (37) that a bound protein inside a loop can alter the probability of looping in a manner proportional to its size, when the chain length is of the order of the Kuhn length (i.e.

< 300 bp). The simulations and theory presented here extend this result to much longer chain lengths. In particular, our model predicts a decrease in the probability of looping that is independent of chain length for long chains in the entropic regime (i.e. $L \gg 300$ bp), provided that a sufficiently large protrusion oriented in-phase with the looping volume δr is positioned within one Kuhn length of one of the chain termini. We further showed that the reduction in looping probability resembles an eclipse-like phenomenon, where the protrusion blocks the line of sight of one chain terminus from the other.

Our model provides a biophysical mechanism for so-called short-range repression or quenching in enhancers (16–19), which successfully captures many of this phenomenon’s salient features. These include lack of dependence on chain length for sufficiently long chains ($L \gg b$), symmetry with respect to binding-site positioning inside or outside of the looping segment, and the dependence of the regulatory effect on both the size of the bound complex and the distance of the TF from the nearest terminus (~ 150 bp) for which significant quenching can be generated. To explore the relevance of our model to natural systems, we computed the looping probability ratio for the *D. melanogaster eve 3/7* enhancer-promoter system in the segment of the embryo where the repressor Knirps is expressed. We found that using realistic repressor-co-repressor-HDAC complex sizes for the simulated protrusions, a significant reduction in looping probability can be obtained, showing that the looping-based mechanism can account for at least some of the quenching generated by Knirps in the context of early fly development. Since there is also a substantial body of evidence supporting HDAC catalysis as a major mechanism for repression (44), we conclude that it is likely that both HDAC activity and looping-related effects combine to generate the quenching observed in *D. melanogaster* and other organisms.

There have been previous attempts to model enhancer regulatory logic for the gap genes of early fly development (45–47). Using a semi-empirical approach, which coupled a thermodynamic model to an empirically-based regulatory scoring function, these works demonstrated that TF occupancy of enhancers can determine the gene-expression outcome. By contrast, our model does not compute the actual expression pattern, but rather the probability of looping for a given occupancy state. For this computation, our model requires only knowledge of the enhancer structure (TF binding sites, protein-complex sizes, and over-all enhancer-promoter distance). This implies that no gene-expression experimental data is needed as an input to predict the down-regulatory effect of bound TFs.

Finally, we consider the range of applicability of our model for natural genomes. While this work focused on bare dsDNA with a protrusion, our results are applicable to any linear polymer as long as its length is significantly larger than its Kuhn length. However, to apply our results to chromosomal DNA, other constraints must be taken into account, some of which we attempted to address in this paper. First, we showed that looping of a linear segment with additional upstream and downstream flanking segments qualitatively resembles looping without flanking segments. Second, chromosomal DNA is subject to confinement, which results in a globular state of the polymer. While there is evidence that active enhancers and their surrounding regions are linear, they are confined by the rest of the chromosome. We attempted to model this constraint using a confining sphere, which allowed us to show that the over-all regulatory effect remains unchanged up to a chain length of ~ 10 kbp for sphere sizes corresponding to realistic blob sizes. Third, DNA is typically a mixture of bare and chromatinized DNA. While we did not address the issue of “mixed” DNA directly, we do know that chromatinized DNA is more densely packed than bare DNA. Thus, the ~ 10 kbp looping results for bare DNA correspond to much longer lengths (up

to \sim 100 kbp) of “mixed” DNA. We can conclude that the 10 kbp range for bare DNA addressed by our simulations is a lower bound on the actual range of validity of our model for “mixed” or chromatinized DNA. Our work indicates that we should consider three ranges for the physics of looping of chromatinized DNA. For short ranges which we studied previously (14, 37), looping is dominated by elastic energy. As shown in this work, there is an intermediate entropic looping range of up to \sim 10 kbp of bare DNA (or \sim 100 kbp of “mixed” DNA). Recent studies indicate that there may be an additional long-range regime in which interactions between neighboring regions of the globular DNA must be taken into account (33, 48, 49).

AUTHOR CONTRIBUTIONS

YP devised and wrote the simulations. YP, SG, and RA analyzed the data and wrote the manuscript.

ACKNOWLEDGEMENTS

This project was funded by the European Union’s Horizon 2020 Research And Innovation Programme under grant agreement 664918—MRG-Grammar, by the Israel Science Foundation through grant 1677/12, by the I-CORE Program of the Planning and Budgeting Committee, and by the Israel Science Foundation (grant 152/11). Y.P. acknowledges support provided by the Russell Berrie Nanotechnology Institute, Technion. The authors declare no conflicts of interest.

SUPPLEMENTARY MATERIAL

An online supplement to this article can be found by visiting BJ Online at <http://www.biophysj.org>.

References

1. Marko, J. F., and E. D. Siggia, 1995. Stretching DNA. *Macromolecules* 28:8759–8770.
2. Müller, J., S. Oehler, and B. Müller-Hill, 1996. Repression of lac Promoter as a Function of Distance, Phase and Quality of an Auxiliary lac Operator. *Journal of Molecular Biology* 257:21–29.
3. Cloutier, T. E., and J. Widom, 2005. DNA twisting flexibility and the formation of sharply looped protein-DNA complexes. *Proceedings of the National Academy of Sciences* 102:3645–3650.
4. Wiggins, P. A., T. van der Heijden, F. Moreno-Herrero, A. Spakowitz, R. Phillips, J. Widom, C. Dekker, and P. C. Nelson, 2006. High flexibility of DNA on short length scales probed by atomic force microscopy. *Nature Nanotechnology* 1:137–141.
5. Vafabakhsh, R., and T. Ha, 2012. Extreme Bendability of DNA Less than 100 Base Pairs Long Revealed by Single-Molecule Cyclization. *Science* 337:1097–1101.

6. Jacobson, H., and W. H. Stockmayer, 1950. Intramolecular Reaction in Polycondensations. I. The Theory of Linear Systems. *The Journal of Chemical Physics* 18:1600–1606.
7. Flory, P. J., 1969. Statistical Mechanics of Chain Molecules. Interscience Publishers. We validated this for our simulations.
8. Gennes, P.-G., 1979. Scaling Concepts in Polymer Physics. Cornell University Press, Ithaca, NY.
9. Shimada, J., and H. Yamakawa, 1984. Ring-closure probabilities for twisted wormlike chains. Application to DNA. *Macromolecules* 17:689–698.
10. Crothers, D. M., J. Drak, J. D. Kahn, and S. D. Levene, 1992. DNA bending, flexibility, and helical repeat by cyclization kinetics. In J. E. D. David M.J. Lilley, editor, Methods in Enzymology, Academic Press, volume Volume 212 of *DNA Structures Part B: Chemical and Electrophoretic Analysis of DNA*, 3–29.
11. Yan, J., R. Kawamura, and J. F. Marko, 2005. Statistics of loop formation along double helix DNAs. *Physical Review E* 71:061905.
12. Buck, M., M.-T. Gallegos, D. J. Studholme, Y. Guo, and J. D. Gralla, 2000. The Bacterial Enhancer-Dependent σ 54 Transcription Factor. *Journal of Bacteriology* 182:4129–4136.
13. Su, W., S. Porter, S. Kustu, and H. Echols, 1990. DNA-looping and enhancer activity: association between DNA-bound NtrC activator and RNA polymerase at the bacterial glnA promoter. *Proceedings of the National Academy of Sciences of the United States of America* 87:5504–5508.
14. Amit, R., H. G. Garcia, R. Phillips, and S. E. Fraser, 2011. Building Enhancers from the Ground Up: A Synthetic Biology Approach. *Cell* 146:105–118.
15. Davidson, E. H., 2006. The Regulatory Genome. Elsevier.
16. Gray, S., P. Szymanski, and M. Levine, 1994. Short-range repression permits multiple enhancers to function autonomously within a complex promoter. *Genes & Development* 8:1829–1838.
17. Gray, S., and M. Levine, 1996. Short-range transcriptional repressors mediate both quenching and direct repression within complex loci in Drosophila. *Genes & Development* 10:700–710.
18. Arnosti, D. N., S. Barolo, M. Levine, and S. Small, 1996. The eve stripe 2 enhancer employs multiple modes of transcriptional synergy. *Development* 122:205–214.
19. Arnosti, D., S. Gary, S. Barolo, J. Zhou, and M. Levine, 1996. The gap protein knirps mediates both quenching and direct repression in the Drosophila embryo. *The EMBO Journal* 15:3659–3966.
20. Nibu, Y., H. Zhang, and M. Levine, 1998. Interaction of Short-Range Repressors with Drosophila CtBP in the Embryo. *Science* 280:101–104.

21. Nibu, Y., H. Zhang, E. Bajor, S. Barolo, S. Small, and M. Levine, 1998. dCtBP mediates transcriptional repression by Knirps, Krüppel and Snail in the Drosophila embryo. *The EMBO Journal* 17:7009–7020.
22. Strunk, B., P. Struffi, K. Wright, B. Pabst, J. Thomas, L. Qin, and D. N. Arnosti, 2001. Role of CtBP in Transcriptional Repression by the Drosophila giant Protein. *Developmental Biology* 239:229–240.
23. Struffi, P., and D. N. Arnosti, 2005. Functional Interaction between the Drosophila Knirps Short Range Transcriptional Repressor and RPD3 Histone Deacetylase. *Journal of Biological Chemistry* 280:40757–40765.
24. Spain, M. M., J. A. Caruso, A. Swaminathan, and L. A. Pile, 2010. Drosophila SIN3 Isoforms Interact with Distinct Proteins and Have Unique Biological Functions. *Journal of Biological Chemistry* 285:27457–27467.
25. Keller, S. A., Y. Mao, P. Struffi, C. Margulies, C. E. Yurk, A. R. Anderson, R. L. Amey, S. Moore, J. M. Ebels, K. Foley, M. Corado, and D. N. Arnosti, 2000. dCtBP-Dependent and -Independent Repression Activities of the Drosophila Knirps Protein. *Molecular and Cellular Biology* 20:7247–7258.
26. Ryu, J.-R., and D. N. Arnosti, 2003. Functional similarity of Knirps CtBP-dependent and CtBP-independent transcriptional repressor activities. *Nucleic Acids Research* 31:4654–4662.
27. Mannervik, M., and M. Levine, 1999. The Rpd3 histone deacetylase is required for segmentation of the Drosophila embryo. *Proceedings of the National Academy of Sciences of the United States of America* 96:6797–6801.
28. Pollak, Y., S. Goldberg, and R. Amit, 2014. Self-avoiding wormlike chain model for double-stranded-DNA loop formation. *Physical Review E* 90:052602.
29. Chen, Z. Y., and J. Noolandi, 1992. Renormalization-group scaling theory for flexible and wormlike polymer chains. *The Journal of Chemical Physics* 96:1540–1548.
30. Tree, D. R., A. Muralidhar, P. S. Doyle, and K. D. Dorfman, 2013. Is DNA a Good Model Polymer? *Macromolecules* 46:8369–8382.
31. Glotzer, S. C., and W. Paul, 2002. Molecular and Mesoscale Simulation Methods for Polymer Materials. *Annual Review of Materials Research* 32:401–436.
32. Kotelyanskii, M., and D. N. Theodorou, 2004. Simulation Methods for Polymers. CRC Press, New York, 1 edition edition.
33. Lieberman-Aiden, E., N. L. van Berkum, L. Williams, M. Imakaev, T. Ragoczy, A. Telling, I. Amit, B. R. Lajoie, P. J. Sabo, M. O. Dorschner, R. Sandstrom, B. Bernstein, M. A. Bender, M. Groudine, A. Gnirke, J. Stamatoyannopoulos, L. A. Mirny, E. S. Lander, and J. Dekker, 2009. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science (New York, N.Y.)* 326:289–293.

34. Vettorel, T., A. Y. Grosberg, and K. Kremer, 2009. Statistics of polymer rings in the melt: a numerical simulation study. *Physical Biology* 6:025013.
35. Halverson, J. D., J. Smrek, K. Kremer, and A. Y. Grosberg, 2014. From a melt of rings to chromosome territories: the role of topological constraints in genome folding. *Reports on Progress in Physics* 77:022601.
36. Rosenbluth, M. N., and A. W. Rosenbluth, 1955. Monte Carlo Calculation of the Average Extension of Molecular Chains. *The Journal of Chemical Physics* 23:356–359.
37. Brunwasser-Meiriom, M., Y. Pollak, S. Goldberg, L. Levy, O. Atar, and R. Amit, 2016. Using synthetic bacterial enhancers to reveal a looping-based mechanism for quenching-like repression. *Nature Communications* 7:10407.
38. Struffi, P., M. Corado, L. Kaplan, D. Yu, C. Rushlow, and S. Small, 2011. Combinatorial activation and concentration-dependent repression of the Drosophila even skipped stripe 3+7 enhancer. *Development* 138:4291–9.
39. Maeshima, K., K. Kaizu, S. Tamura, T. Nozaki, T. Kokubo, and K. Takahashi, 2015. The physical size of transcription factors is key to transcriptional regulation in chromatin domains. *Journal of Physics: Condensed Matter* 27:064116.
40. Khokhlov, A. R., A. Y. Grosberg, and V. S. Pande, 2002. Statistical Physics of Macromolecules. American Institute of Physics, New York, 1994 edition edition.
41. Duan, Z., M. Andronescu, K. Schutz, S. McIlwain, Y. J. Kim, C. Lee, J. Shendure, S. Fields, C. A. Blau, and W. S. Noble, 2010. A three-dimensional model of the yeast genome. *Nature* 465:363–367.
42. Walhout, M., M. Vidal, and J. Dekker, editors, 2012. Handbook of Systems Biology: Concepts and Insights. Academic Press, London, 1 edition edition.
43. Heintzman, N. D., R. K. Stuart, G. Hon, Y. Fu, C. W. Ching, R. D. Hawkins, L. O. Barrera, S. Van Calcar, C. Qu, K. A. Ching, W. Wang, Z. Weng, R. D. Green, G. E. Crawford, and B. Ren, 2007. Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nature Genetics* 39:311–318.
44. Perissi, V., K. Jepsen, C. K. Glass, and M. G. Rosenfeld, 2010. Deconstructing repression: evolving models of co-repressor action. *Nature Reviews Genetics* 11:109–123.
45. Janssens, H., S. Hou, J. Jaeger, A.-R. Kim, E. Myasnikova, D. Sharp, and J. Reinitz, 2006. Quantitative and predictive model of transcriptional control of the Drosophila melanogaster even skipped gene. *Nat Genet* 38:1159–65.
46. Segal, E., T. Raveh-Sadka, M. Schroeder, U. Unnerstall, and U. Gaul, 2008. Predicting expression patterns from regulatory sequence in Drosophila segmentation. *Nature* 451:535–40.

47. Kim, A.-R., C. Martinez, J. Ionides, A. F. Ramos, M. Z. Ludwig, N. Ogawa, D. H. Sharp, and J. Reinitz, 2013. Rearrangements of 2.5 kilobases of noncoding DNA from the *Drosophila* even-skipped locus define predictive rules of genomic cis-regulatory logic. *PLoS Genet* 9:e1003243.
48. Goloborodko, A., J. F. Marko, and L. A. Mirny, 2016. Chromosome Compaction by Active Loop Extrusion. *Biophysical Journal* 110:2162–2168.
49. Grosberg, A. Y., 2016. Extruding Loops to Make Loopy Globules? *Biophysical Journal* 110:2133–2135.