Enhancer-Insulator Pairing Reveals Heterogeneous Dynamics in Long-Distance 3D Gene Regulation

Lucas Hedström [1,*] Ralf Metzler [2,3,†] and Ludvig Lizana [1,‡]

[1]*Integrated Science Lab, Department of Physics, Umeå University, SE-901 87 Umeå, Sweden*
[2]*Institute for Physics & Astronomy, University of Potsdam, DE-144 76 Potsdam-Golm, Germany*
[3]*Asia Pacific Centre for Theoretical Physics, Pohang 37673, South Korea*

Cells regulate fates and complex body plans using spatiotemporal signaling cascades that alter gene expression. Short DNA sequences, known as enhancers (50–1500 base pairs), help coordinate these cascades by attracting regulatory proteins that enhance the transcription by binding to distal gene promoters. In humans, there are hundreds of thousands of enhancers dispersed across the genome, which poses a challenging coordination task to prevent unintended gene activation. To mitigate this problem, the genome contains insulator elements that block enhancer-promoter interactions. However, there is an open problem with how the insulation works, especially as enhancer-insulator pairs may be separated by millions of base pairs. Based on recent empirical data from Hi-C experiments, this paper proposes a new mechanism that challenges the common paradigm that rests on specific insulator-insulator interactions. Instead, this paper introduces a stochastic looping model where insulators bind weakly to chromatin rather than other insulators. After calibrating the model to experimental data, we use simulations to study the broad distribution of hitting times between an enhancer and a promoter when insulators are present. We find parameter regimes with large differences between average and most probable hitting times. This makes it difficult to assign a typical timescale and hints at highly defocused regulation times. We also map our computational model onto a resetting problem that allows us to derive several analytical results. Besides offering new insights into enhancer-insulator interactions, our paper advances the understanding of gene regulatory networks and causal connections between genome folding and gene activation.

I. INTRODUCTION

Cell fates and complex body plans are established through signaling cascades that turn genes on and off in complex spatiotemporal patterns. One of the critical genetic elements that help coordinate these cascades is enhancers. These are short regulatory DNA sequences [50–1500 base pairs (bp)] that attract proteins, such as transcription factors, to "enhance" the transcription of select genes [Fig. 1(a)]. Enhancer elements are often far from the target gene start, sometimes as far as millions of base pairs apart. Yet, experiments show they appear close in 3D to regulate transcription [3].

In humans, the genome harbors hundreds of thousands of dispersed enhancers supporting gene expression networks [4]. Notably, these enhancers do not necessarily act on the closest promoter and may regulate multiple genes [5,6]. This posits a challenging coordination task of all these distal 3D interactions to protect genes from unintended activation. One way cells manage this task is by using insulators. Like enhancers, insulators contain clusters of binding sites for

sequence-specific DNA-binding proteins that block enhancer-promoter interactions. However, insulators are usually a bit larger, spanning ∼300–2000 bp. When first discovered in *Drosophila melanogaster* [7–9], the insulators appeared to define boundaries between different chromatin states. However, researchers soon found they could block enhancer action when inserted at specific genomic loci and that gene activity depended on specific DNA binding proteins associating with the insulator element.

From a genetic point of view, insulators are simply some DNA piece that activates a gene when removed. But the question is how this insulation works mechanistically, especially as some enhancer-promoter distances are so large [3]. One of the most popular mechanistic descriptions is the topological model. This model suggests that two or more insulating elements bind each other to form loops [10]. This idea agrees with Hi-C data from mammals, where CTCF insulator elements (CCCTC-binding factor) make three-dimensional (3D) contacts and often define borders of shielded chromatin communities, so-called topologically associated domains [11,12]. This is further consistent with the loop-extrusion model, where a handcuff-shaped protein (Cohesin) binds and extrudes DNA through itself until it reaches a CTCF site, thus creating a loop with CTCF as the anchor points [13,14]. The topological model also agrees with extensive polymer simulation studying insulation in varying enhancer-insulator-promoter configurations (measured by reduced 3D contact probabilities) [15].

However, there are recent experimental data that challenge this paradigm [16]. This data set comes from Hi-C

---

*Contact author: lucas.hedstrom@umu.se
†Contact author: rmetzler@uni-potsdam.de
‡Contact author: ludvig.lizana@umu.se
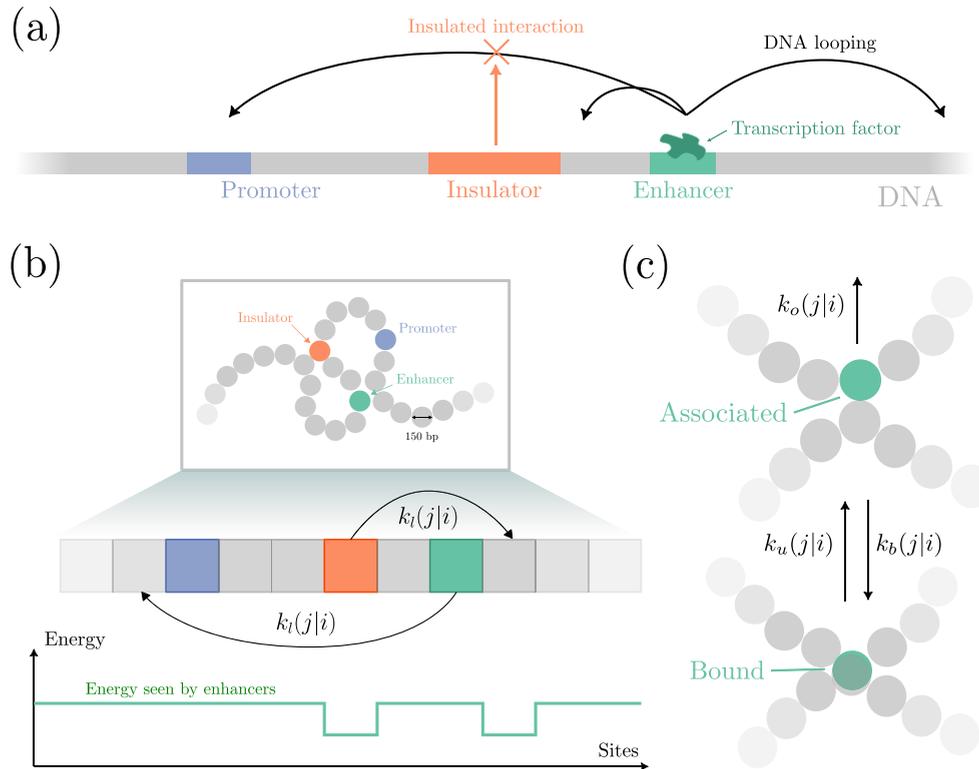
FIG. 1. Schematics of enhancer-promoter-insulator constructs on DNA (a) and our three-state looping model (b), (c). (a) The insulator region blocks long-ranged enhancer-promoter interactions. Reducing such interactions causes "insulation" as the transcription factor cannot physically assist the transcription machinery (e.g., RNA polymerases) that interacts with the promoter. (b) Three-state lattice model. We model the chromatin as a sequence of N sites. A few of these sites represent enhancers (green) or insulators (orange). They can loop and associate to any other site with the rate $k_l(j|i)$. This rate balances looping entropy and binding energy, where this energy changes with their time-varying relative positions. The line graph below shows a snapshot of the energy landscape experienced by the enhancer. This landscape has two "dips" corresponding to the insulator loop anchor. The other sites represent "inactive" chromatin. Unlike enhancers and insulators, we consider chromatin as an ensemble of interacting loops and include them by adjusting the enhancer-promoter looping exponent [1,2]. We color the promoter in blue. In the model, we treat the promoter as an absorbing target. (c) Local binding dynamics. After looping, the enhancer and insulator can either strongly associate (bind) to the new site by k_b or dissociate back to its original position by k_o (similar to a resetting process). If bound, then the return rate to the associated state is k_u.

experiments from *D. melanogaster* and contains the contact counts between thousands of insulator-insulator pairs. By clustering these pairs based on their genomic separation, the paper convincingly shows that insulators do not contact each other more frequently than with general chromatin. This suggests the topological insulator model needs revision (at least in *Drosophila*).

Building on this observation, our paper proposes an alternative mechanism where insulators bind weakly to surrounding chromatin rather than other insulators. We formulate our model on a lattice with stochastic looping dynamics where we calibrate the rates to existing experimental data and benchmark to existing measured contacts from Ref. [16]. Next, we use our model to study the dynamics of enhancer-promoter hitting times and show that the average may deviate substantially from the typical (most probable). We also map our computational model onto a resetting problem and provide several analytical results. Our work offers new insights into the enhancer-insulator mechanics. Besides yielding a better understanding of gene regulatory networks, knowing how insulators work may help unveil causal relationships connecting gene expression and genome folding [17,18].

II. METHODS

A. Looping model for enhancer-insulator dynamics

We represent the chromatin as an array of sites $i = 1, \ldots, N$, where each site symbolizes a nucleosome ($\approx$175 bp), the basic chromatin unit. This choice is natural from a biological (or epigenetic) point of view [19–21] but not critical for our general framework. The array has four site types: enhancers, promoters, insulators, and regular chromatin (Fig. 1). In most simulations, we consider one enhancer and one promoter, typically placed 20–50 array indices between one another ($\approx$3.5–8.5 $\times 10^3$ kb), and up to 20 insulators. However, actual gene clusters usually have a much richer arrangement where several enhancers and insulators act in concert to ensure proper gene expression of many genes. But to better appreciate the model, we study simpler configurations.

As mentioned above, enhancers are DNA elements that attract regulatory proteins, such as transcription factors. While being attached to DNA, the transcription factors try to find the enhancers associated with the designated promoter to regulate transcription. In other words, protein-bound enhancers

make repeated looping attempts with surrounding chromatin until they reach the target site. Our model captures these dynamics by considering that the enhancers are always bound by transcription factors and that the enhancers and insulators loop and interact with the nearby chromatin at rates we define below.

Our model treats the promoter as an absorbing point, and the simulation stops once the enhancer complex reaches there (i.e., infinite reaction rate). This contrasts the interaction with surrounding chromatin which is much weaker. We model enhancer-chromatin binding using standard transition-state theory where the enhancer may be bound with energy E_b or loosely associated (E_a) from where it may detach (Fig. 1). These two states ("bound" and "associated") are similar to "search" and "recognition" modes often used to represent two-state searchers to explain the so-called speed-stability paradox in DNA-target search problems [22,23]. Furthermore, by separating the model where it either loops or binds to the sequence, we avoid having to tune every parameter in correlation with each other, simplifying the process of determining the parameter space.

In addition to E_a and E_b, there is an energy barrier ΔG_b separating the bound and associated states (called activation energy in transition-state theory). Below (Sec. III A), we estimate these parameters from actual transcription factor binding data.

In addition to surrounding chromatin, the enhancer also binds to insulators. We model this binding using transition state theory but assign different values for E_a, E_b, and ΔG_b. Importantly, because the insulators are dynamic objects like the enhancer, they form loops with surrounding chromatin (discussed below), meaning these energies change with site index j (and time). Therefore, for some fixed insulator configurations, we express the unbinding and binding rates as

$$k_u(j) = \gamma e^{-[\Delta G_b(j) - E_b(j)]},$$
$$k_b(j) = \gamma e^{-[\Delta G_b(j) - E_a(j)]}, \tag{1}$$

where γ is the basal binding rate, on the order 10^7 s^{-1} [24,25], and the thermal energy $k_B T$ is set to unity. Since $k_b(j)/k_u(j) = \exp[E_a(j) - E_b(j)]$, these rates obey detailed balance.

Next, we discuss the enhancer's looping rates. Similar to previous work [26–29], these rates depend on the entropy cost of forming the loop $k_B \ln[(\ell/\ell_0)^{-\alpha}]$ where ℓ is the loop length, α is the looping exponent, and ℓ_0 is a characteristic loop scale [2]. To calibrate constants to obey typical DNA-looping times, we again use transition-state theory and add a small binding activation energy ΔG_l. Given these parameters, the looping on-rate from lattice site i to j is

$$k_l(j|i) = \begin{cases} \delta e^{-\alpha \ln(\ell_{ij}/\ell_0) - \Delta G_l} & \text{if } i \neq j, \\ 0 & i = j, \end{cases} \tag{2}$$

where δ is the basal looping frequency, which is on the order of $10^3 \mathrm{s}^{-1}$ [24], and $\ell_{ij} = d_{\text{nucl.}} \times |i - j|$ is the loop length; $d_{\text{nucl.}}$ is the nucleosome diameter (~ 10 nm, or 175 bp).

The looping off-rate follows a similar formula as Eq. (1), but it contains the energy activation associated with the looped state ΔG_l instead of $\Delta G_b(j)$ since we imagine this is the only state from which the loop can break apart [it is also unphysical

that the loop rate depends on the position-dependent $E_a(j)$]. Thus, to unloop from the bound state [$E_b(j)$], the enhancer must first become "associated" [$E_a(j)$] and then unloop. In summary, the looping off-rate is

$$k_o(j|i) = \delta e^{-[\Delta G_l - E_a(j)]}, \tag{3}$$

where the two looping rates obey detailed balance $k_l(j|i)/k_o(j|i) = \exp\{-E_a(j) - [-\alpha \ln(\ell_{ij}/\ell_0)]\}$. In our model, the site returns to its starting position after looping, similar to other models [27,30,31]. We tune the time between looping events using l_0 and ΔG_l to account for the noncorrelated loops, which we discuss later.

Before showing how we calibrate the binding energies to experimental data, we make three comments. First, even if the discussion above was mostly about enhancers, we assume insulators to follow the same dynamics [Eqs. (1)–(3)], albeit with slightly different energy parameters. For example, we assume that insulators interact weekly with chromatin. They do so with a binding constant that should not be smaller than $K \sim 10^2$ μM, which is the characteristic scale for specific binding (estimated from *Escherichia coli* [32]). While our model does not rest on specific insulator-insulator interactions, they are not excluded. But they cannot be significantly larger than for general chromatin. If they were, then insulator pairs would appear as high-contact stripes in Hi-C maps. This was not observed in Ref. [16]. Therefore, we let these interactions have the same strength as regular insulator-chromatin interactions.

We also point out that the insulators in our simulations constantly form loops and, therefore, appear on two different lattice positions [Fig. 1(b)] from the point of view of the enhancer. One position is always fixed and coincides with the insulator's designated DNA segment. The second one represents the other end of the insulator-chromatin loop. As this loop is short-lived, the second position is highly dynamic and switches frequently and symmetrically around the insulator's primary position during the simulation (Fig. 7). This effect implies that the enhancer has two possibilities to bind the insulator [Fig. 1(c)].

Second, thus far, we have discussed DNA looping, omitting regular chromatin. But in reality, chromatin also fluctuates. Instead of introducing specific looping rates akin to Eqs. (2) and (3), we treat chromatin as an ensemble of interacting loops and modify the looping exponent α for enhancers and insulators accordingly. For long self-avoiding chains, the exponent (often denoted "ring factor") is

$$\alpha \approx d\nu - 2\sigma, \tag{4}$$

where d is the embedding dimension, ν is the exponent associated with the polymer's radius of gyration, and σ is the "scaling factor" ($\sigma = 0$ for noninteracting loops) [33]. Using $d = 3$, $\nu = 0.588$, and $\sigma = -0.175$ gives $\alpha = 2.114$ [1,34].

Third, some papers include a bending term $\propto 1/\ell_{ij}$ in the looping rates [27]. This contribution accounts for the bending energy cost when creating short loops. We omit this since we only consider loops that are much longer than the DNA's Kuhn length (≈ 300 bps).

Fourth, we acknowledge that there are different choices for discretizing chromatin, each leading to unique downstream effects. First, there are intrinsic limits on how small we can set the lattice spacing without adding new biophysical

mechanisms. As pointed out above, if the spacing is smaller than the DNAs persistence (or Kuhn) length, then we would have to incorporate a bending rigidity into the free energy calculations for loop formation. Conversely, if the lattice spacing is too large, then such a model fails to capture the decaying looping probabilities along the DNA chain, resulting in large regions with uniform probabilities in conflict with Hi-C data. Finally, there is no straightforward method to map different discretization levels while keeping microscopic variables like the mean hitting time constant due to position-dependent resetting rates. Hence, any change in lattice spacing requires a re-calibration of the models parameters.

III. RESULTS

A. Matching model parameters to data

To calibrate the model to empirical data, we estimate $E_a(j)$, $E_b(j)$, and $\Delta G_b(j)$ for binding and association, and l_0 and ΔG_l for the looping process. To this end, we use comprehensive binding data for transcription factors and measured *in vitro* looping rates. We start with association and binding.

Previous work calculated the energy (and free-energy) landscape along DNA for individual transcription factors (TFs) that bind to 10–30 bp long sequence-specific motifs [24,35,36]. But because DNA is so much longer (and there are only four base pair types), there are many instances of almost similar binding sequences. This results in fluctuating genome-wide binding profiles interpreted as the TF's spatial probability density $p_{TF}(x)$, where x denotes the DNA coordinate. From this probability density, it is common to define the energy profile $\mathcal{E}(x)$, assuming that $p_{TF}(x) \sim \exp[-\mathcal{E}(x)]$. We defer to Supplemental Material Sec. SII ([37]) for technical details about obtaining $\mathcal{E}(x)$ from a given TF target sequence. In Supplemental Material Sec. SII [37], we plot the average binding data for about 300 TFs that we use to estimate $E_a(j)$, $E_b(j)$, and $\Delta G_b(j)$ for the enhancers and insulators. In our lattice notation $\mathcal{E}(x) \to E(j)$.

To estimate $E_a(j)$, $E_b(j)$, and $\Delta G_b(j)$ for the enhancer, we use the formalism developed [24] for a two-state TF flipping between "search" and "recognition" mode while searching for a DNA-target sequence. When in recognition mode, the TF is immobile, and the residence time depends on how similar the local and target sequences are to each other. In other words, the time depends on the depth of the landscape $E(j)$. When in search mode, the TF diffuses and only weakly interacts with the DNA. Reference [24] also assumes that the effective energy landscape during the search mode is a scaled version of $E(j)$, i.e., $\rho E(j)$, where ρ is adjusted to agree with measured 1D diffusion constants. In practice, this means setting $\rho \leqslant 0.3$; we use $\rho = 0.3$. Last, there is a free energy barrier ΔG_{RS} separating "search" and "recognition" mode that we also extract from TF-binding data (see below and in Supplemental Material Sec. SII [37]). In summary, the equations to calculate $E_a(j)$, $E_b(j)$, and $\Delta G_b(j)$ read

$$E_a(j) = \rho E(j),$$
$$E_b(j) = \Delta G_{RS} + E(j), \qquad (5)$$
$$\Delta G_b(j) = \Delta G_{RS} + \frac{1+\rho}{2} E(j),$$

where all the parameters on the right-hand side come from empirical data. In particular, we used TF data from the JASPAR database [38] to extract $E(j)$ and ΔG_{RS}. We apply these formulas to the three binding instances we have in our problem: enhancer-insulator, insulator-insulator, and unspecific (enhancer-chromatin and insulator-chromatin.) We used $\Delta G_{RS} = 10.13$ in all three cases.

1. Enhancer-insulator binding

We calculated $\langle E(j) \rangle_{\text{binding sites}}$ for all human chromosomes and available transcription factors from the JASPAR database [38]. Using Eq. (5), we find that the population median [Md(·)] is $\text{Md}(\langle E(j) \rangle_{\text{binding sites}}) = -21.13$ (see Supplemental Material Sec. SII [37]). This gives $E_a(j) = -6.34$, $E_b(j) = -11.0$, and $\Delta G_b(j) = -3.60$, which is a system having low binding energy, corresponding to a low association energy and energy barrier [see Fig. 2(a), lower orange line].

2. Unspecific binding for enhancers

Here we set $E(j) = 0$. Plugging this into Eq. (5) gives $E_a(j) = 0$, $E_b(j) = 10.13$ and $\Delta G_b(j) = 10.13$. This situation is where strong association (binding) is rare and weak [Fig. 2(a), green upper line].

3. Unspecific binding for insulators

Here we set $E(j) = -15.0$. Plugging this into Eq. (5) gives $E_a(j) = -4.5$, $E_b(j) = -4.9$ and $\Delta G_b(j) = 0.39$.

Next, we match the looping rates for enhancers and insulators. To set the corresponding parameters l_0 and ΔG_l, we use experimental data from [39–41]. These report that typical looping times for 300 bps long loops are 10^1–10^2 seconds (LacI protein and restriction enzymes NaeI and NarI), and 10^3 seconds for 3500 bp long loops. Matching with our simulation gives $l_0 \approx 0.04$ and $\Delta G_l = 0.0$. However, the looping times *in vivo* may deviate substantially in crowded cell conditions, where far-away sections come in contact faster than expected due to compartmentalization [18]. To this end, we vary l_0 to study the fast- and slow-looping regimes.

B. Sticky insulators reproduce measured contact differences in *Drosophila melanogaster* embryos

In this section, we benchmark our model to empirical data from Ref. [16]. Using Hi-C experiments, this paper quantifies how effectively insulators block 3D interactions between flanking DNA segment pairs. By collecting contact profiles for hundreds of insulator positions in mutant and wild-type Fruit fly embryos, the authors made two key observations. First, the insulators block 3D interactions over distances up to 200 kb. We replot the empirical data in Fig. 2(b) for two replicate experiments (Rep. 1 and Rep. 2). The background is derived from the same experiments but uses random loci far from any insulators. Second, they also measured 3D contacts between insulator pairs and could not detect any specific binding, which contradicts the standard topological model.

To reproduce the measured contact decay in Fig. 2(b) using our model, we constructed a large-scale version where each lattice site matches the resolution of the data (i.e., 5 kb), rather than a single nucleosome (~0.2 kb). In the middle of
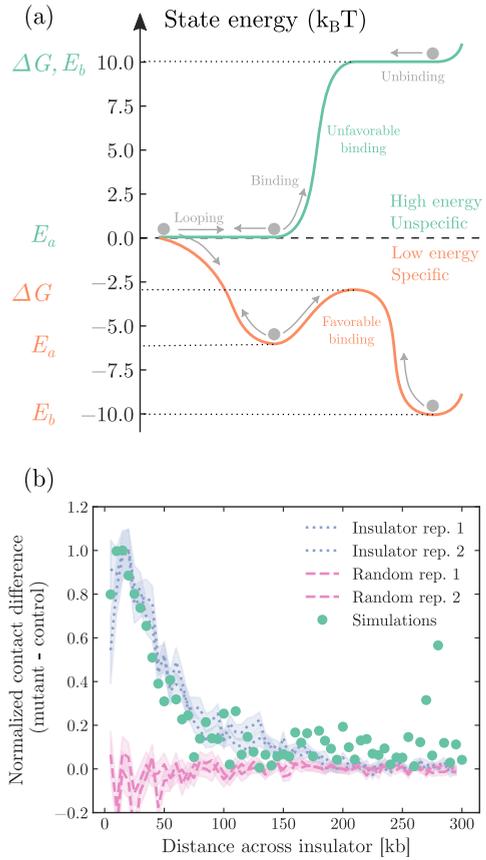
FIG. 2. Energy levels and fit to experimental data [16]. (a) Two typical interaction cases: high and low interaction energy (green and orange). The energy is high for nonspecific sites (gray sites in Fig. 1) that corresponds to a weak association energy (E_a), a weaker bound energy (E_b), and a large barrier ΔG separating E_a and E_b. This is a fleeting binding event that case differs from the enhancer-insulator interactions, where the interaction energy is low (orange). Here, the association and bound energy are both small and separated by a smaller energy barrier. (b) Recent experiments found that the contact difference between DNA regions separated by insulator-rich segments increases if the insulators are compromised (i.e., genetically engineer a mutant lacking the CP190 binding factor that binds to insulators) [16]. We replot the empirical data here (with permission) as dotted blue and purple lines with $\pm$ one standard deviation (shading). To match the data, we simulated a center-placed insulator region and calculated the contact probability with each flanking site. By setting the lattice spacing to the Hi-C data resolution (5 kb), we found an excellent agreement between our simulations and the data if the insulator-chromatin interaction energy is $E(j) = -17$ ($\forall j$) and the looping scale is $l_0 = 0.4$. We normalized the y-axis by dividing by the maximum value of our simulations and the experimental data.

the lattice, we placed an insulator-rich region (one site) that weakly associates with all other sites with the same energy [$E(j) = \text{const.}$] and forms loops. To get the contact frequencies, we simulate repeated looping events of all flanking sites across the insulator according to rates $k_l(i|j)$ and k_o [Eqs. (2) and (3)] and record the residence times between all site pairs using the Gillespie method implemented in Julia, see Supplemental Material Sec. SI ([37]). Next, we collected these

residence times over 10^2 simulations ($\approx 10^6$ time steps each) and used them as a proxy for contact probabilities. To fit the model to the dotted lines in Fig. 2(b), we calculated the relative difference in these probabilities with and without the insulator [Fig. 2(b), green dots].

We note that our model agrees well with the empirical data using $E(j) = -17$ as unspecific binding for insulators and $l_0 = 0.4 \leftrightarrow 2000$ bp. This length scale agrees with general coarse-grained "beads-on-a-string" polymer models for chromatin. For example, if modeled as a freely jointed chain, each monomer should contain 2000–25000 bp of DNA [42]. This value differs slightly from our previous fitting using *in vitro* looping times, where $l_0 = 0.2 \leftrightarrow 1000$ bp. In the remaining part of the paper, we return to a nucleosome-centric model and use $l_0 = 0.24$ and the unspecific binding $E(j) = -15$ as the standard settings (albeit we also study the effects of l_0 variations). But, in summary, this section shows that our model can reproduce empirical data from Hi-C measurements across an ensemble of insulators in *D. melanogaster*.

C. Insulator densities strongly affect enhancer-promoter hitting frequencies

One of the paper's key aims is to better understand the hitting dynamics between the enhancer and promoter elements under insulation, as such an encounter is the primary step in transcription. In particular, we wish to calculate the distribution of time interval lengths between hitting events and study how they change with key variables such as insulator densities, positions and binding strengths. To this end, we perform Gillespie simulations using the rates outlined in Sec. II A. The simulations use a 200 lattice site system (≈ 35 kbp), where the enhancer and promoter reside in the middle, separated by 30 sites in the standard setting (≈ 5 kbp) [Fig. 3(a)]. On these 30 sites, we put an insulator region with length n_{ins}. A typical simulation produces $\approx 10^5$ samples, where we record the time t_a for the enhancer to reach the promoter site for the first time.

We show several simulated histograms in Fig. 3(b) with varying insulator density ($\sigma_{ins} = n_{ins}/20 = 0.0, 0.1, 0.5, 1.0$) (see legend). We portray the histogram in two ways to highlight different features. The lower panels show histograms $\rho(t_a)$, where the area is normalized to unity ("Density"). The upper panels ("Probability") show the enhancer's probability $\mathcal{P}(t_a)$ of finding the promoter within a specific time interval Δt. We calculate this probability from the density ρ as

$$\mathcal{P}(t_a) = \int_{t_a}^{t_a + \Delta t} \rho(u)du \approx \rho(t_a) \times \Delta t. \tag{6}$$

Above the histograms, we also show the accumulated probability as colored stripes. These stripes provide an intuition for the weight of data points, where each shift into a shaded color indicates 10 percentiles of the data. Last, we indicate the average and median hitting times by filled arrows above the $\mathcal{P}(t_a)$ curves, where the larger arrow indicates the mean.

Consider the leftmost histogram column. Plotted in logscale, we note that the hitting-time distributions are broad and that most data points follow the same trend until $t_a \sim 10^{-1}$
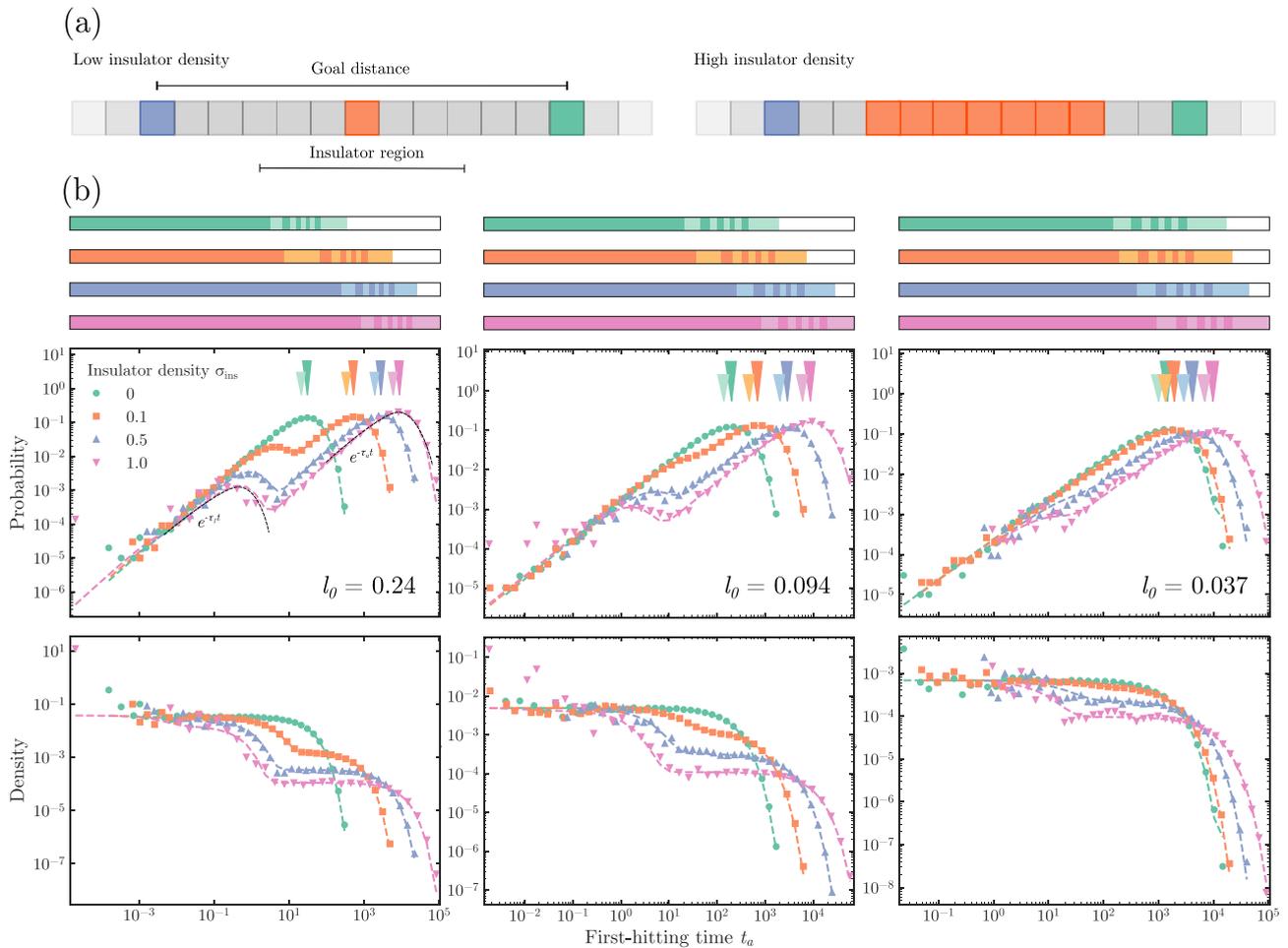
FIG. 3. Histograms of the first-hitting times for varying insulator density σ_{ins} and looping scale l_0. (a) Schematics of enhancer-insulator promoter configurations for two insulator densities ("low" and "high"). The distance between the enhancer (left, blue) and the promoter (green, right) is 30 sites. The insulators are placed in the middle (orange), where we increase the density by extending the insulator region (from 0 to 20). We used a spacing of 5 sites between the insulators and the enhancer and promoter to avoid effects simply due to proximity. (b) Log-binned histograms of the first-hitting times for different insulator densities. We portray the histograms in two ways: area-normalized (bottom, "Density") and hitting-time probability [see Eq. (6)] (top, "Probability"). The colors and markers correspond to a varying amount of insulators (a). The colored bars above show the 10th percentiles of all data points. The filled and shaded arrows above the lines indicate the mean (filled) and median (shaded) first-hitting times, respectively. We observe a few peaks and valleys in the histograms, corresponding to different search trajectories, such as the enhancer finding the promoter on the first try (potentially with a few intermediate fleeting chromatin interactions) or getting sequestered by the insulator for a significant period. Note the clear exponential shoulders in the probability distribution. These correlate to the exponential decay of the wide plateaus of constant first-hitting time probabilities, seen in the density.

and then spread apart. This means the short-time dynamics are insulator-independent. We interpret this as the enhancer loops into the target after a few short-lived encounters with surrounding chromatin without touching the insulator.

Plotting $\mathcal{P}(t_a)$ instead of $\rho(t_a)$ unveils a series of peaks corresponding to different types of search trajectories. The left peaks are associated with enhancer-promoter contact events not involving insulators (as explained above). This contrasts with the right peaks, representing the most probable hitting time. These peaks shift towards larger t_a with growing insulator densities, making the distributions orders of magnitude broader.

Apart from simulated data (symbols), the plots contain several dashed lines with identical colors. These lines closely follow the data points and represent a numerical inverse

Laplace transform of a theoretical search-and-resetting model for $\rho(t_a)$ [Eq. (17)], that we derive in Sec. IV. In addition, two black dashed lines show local exponential fits for the two peaks. In particular, the right black line follows the simple relationship $\rho(t_a) \sim \exp(-t_a/\tau_a)$, where τ_a is the mean hitting-time derived analytically [Eq. (19), Sec. IV]. All dashed curves show a good agreement between simulations and the analytics.

Let us consider all the histograms in Fig. 3(b). Each column shows the hitting-time distribution for varying looping scale l_0 (values are indicated in the bottom right). While this parameter is often interpreted as the Kuhn length, it is also a proxy for chromatin compaction. As discussed in Ref. [18], typical looping scales differ significantly for different types of chromatin folding and density (e.g., space-filling versus
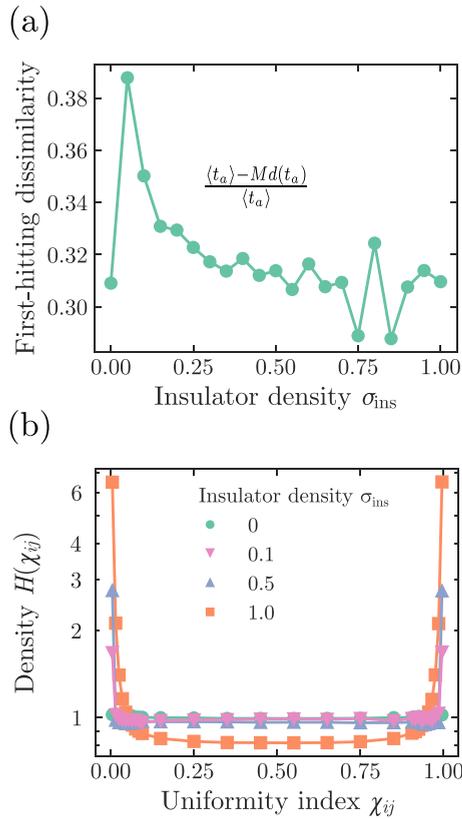
FIG. 4. Heterogenous search trajectories. (a) Relative difference between mean and median (see the inset equation) from Fig. 3(b) (filled and shaded arrows). Note the nonmonotonic growth and decay. (b) Uniformity index histograms (normalized), $H(\chi_{ij})$ for varying insulator densities. The bimodal shapes indicate heterogeneous enhancer-promoter search trajectories.

self-avoiding). By lowering l_0 in our model, thus mimicking less dense folding, we note that the left peaks disappear and that the histograms become insensitive to the insulator density. If l_0 increases, then the peaks become more pronounced, indicating increasing enhancer-insulator interactions due to less time spent looping (and thus allowing for more possible contacts), resulting in greater insulation.

D. Enhancer-promoter search trajectories are highly dissimilar

In the previous section, we noted a broad distribution of hitting times t_a [Fig. 3(b)]. This suggests that there is a significant difference between average and typical search times [indicated by filled arrows in Fig. 3(b)]. To quantify these differences, we plot their relative distance with increasing insulator density σ_{ins} in Fig. 4(a) (for fixed $l_0 = 0.24$). Without insulators, the mean and median are relatively similar (the mean is slightly larger). But as the insulator count increases, the average and the mean start to deviate until reaching some threshold density, where the trend reverses. This threshold density occurs at a relatively low value (around 0.1), which might change depending on the system configuration, such as insulator strength, placement or discretization. We interpret this trend change as a result of two competing timescales—

one short, governed by chromatin interactions, and one long, dominated by insulator-enhancer interactions.

To further explore the heterogenous difference between search times, we adopt a theoretical framework based on the uniformity index ω_{ij} [43–46]. This index is defined as the ratio of one search time versus the sum of two search times for two randomly chosen trajectories i and j:

$$\chi_{ij} = \frac{t_a^i}{t_a^i + t_a^j}. \tag{7}$$

If most samples are similar, then the histogram of χ_{ij} values, $H(\chi_{ij})$, follow a bell-shaped curve centered around $1/2$. However, if the samples are dissimilar, where some are short and others are long, then the distribution is broad or even bimodal with two peaks $\chi_{ij} = 0$ and $\chi_{ij} = 1$. We plotted the similarity distribution $H(\chi_{ij})$ in Fig. 4(b) for several insulator densities. We see that the histograms are never bell-shaped. Instead, they are bimodal and become increasingly so for growing insulator density (σ_{ins}). This shows yet again that there is no well-defined scale describing enhancer-promoter hitting times. The distribution indicates that there is a significant portion of ultrashort search times where the enhancer immediately loops to the promoter and a large fraction of orders-of-magnitude longer trajectories [note the t_a variation in Fig. 3(b)], where the enhancer gets sequestered at the insulator region, possibly several times, before reaching the promoter site. Because of the heterogeneous dynamics, it is difficult to assign a typical enhancer-promoter hitting timescale describing the breadth of timescales. Finally, we point out that this heterogeneity gets reinforced with unsuccessful binding enhancer-promoter binding attempts when they must come together several times before forming a stable complex. However, we do not explore this aspect further here.

E. Insulation efficiency

The histograms in Fig. 2(b) show that the distribution of hitting times changes with insulator density σ_{ins} and looping scale l_0. But there is yet another critical variable: the interaction energy E between insulators and enhancers. Here, we explore the relationship between E and the average hitting time τ_a and where it is most sensitive.

To study this question, we simulated and calculated τ_a analytically [Eq. (19)] for several values of E and σ_{ins}, keeping l_0 fixed [Fig. 7(a)]. Simulations for very strong binding energies are omitted due to the very long simulation times. The figure depicts that increasing the insulator density generally leads to higher τ_a. However, the slope is much higher for stronger enhancer-insulator interaction (large negative E values). This suggests that the insulator density has little influence on τ_a when the interactions are too weak.

To better illustrate this finding, we extracted the slope from the end of each curve [see Fig. 5(a)] and plotted it against E [Fig. 5(b)]. Repeating this procedure for three different l_0 values, we note a sigmoidlike behavior, where the sensitivity switches significantly in a tight E-range and remains constant outside. This indicates that there is a critical lowest binding energy E^* required to tune repression using the insulator size; insulator lengths vary about one order of magnitude in cells (200–3000 bp). The sigmoid curves also show that increasing
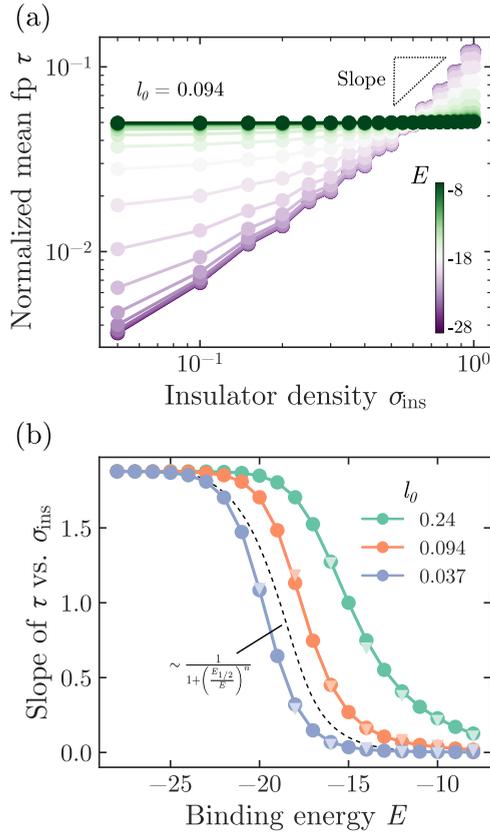
FIG. 5. The efficiency of insulation. (a) Analytical average enhancer-promoter hitting time for varying enhancer-insulator interaction energies (E, see color bar) and insulator density. To increase readability, we normalized the curves so the y values sum to 1. Note how the slope increases with strong interactions (smaller E). (b) Slope sensitivity of τ_a to E, calculated from the large σ_{ins} behavior [see panel (a)]. Analytical solutions are shown as circles and lines, with simulations shown as triangles. For all three l_0 values, the curves show the same sigmodal behavior with a strong E. The dashed line represents an inverted Hill curve, which after fitting gives $E_{1/2} = -14.4, -17.4, -19.6$ and $n = 8.2, 14.1, 18.8$ for the three $l_0 = 0.24, 0.094, 0.037$ values.

the enhancer-insulator interaction energy beyond some value does not lead to greater sensitivity. Moderate energies are enough, which happen to coincide with actual transcription factor binding energies (see Supplemental Material Fig. S1 [37]).

To appreciate the magnitude of the switching behavior, we fitted an inverted Hill curve $\sim[1 + (E_{1/2}/E)^n]^{-1}$ to the data in Fig. 5(b). The fitting yields quite extreme Hill coefficients in the range $n = 10 - 20$. This indicates ultrasensitive switch-like behavior when $E \approx E_{1/2}$. The fitted Hill coefficients are much higher than observed for TF-operator sites, which are usually smaller than 4 [47].

Furthermore, we note that the overall sensitivity decreases with shorter looping times (E^* or $E_{1/2}$ becomes lower with l_0). We envision this reflects cases when the enhancer can explore several sites before finding the promoter, which increases the chance of finding an insulator and, thus, increases insulation.

IV. EFFECTIVE RESETTING THEORY FOR ENHANCER-PROMOTER DYNAMICS

In this section, we derive an analytical theory that allows us to calculate several quantities and better understand our simulation results. For example, we calculate an exact expression for the first-hitting time density in Laplace transform space $\rho_a(s)$, its first and second moments ($\langle t_a \rangle$ and $\langle t_a^2 \rangle$), and the enhancer's position probability density function $P_i(t)$ ($i = 1, \ldots, N$). To achieve this, we express our model as a target-search problem with resetting in a random energy landscape.

We envision an enhancer sitting at site i_0 looping out to a site i to find a target at $i = a$. When reaching site $i (\neq a)$, the enhancer stays associated for a short while and may either return directly to i_0 or become bound at i. Regardless of which, the enhancer eventually returns to i_0 like a resetting event, see the schematic in Fig. 6. The difference between the two scenarios is that the residence time is longer if bound, where the enhancer must pass through the "associated" state before it may reset to site i_0. Without insulators, the resetting rate r is the same across all lattice sites, and this system constitutes a standard resetting problem studied by several authors [48–52]. In our notation, the master equation for $P_i(t)$ when the resetting rate is constant reads

$$\frac{dP_i(t)}{dt} = \sum_j \omega(i, j) P_j(t) - r P_i(t)$$
$$+ \sum_{j \neq a} r P_j(t) \delta_{i,i_0} - \rho_a(t) \delta_{i,a}, \qquad (8)$$

where we write the jumping rates as

$$\omega(i, j) = \begin{cases} -\sum_{k \neq i_0} k_l(k|j), & j = i = i_0, \\ k_l(i|j), & j = i_0, \\ 0, & \text{otherwise.} \end{cases} \qquad (9)$$

Here, the diagonal element $\omega(i_0, i_0) = -\sum_{k \neq i_0} k_l(k|i_0)$ represents all outgoing loops from the enhancer's position i_0 to anywhere on the lattice, and $\omega(i_0, j) = k_l(j|i_0)$ is the loop from i_0 to j. In addition to $\omega(i, j)$ and r, the master equation includes a sink term $\rho_a(t)\delta_{i,a}$ for the target, where $\rho_a(t)$ is the first-hitting time distribution [53,54]. Because of this sink, we must compensate with another term proportional to the survival probability $\sum_{j \neq a} P_j(t)$ (third term, right-hand side) that ensures there is no resetting if the enhancer already reached the target.

Equation (8) is solvable using standard methods. However, the situation changes when there are insulators. It changes because enhancers and insulators may bind each other, thus creating a nonuniform binding landscape. This manifests as a position-dependent resetting rate $r(j|i)$ at site j given that the searcher started at some site i. This is a more complex problem than Eq. (8) that cannot be solved with standard methods.

To find $r(j|i)$ as a function of the rates and energies outlined in Sec. II A, we assume that the "bound" and "associated" states are in equilibrium. This assumption gives
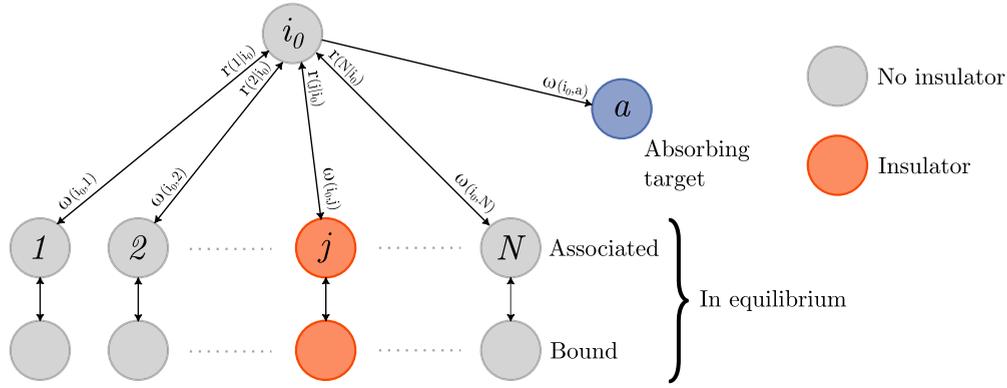
FIG. 6. Schematic of the effective model that captures the general model in Sec. II A. All rates are shown for an enhancer (or insulator) that starts at i_0, jumps to some node j with a rate $\omega(i_0, j)$, and resets back to i_0 rate $k(j|i_0)$. Note that the resetting rate is a combination of rates, assuming equilibrium between the associated and bound states (see Sec. II A and Supplemental Material Sec. SV [37]). At some of the sites, an insulator might be present (marked in orange), which affects the resetting rate.

(see derivation in Supplemental Material Sec. SIII [37])

$$r(j|i) = \frac{k_o(j|i)}{1 + \frac{k_b(j|i)}{k_u(j|i)}} = \delta \frac{e^{-[\Delta G_l - E_a(j)]}}{1 + e^{E_a(j) - E_b(j)}}. \quad (10)$$

It is essential to realize that this resetting rate depends critically on the insulators' positions as they determine the binding landscape $E_b(j)$ when forming loops with the surrounding chromatin. We show below how we estimate $E_b(j)$ assuming the insulators' spatial probability density is in equilibrium. But first, we reformulate the master equation Eq. (8) with a position-dependent resetting rate in an analytically solvable form.

To this end, we write the resetting rate as

$$r(j|i) = r^* + \Delta r(j|i), \quad (11)$$

where $r^* = \min_j [r(j|i)]$ is the smallest resetting rate, and redefine the looping matrix to

$$\nu(i, j) = \begin{cases} \omega(i, j) - \Delta r(j|i_0), & j = i, \\ \omega(i, j) + \Delta r(j|i_0), & i = i_0, \\ \omega(i, j) & \text{otherwise.} \end{cases} \quad (12)$$

Using this in Eq. (8) gives

$$\frac{dP_i(t)}{dt} = \sum_j \nu(i, j)P_j(t) - r^* P_i(t)$$
$$+ r^* Q_a(t)\delta_{i,i_0} - \rho_a(t)\delta_{i,a}, \quad (13)$$

where $Q_a(t) = \sum_{j \neq a} P_j(t)$ is the survival probability. This master equation is now in a standard form and analytically solvable. It represents one of our paper's main results. But before presenting the analytical solution, we outline the basic arguments for including the insulator dynamics in the resetting rate.

In our simulations, we update the positions of the enhancer and the insulators by drawing loops with rates $k_l(j|i)$. If the enhancer and one of the insulators happen to end up on the same lattice site, then they form a complex with rate $k_b(j)$. To include this process in the resetting rate $r(j|i)$, we make two assumptions. First, we take advantage of the timescale separation between the insulators' probability density function

and characteristic search times (the quantity of main interest) and assume that the insulators' probability density function is stationary. We calculate this probability distribution $P_j^{(\text{ins})}$ analytically from Eq. (8) by leaving out the target [see Eq. (21)]. We overlay a few examples of $P_j^{(\text{enh})}$ alongside simulated data in Fig. 7(a); the agreement is excellent.

Second, we assume that the resetting rate $r(j|i)$ for an enhancer is a weighted combination of the resetting rate with or without an insulator at site j. Denoting the rates for these cases as $r(j|i)_{\text{ins}}$ and $r(j|i)_{\text{no ins}}$, we obtain

$$\frac{1}{r(j|i)} \approx \frac{1 - p_{\text{ins}}(j)}{r(j|i)_{\text{ins}}} + \frac{p_{\text{ins}}(j)}{r(j|i)_{\text{no ins}}}. \quad (14)$$

We estimate $p_{\text{ins}}(j)$ as the sum of all the insulator's probability distributions at site j, times the probability of a binding event occurring before the insulator leaves

$$p_{\text{ins}}(j) = \sum_{\text{ins}} P_j^{(\text{ins})}(t \to \infty) \frac{r(j|i)_{\text{ins}}}{r(j|i)_{\text{ins}} + r(k|j)}, \quad (15)$$

where k denotes the insulator's starting site (similar to i_0 for the enhancer).

We point out that approximation (14) tends to overestimate the resetting probability. It also changes when insulators have a varying $E^{(\text{ins})}$ since it sums over all possible binding energies when, in reality, the enhancer can only be bound to one insulator.

In summary, our theoretical model is completely defined by Eqs. (10) and (13)–(15) that allow us to calculate $\rho_a(t)$ analytically. However, we note that some of the above assumptions start to fail when the site-specific binding is so strong that the residence times are comparable with typical target-search times (e.g., the free-energy barrier $\Delta G_b(j)$ should not be too high; see Supplemental Material Fig. S2 [37]).

A. Analytical solution and asymptotic expansion for the first-hitting density

In this section, we solve Eq. (13) analytically. First, we diagonalize the matrix ν [with elements in Eq. (12)] into $\nu = \mathbf{V} \mathbf{D} \mathbf{V}^{-1}$ where each column $\mathbf{V}$ is the eigenvector of ν and $\mathbf{D}$ has the eigenvectors $\lambda_1, \ldots, \lambda_N$ (sorted from smallest to
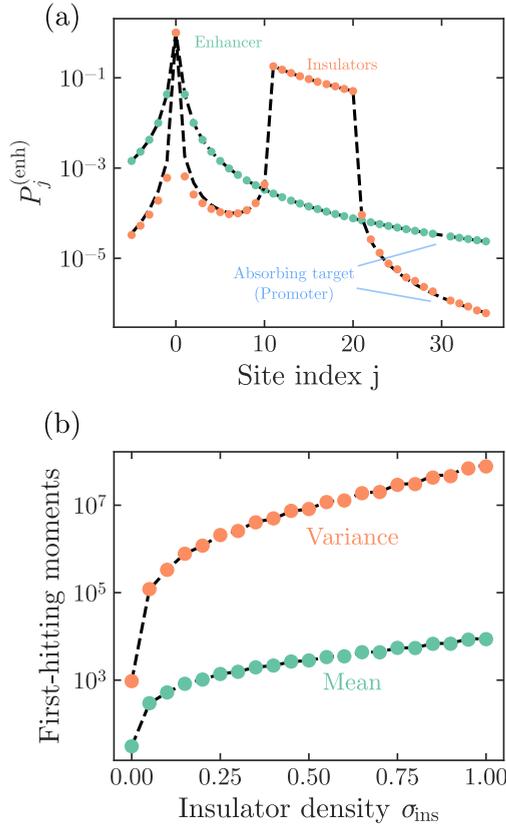
FIG. 7. Benchmarking simulations (symbols) and resetting theory (dashed black lines). (a) The probability density function of the enhancer positioned at lattice site $i = 0$ (indicated by the peak in the distribution) for two cases, one without insulators (green dots) and ten insulators (orange dots) placed in between the enhancer and the absorbing boundary at $i = 30$ (the promoter). The insulators are placed following the schematic in Fig. 3(a), representing the case when the density is 0.5. In the case without insulators, the PDF follows a power-law dropoff, as expected. In the case when insulators are present, we note that the probability of being bound to the insulators decreases the probability at other indices, effectively slowing down search times. (b) The mean and variance of the first-hitting time between enhancer and promoter for varying insulator density. Both of the quantities increase with an increase in density.

largest) along the diagonal (i.e., $D_{ii} = \lambda_i$ and zero otherwise). Next, we introduce the new variable $q_j(t) = \sum_i V_{ji}^{-1} P_i(t)$ and take a Laplace transform ($\mathcal{L}\{f\}(s) = \int_0^\infty f(t)e^{-st}t$). This gives

$$
sq_j(s) - V_{ji_0}^{-1} = \lambda_j q_j(s) - r^* q_j(s) \\
+ \frac{r^*}{s} V_{ji_0}^{-1} - \rho_a(s)\left(\frac{r^*}{s} V_{ji_0}^{-1} + V_{ja}^{-1}\right), \quad (16)
$$

where we used that the survival probability $Q_a(t) = 1 - \int_0^t \rho_a(t)dt$ is $Q_a(s) = [1 - \rho_a(s)]/s$. As the final steps we obtain a closed form expression for $\rho_a(s)$ by solving Eq. (16) for $q_j(s)$ and use the absorbing boundary condition

$\sum_j V_{aj} q_j(s) = P_a(s) = 0$. This gives

$$
\rho_a(s) = \frac{(s + r^*) \sum_j \frac{A_j}{s+r^*-\lambda_j}}{\sum_j \frac{r^* A_j + s B_j}{s+r^*-\lambda_j}}, \quad (17)
$$

where $A_j = V_{aj} V_{ji_0}^{-1}$ and $B_j = V_{aj} V_{ja}^{-1}$. We remind that this analytical solution holds for any enhancer-promoter-insulator configuration (captured in $r(j|i)$ and $\boldsymbol{v}$) under the assumption of fast binding dynamics.

This solution allows us to extract the average hitting time analytically in terms of $\boldsymbol{v}$. By expanding $\rho_a(s)$ for small s (long-time limit) and using

$$
\rho_a(s) \simeq 1 - \langle t_a \rangle s + \frac{s^2}{2}\langle t_a^2 \rangle - \dots, \quad (18)
$$

we obtain to first-order in s that [55]

$$
\langle t_a \rangle = \frac{\sum_{j \neq 1} \frac{B_j - A_j}{r^* - \lambda_j}}{A_n + r^* \sum_{j \neq 1} \frac{A_j}{r^* - \lambda_j}}. \quad (19)
$$

A similar expansion focusing on the second-order terms yields the second moment

$$
\langle t_a^2 \rangle = \frac{2}{\left(A_1 + \sum_{j \neq 1} \frac{r^* A_j}{r^* - \lambda_j}\right)^2} \\
\times \left\{\left[\sum_{j \neq 1}\left(-\frac{r^* A_j}{(r^* - \lambda_j)^2} + \frac{B_j}{r^* - \lambda_j}\right)\right]\left(\sum_{j \neq 1}\frac{B_j - A_j}{r^* - \lambda_j}\right) \\
+ \left(A_1 + \sum_{j \neq 1}\frac{r^* A_j}{r^* - \lambda_j}\right)\left(\sum_{j \neq 1}\frac{B_j - A_j}{(r^* - \lambda_j)^2}\right)\right\}. \quad (20)
$$

In the Supplemental Material ([37], see Sec. SIV), we provide an explicit formula for the variance $\sigma^2 = \langle t_a^2 \rangle - (\langle t_a \rangle)^2$.

To test the theory against simulations, we compare the variance and mean in Fig. 7(b). The simulations (symbols) and the theory (black dashed lines) show excellent agreement.

As a final result, we show that the analytical result for the probability density $P_i(t)$ is valid for both enhancers and insulators. If ignoring the absorbing target (thus setting $Q_a(s) = 1$), then Eq. (16) becomes

$$
P_i(t) = \sum_j V_{ij} V_{ji_0}^{-1} \frac{r^* - \lambda_j e^{(\lambda_j - r^*)t}}{r^* - \lambda_j}, \quad (21)
$$

where the initial condition here is $P_i(0) = \delta_{i,i_0}$. In the steady state, this equation simplifies to

$$
P_i(t \to \infty) = V_{i1} V_{1i_0}^{-1} + \sum_{j \neq 1} \frac{r^*}{r^* - \lambda_j} V_{ij} V_{ji_0}^{-1}, \quad (22)
$$

which we used to estimate the insulators' positions in the resetting rate $r(j|i)$. We show the complete derivation in the Supplemental Material ([37], Sec. SV).

V. DISCUSSION AND CONCLUSION

Cells use a complex web of enhancer-insulator interactions to support gene regulatory networks and orchestrate signaling cascades during development. Insulator elements are critical to prevent unintended gene activation and are relatively simple

from a genetic point of view—removing them from DNA or abolishing the associated transcription factors causes gene activation. Yet, it remains unclear how this happens because of recent conflicting empirical observations [16]. This paper presents a new mechanistic model where insulators bind weakly to surrounding chromatin rather than to themselves, which is the common assumption.

This assumption largely derives from so-called insulator bypass [56–58]. This phenomenon refers to a family of genetic experiments aiming to neutralize insulator-enhancer blocking by genetically inserting a piece of foreign DNA. Known as transgenic constructs, these DNA pieces contain yet another insulator (specifically, a short DNA sequence that attracts insulator-binding factors), and several studies show that these constructs help remove the blocking when sandwiched between the insulator-enhancer pair. The governing explanation for these observations is that insulators pair up and form a loop. This theory has yet further support from measurements showing that some insulator-binding factors can bind each other (e.g., Ref. [58]). However, most transgenic experiments study short-ranged interactions, usually less than 5 kb. While insulator pairing could be the primary insulation mechanism at short distances, it is doubtfully so over long distances. This is the principal observation in [16] using Hi-C data, having 5 kb as the lower resolution limit. They could not detect notable insulator-insulator interactions across thousands of pairs.

The observations from Ref. [16] are the starting point of this work. We aimed to establish a new biophysical model that did not rely on specific insulator-insulator interactions. Instead, it rests on generic but weak insulator-chromatin interactions. From a few assumptions, we calibrated the model to a few independent empirical datasets and derived analytical results that fit empirical observations.

One assumption that likely represents an oversimplification is that the insulators have equal binding strengths to enhancers and the surrounding chromatin. We also limit this study to simple enhancer-insulator-promoter configurations. But in reality, gene clusters have more complex arrangements, including many enhancers, insulators, and promoters, all having heterogeneous binding strengths. It would be instructive to investigate how hitting frequencies between select enhancer-promoter pairs respond to changes in these variables. For example, will one strong insulator do the same job as a few weak ones? We leave these questions for future work.

Another finding emerging from our simulations is a broad range of enhancer-promoter hitting times. This finding connects to the broader discussion of whether the mean-first-passage-time alone is a sufficient measure for biochemical reactions. For example, in reactions with both reaction and diffusion control, mean times split into two additive contributions under macroscopic reactant concentrations. However, this approach fails when specific proteins exist at very low and fluctuating concentrations, such as signaling in biological cells. In such scenarios, e.g., in gene-gene communication, distinct distance dependencies between the communicating genes emerge [59]. This also emerges in generic settings, where reaction times may span orders of magnitude [44,60]. These findings argue that mean reaction times lose their meaning when measuring individual realizations from one cell to another. What we observe here is a similar situation. The enhancer-promoter hitting times span a pronounced range rather than being nicely synchronized. According to our results, this variability indicates that the system is more "plastic" and imprecise in its individual steps. Whether this has any particular biological advantage for enhancer-promoter regulation remains an open problem, but it is a notable consequence of our model.

To close, mammalian genomes harbor $\sim$20 000 genes regulated by $\sim$900 000 enhancer-like elements interspersed with $\sim$30 000 CTCF-bound sites, many of which act as insulators [61]. These elements represent critical components of gene regulatory networks that also seem to shape DNA's spatial organization by forming complex networks of 3D interactions and semihierarchical 3D communities [62] (e.g., topologically associated domains and A/B compartments). Thus, unveiling the mechanisms of insulation is a critical step to understanding the causal mechanisms of the structure-function relationship of interphase chromosomes.

[1] L. Schäfer, C. von Ferber, U. Lehr, and B. Duplantier, Renormalization of polymer networks and stars, Nucl. Phys. B **374**, 473 (1992).

[2] A. Hanke and R. Metzler, Entropy loss in long-distance DNA looping, Biophys. J. **85**, 167 (2003).

[3] C. Chakraborty, I. Nissen, C. A. Vincent, A.-C. Hägglund, A. Hörnblad, and S. Remeseiro, Rewiring of the promoter-enhancer interactome and regulatory landscape in glioblastoma orchestrates gene expression underlying neurogliomal synaptic communication, Nat. Commun. **14**, 6446 (2023).

[4] A. Panigrahi and B. W. OMalley, Mechanisms of enhancer action: The known and the unknown, Genome Biol. **22**, 108 (2021).

[5] P. Geyer and I. Clark, Protecting against promiscuity: The regulatory role of insulators, Cell. Mol. Life Sci. **59**, 2112 (2002).

[6] M. Mohrs, C. M. Blankespoor, Z.-E. Wang, G. G. Loots, V. Afzal, H. Hadeiba, K. Shinkai, E. M. Rubin, and R. M. Locksley, Deletion of a coordinate regulator of type 2 cytokine expression in mice, Nature Immunol. **2**, 842 (2001).

[7] A. Udvardy, E. Maine, and P. Schedl, The 87A7 chromomere: Identification of novel chromatin structures flanking the heat shock locus that may define the boundaries of higher order domains, J. Mol. Biol. **185**, 341 (1985).

[8] C. Holdridge and D. Dorsett, Repression of HSP70 heat shock gene transcription by the suppressor of hairy-wing protein of *Drosophila melanogaster*, Mol. Cell. Biol. **11**, 1894 (1991).

[9] A. M. Bushey, E. Ramos, and V. G. Corces, Three subclasses of a *Drosophila* insulator show distinct and cell type-specific genomic distributions, Genes Dev. **23**, 1338 (2009).

[10] J. R. Raab and R. T. Kamakaka, Insulators and promoters: Closer than we think, Nat. Rev. Genet. **11**, 439 (2010).

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

[12] J. R. Dixon, S. Selvaraj, F. Yue, A. Kim, Y. Li, Y. Shen, M. Hu, J. S. Liu, and B. Ren, Topological domains in mammalian genomes identified by analysis of chromatin interactions, Nature (London) **485**, 376 (2012).

[13] A. L. Sanborn, S. S. Rao, S.-C. Huang, N. C. Durand, M. H. Huntley, A. I. Jewett, I. D. Bochkov, D. Chinnappan, A. Cutkosky, J. Li *et al.*, Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes, Proc. Natl. Acad. Sci. USA **112**, E6456 (2015).

[14] K. Polovnikov and B. Slavov, Topological and nontopological mechanisms of loop formation in chromosomes: Effects on the contact probability, Phys. Rev. E **107**, 054135 (2023).

[15] B. Doyle, G. Fudenberg, M. Imakaev, and L. A. Mirny, Chromatin loops as allosteric modulators of enhancer-promoter interactions, PLoS Comput. Biol. **10**, e1003867 (2014).

[16] T. G. Kahn, M. Savitsky, C. Kuong, C. Jacquier, G. Cavalli, J.-M. Chang, and Y. B. Schwartz, Topological screen identifies hundreds of Cp190-and CTCF-dependent *Drosophila* chromatin insulator elements, Sci. Adv. **9**, eade0090 (2023).

[17] Y. B. Schwartz and G. Cavalli, Three-dimensional genome organization and function in *Drosophila*, Genetics **205**, 5 (2017).

[18] J. Dekker and L. Mirny, The 3D genome as moderator of chromosomal communication, Cell **164**, 1110 (2016).

[19] L. Lizana, N. Nahali, and Y. B. Schwartz, Polycomb proteins translate histone methylation to chromatin folding, J. Biol. Chem. **299**, 105080 (2023).

[20] M. J. Lundkvist, L. Lizana, and Y. B. Schwartz, Forecasting histone methylation by Polycomb complexes with minute-scale precision, Sci. Adv. **9**, eadj8198 (2023).

[21] S. Berry, C. Dean, and M. Howard, Slow chromatin dynamics allow polycomb target genes to filter fluctuations in transcription factor activity, Cell Systems **4**, 445 (2017).

[22] O. Bénichou, C. Loverdo, M. Moreau, and R. Voituriez, Intermittent search strategies, Rev. Mod. Phys. **83**, 81 (2011).

[23] M. P. Kochugaeva, A. A. Shvets, and A. B. Kolomeisky, How conformational dynamics influences the protein search for targets on DNA, J. Phys. A: Math. Theor. **49**, 444004 (2016).

[24] M. Cencini and S. Pigolotti, Energetic funnel facilitates facilitated diffusion, Nucleic Acids Res. **46**, 558 (2018).

[25] C. G. Kalodimos, R. Boelens, and R. Kaptein, A residue-specific view of the association and dissociation pathway in protein–DNA recognition, Nat. Struct. Biol. **9**, 193 (2002).

[26] O. Bénichou, Y. Kafri, M. Sheinman, and R. Voituriez, Searching fast for a target on dna without falling to traps, Phys. Rev. Lett. **103**, 138102 (2009).

[27] C. Felipe, J. Shin, and A. B. Kolomeisky, DNA looping and DNA conformational fluctuations can accelerate protein target search, J. Phys. Chem. B **125**, 1727 (2021).

[28] A. A. Shvets and A. B. Kolomeisky, The role of DNA looping in the search for specific targets on DNA by multisite proteins, J. Phys. Chem. Lett. **7**, 5022 (2016).

[29] S. Vemulapalli, M. Hashemi, A. B. Kolomeisky, and Y. L. Lyubchenko, DNA looping mediated by site-specific SfiI–DNA interactions, J. Phys. Chem. B **125**, 4645 (2021).

[30] Q. Lu, D. Bhat, D. Stepanenko, S. Pigolotti *et al.*, Search and localization dynamics of the CRISPR-Cas9 system, Phys. Rev. Lett. **127**, 208102 (2021).

[31] M. A. Lomholt, T. Ambjörnsson, and R. Metzler, Optimal target search on a fast-folding polymer chain with volume exchange, Phys. Rev. Lett. **95**, 260603 (2005).

[32] R. Phillips, J. Kondev, J. Theriot, and H. Garcia, *Physical Biology of the Cell* (Garland Science, New York, NY, 2012).

[33] Y. Kafri, D. Mukamel, and L. Peliti, Denaturation and unzipping of DNA: Statistical mechanics of interacting loops, Physica A **306**, 39 (2002).

[34] We note that this exponent is not necessarily the same for all regions across the genome. To this end, we re-created some of the main results using a different exponent $\alpha = 3 \times 0.588 \approx 1.764$, where 0.588 is the Flory exponent in 3D. This is shown in Supplemental Material ([37]), Sec. SVI. We found that there is no qualitative difference in our results between the exponents.

[35] G. D. Stormo and D. S. Fields, Specificity, free energy and information content in protein–DNA interactions, Trends Biochem. Sci. **23**, 109 (1998).

[36] O. G. Berg and P. H. von Hippel, Selection of DNA binding sites by regulatory proteins: Statistical-mechanical theory and application to operators and promoters, J. Mol. Biol. **193**, 723 (1987).

[37] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PRXLife.2.033008 for further details on the simulations, certain derivations and analysis, and which includes Refs. [24,25,33,38,63–67].

[38] O. Fornes, J. A. Castro-Mondragon, A. Khan, R. Van der Lee, X. Zhang, P. A. Richmond, B. P. Modi, S. Correard, M. Gheorghe, D. Baranašić *et al.*, JASPAR 2020: Update of the open-access database of transcription factor binding profiles, Nucleic Acids Res. **48**, D87 (2020).

[39] L. Finzi and J. Gelles, Measurement of lactose repressor-mediated loop formation and breakdown in single DNA molecules, Science **267**, 378 (1995).

[40] B. v. d. Broek, F. Vanzi, D. Normanno, F. S. Pavone, and G. J. Wuite, Real-time observation of DNA looping dynamics of Type IIE restriction enzymes NaeI and NarI, Nucleic Acids Res. **34**, 167 (2006).

[41] A. S. Weintraub, C. H. Li, A. V. Zamudio, A. A. Sigova, N. M. Hannett, D. S. Day, B. J. Abraham, M. A. Cohen, B. Nabet,

D. L. Buckley *et al.*, YY1 is a structural regulator of enhancer-promoter loops, Cell **171**, 1573 (2017).

[42] L. A. Mirny, The fractal globule as a model of chromatin architecture in the cell, Chrom. Res. **19**, 37 (2011).

[43] T. G. Mattos, C. Mejía-Monasterio, R. Metzler, G. Oshanin, and G. Schehr, Trajectory-to-trajectory fluctuations in first-passage phenomena in bounded domains, in *First-Passage Phenomena and Their Applications* (World Scientific, Singapore, 2014), pp. 203–225.

[44] D. S. Grebenkov, R. Metzler, and G. Oshanin, Strong defocusing of molecular reaction times results from an interplay of geometry and reaction control, Commun. Chem. **1**, 96 (2018).

[45] T. G. Mattos, C. Mejía-Monasterio, R. Metzler, and G. Oshanin, First passages in bounded domains: When is the mean first passage time meaningful? Phys. Rev. E **86**, 031143 (2012).

[46] C. Mejía-Monasterio, G. Oshanin, and G. Schehr, First passages for a search by a swarm of independent random searchers, J. Stat. Mech.: Theory Exp. (2011) P06022.

[47] U. Alon, *An Introduction to Systems Biology: Design Principles of Biological Circuits* (CRC Press, Boca Raton, FL, 2019).

[48] M. R. Evans and S. N. Majumdar, Diffusion with stochastic resetting, Phys. Rev. Lett. **106**, 160601 (2011).

[49] A. Chechkin and I. M. Sokolov, Random search with resetting: A unified renewal approach, Phys. Rev. Lett. **121**, 050601 (2018).

[50] X. Durang, S. Lee, L. Lizana, and J.-H. Jeon, First-passage statistics under stochastic resetting in bounded domains, J. Phys. A: Math. Theor. **52**, 224001 (2019).

[51] M. R. Evans, S. N. Majumdar, and G. Schehr, Stochastic resetting and applications, J. Phys. A: Math. Theor. **53**, 193001 (2020).

[52] C. Di Bello, A. V. Chechkin, A. K. Hartmann, Z. Palmowski, and R. Metzler, Time-dependent probability density function for partial resetting dynamics, New J. Phys. **25**, 082002 (2023).

[53] A. V. Chechkin, R. Metzler, V. Y. Gonchar, J. Klafter, and L. V. Tanatarov, First passage and arrival time densities for Lévy flights and the failure of the method of images, J. Phys. A: Math. Gen. **36**, L537 (2003).

[54] V. V. Palyulin, A. V. Chechkin, and R. Metzler, Lévy flights do not always optimize random blind search for sparse targets, Proc. Natl. Acad. Sci. USA **111**, 2931 (2014).

[55] Since the edge case $r^* = 0$ is unsolvable if $\lambda_j = 0$, we remove $\lambda_1 = 0$ from the sum.

[56] E. Muravyova, A. Golovnin, E. Gracheva, A. Parshikov, T. Belenkaya, V. Pirrotta, and P. Georgiev, Loss of insulator activity by paired Su (Hw) chromatin insulators, Science **291**, 495 (2001).

[57] H. N. Cai and P. Shen, Effects of cis arrangement of chromatin insulators on enhancer-blocking activity, Science **291**, 493 (2001).

[58] O. Kyrchanova, D. Chetverina, O. Maksimenko, A. Kullyev, and P. Georgiev, Orientation-dependent interaction between *Drosophila* insulators is a property of this class of regulatory elements, Nucleic Acids Res. **36**, 7019 (2008).

[59] O. Pulkkinen and R. Metzler, Distance matters: The impact of gene proximity in bacterial gene regulation, Phys. Rev. Lett. **110**, 198101 (2013).

[60] A. Godec and R. Metzler, Universal proximity effect in target search kinetics in the few-encounter limit, Phys. Rev. X **6**, 041037 (2016).

[61] J. E. Moore, M. J. Purcaro, H. E. Pratt, C. B. Epstein, N. Shoresh, J. Adrian, T. Kawli, C. A. Davis, A. Dobin *et al.*, Expanded encyclopaedias of DNA elements in the human and mouse genomes, Nature (London) **583**, 699 (2020).

[62] D. Bernenko, S. H. Lee, P. Stenberg, and L. Lizana, Mapping the semi-nested community structure of 3D chromosome contact networks, PLoS Comput. Biol. **19**, e1011185 (2023).

[63] W. J. Kent, C. W. Sugnet, T. S. Furey, K. M. Roskin, T. H. Pringle, A. M. Zahler, and D. Haussler, The human genome browser at UCSC, Genome Res. **12**, 996 (2002).

[64] M. Sheinman, O. Bénichou, Y. Kafri, and R. Voituriez, Classes of fast and specific search mechanisms for proteins on DNA, Rep. Prog. Phys. **75**, 026601 (2012).

[65] J. Bezanson, A. Edelman, S. Karpinski, and V. B. Shah, Julia: A fresh approach to numerical computing, SIAM Rev. **59**, 65 (2017).

[66] D. T. Gillespie, A general method for numerically simulating the stochastic time evolution of coupled chemical reactions, J. Comput. Phys. **22**, 403 (1976).

[67] E. Lieberman-Aiden, N. L. Van Berkum, L. Williams, M. Imakaev, T. Ragoczy, A. Telling, I. Amit, B. R. Lajoie, P. J. Sabo, M. O. Dorschner *et al.*, Comprehensive mapping of long-range interactions reveals folding principles of the human genome, Science **326**, 289 (2009).