

Biophysical Journal, Volume 121

Supplemental information

Chromatin dynamics controls epigenetic domain formation

Marina Katava, Guang Shi, and D. Thirumalai

Supporting Information: Chromatin dynamics controls
epigenetic domain formation

Marina Katava, Guang Shi, D. Thirumalai

1 Energy Function

Energy function: The energy of the chromatin, U_T , is taken to be a sum of bond stretch (U_B), bond angle (U_{KP}), and interactions (U_{LJ}) between the nucleosomes that are separated by at least 3 bonds.

Bond stretch and bond angle potentials: The connectivity of the chromatin thread is taken into account using a harmonic potential,

$$U_B = \frac{k_s}{2}(r - r_0)^2, \quad (1)$$

where r is the distance between the two consecutive nucleosomes, r_0 is the equilibrium bond length, and k_s is the spring constant. The bond angle is constrained using the Kratky-Porod potential in order to control the stiffness of the chain. We assume that,

$$U_{KP} = \frac{k_B T l_k}{2\sigma} \left[1 - \frac{\mathbf{t}_1 \cdot \mathbf{t}_2}{|\mathbf{t}_1| |\mathbf{t}_2|} \right], \quad (2)$$

where k_B is the Boltzmann constant, T is the temperature, $l_k/2$ is the intrinsic persistence length (l_p), and σ is the effective inter nucleosome distance. The length unit is σ . The variables \mathbf{t}_1 and \mathbf{t}_2 are bond vectors connecting nucleosomes $(i, i + 1)$ and $(i + 1, i + 2)$, respectively.

Non-bonded potential: We used the Lennard-Jones (LJ) potential,

$$U_{LJ} = 4\epsilon \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^6 \right], \quad (3)$$

to model interactions between non-bonded loci. In the above equation, r is the distance between the nucleosomes, σ is roughly the size of the nucleosome, and ϵ , which sets the energy scale, is the strength of the interactions. The LJ interaction is truncated at 3σ ($\epsilon = 0$ for $r > 3\sigma$).

2 Modification Probabilities

The modification of the nucleosomes by enzymes is modeled using a two state kinetics,



where k^+ and k^- are the forward and backward reaction rates, respectively. The solutions to the rate equations,

$$\begin{aligned} -\frac{d[U]}{dt} &= k^+[U] - k^-[M] \\ -\frac{d[M]}{dt} &= -k^+[U] + k^-[M]. \end{aligned} \quad (5)$$

are given by,

$$\begin{aligned} P^+(t) &= 1 - \frac{k^- + k^+ e^{-\lambda t}}{\lambda} \\ P^-(t) &= 1 - \frac{k^+ + k^- e^{-\lambda t}}{\lambda}, \end{aligned} \quad (6)$$

where $\lambda = k^+ + k^-$ and t is time. By defining $r = k^+/k^-$ and $\beta = k^+ \tau_r$, where τ_r is the characteristic time scale, the modification probabilities become,

$$P^+(\tau_r) = 1 - \frac{1 + \alpha r e^{-\beta(1+\alpha r)/\alpha r}}{1 + \alpha r}; P^-(\tau_r) = 1 - \frac{\alpha r + e^{-\beta(1+\alpha r)/\alpha r}}{1 + \alpha r}, \quad (7)$$

where $\alpha = 1$ for nucleation site. The relevant parameters that will control spreading are r , α , and β . Note that the probabilities given in Eq.7 reduce to $P^+ = k^+ \tau_r$ and $P^- = k^- \tau_r$ when $\tau_r \rightarrow 0$. When τ_r is not small, Eq.7 give the correct values for transition probabilities.

Finally, the probabilities of the modification and un-modification of each nucleosome are computed based on Eq.7 with the following assumptions: (i) spreading from the bonded neighbor and non-bonded neighbors are independent and contribute equally to the transition (ii) each neighbor nucleosome contributes to the transition independently. Additional details on the computation of spreading probability are given in Figure S1 and Methods in the main text. We assume that epigenetic spreading occurs through a set of enzyme reactions, which presumably involve physical proximity between the enzymes and the nucleosomes. Due to the possible crowding effect would restrict the available space around the nucleosome including those that are nearby. Therefore, we assume that the 1D and 3D spreading cannot occur at the same time, and maybe treated as independent events. For this reason, we used the total transition probability that is given in Figure (S1b).

3 Langevin dynamics

| Parameter | Description | Value |
|-------------------|---|------------------------------|
| σ (Eq 3) | unit of length | 1 |
| ϵ (Eq 3) | strength of LJ interaction | $0.1k_B T$ |
| d_0 | equilibrium bond distance | 1.0σ |
| k_s (Eq 1) | spring constant between connected loci | $\frac{3000\epsilon}{d_0^2}$ |
| l_k (Eq 2) | double the chromatin persistence length | 2σ |
| a^* | value of σ | $27nm$ |

Table S1: Parameters used in the Langevin dynamics simulations. * a is roughly the sum of the size of the nucleosome ($10nm$) and the contour length of the DNA ($\simeq 50bp \cdot 0.33nm/bp \approx 17nm$). Equation numbers in the first column refer to the main text.

We describe the time evolution of the chromatin by the Langevin equation,

$$m_i \frac{d^2 \mathbf{r}_i}{dt^2} = \mathbf{F}_i - \xi \frac{d\mathbf{r}_i}{dt} + \mathbf{R}_i(t), \quad (8)$$

where \mathbf{r}_i is the position of the i^{th} locus whose mass is m_i , $\mathbf{F}_i = -\frac{\partial U_T}{\partial \mathbf{r}_i}$ is the systematic force arising from U_T (see the main text), ξ is the friction coefficient, and \mathbf{R}_i is the random force that satisfies

(a)

| | 1D forward | 3D forward | 1D reverse | 3D reverse | Noise |
|-------------------|---------------|---------------|---------------|---------------|----------|
| | $P_{1D}^+(i)$ | $P_{3D}^+(i)$ | $P_{1D}^-(i)$ | $P_{3D}^-(i)$ | $P_N(i)$ |
| I: 1D+F+R | Y | 0 | Y | 0 | Y |
| II: 3D+F+R | Y | Y | Y | Y | Y |

(b)

$$\text{i. } P^+(i) = P_{1D}^+(i) + P_{3D}^+(i) - P_{1D}^+(i)P_{3D}^+(i)$$

$$\text{ii. } P^-(i) = P_N(i) + (1 - P_N(i))(P_{1D}^-(i) + P_{3D}^-(i) - P_{1D}^-(i)P_{3D}^-(i))$$

$$\text{iii. } P_{1D}^\pm = 1 - [1 - P_\pm(t = \gamma\tau_r)]^{n_\pm^{1D}} \quad P_{3D}^\pm = 1 - [1 - P_\pm(t = \tau_r)]^{n_\pm^{3D}} \quad P_N(i) = P_-(t = \gamma\tau_r)$$

$$n_\pm^{1D} = \sum_{j \in \{i-1, i+1\}} \delta_{s_j, \pm 1} \quad n_\pm^{3D} = \sum_{|i-j| > 2} \Theta(r_c - d_{ij}) \delta_{s_j, \pm 1}$$

Figure S1: **(a)** Columns represent the elementary reactions and the rows show the schemes implemented in the simulations. The five elementary reactions are: 1D forward, 1D reverse, 3D forward, 3D reverse, and the noise. The noise process, with the associated probability $P_N(i)$ where i is the nucleosome label, affects only the modified nucleosomes. The two simulation schemes are: **I** (1D+F+R, (1D forward, 1D reverse, and the noise)), and **II** (3D+F+R, (1D forward, 1D reverse, 3D forward, 3D reverse, and the noise)). Y indicates that the probabilities for the five elementary reactions are computed using equation **iii** in **(b)**, and 0 means that the corresponding probability is zero. For each simulation scheme, the cumulative probability for the forward ($P^+(i)$) and reverse ($P^-(i)$) reactions is computed using Equations **(b)i** and **(b)ii**. For scheme **I** (1D+F+R), the probabilities $P_{3D}^+(i)$ and $P_{3D}^-(i)$ are zero, resulting in $P^+(i) = P_{1D}^+(i)$ and $P^-(i) = P_N(i) + (1 - P_N(i))P_{1D}^-(i)$, where $P_{1D}^\pm(i)$ and $P_N(i)$ are computed using Equations **(b)iii**. The probability for scheme **II**, which includes all the elementary reactions, is given by $P^+(i)$ (Equation **(b)i**) and $P^-(i)$ (Equation **(b)ii**). **(b)** Equations for all the relevant probabilities. (i) and (ii) are the cumulative probability for forward and reverse reaction. (iii) Displays the probabilities for the five elementary reactions. $t = \gamma\tau_r$ ($t = \tau_r$) is the time at which 1D and noise (3D) spreading is considered. The variables n_\pm^{1D} (n_\pm^{3D}) are the number of bonded neighbors (non-bonded neighbors) whose state is U (-1) or M (+1) to i^{th} nucleosome. Non-bonded neighbors (j s) satisfy the criterion $|i - j| \geq 2$ and $r_{ij} < r_c = 1.122\sigma$.

the fluctuation-dissipation theorem. An in-house code was developed to integrate the Langevin equation, with $T = 300K$, using the velocity-Verlet algorithm [1]. After equilibrating the polymer for times that exceed the relaxation time ($\tau_{R_{ee}}$) of the end-to-end vector of the polymer, we performed long simulations ($\gg 10\tau_{R_{ee}}$) so that reliable statistics for computing various quantities

of interest could be generated.

4 Epigenetic modification algorithm flowchart

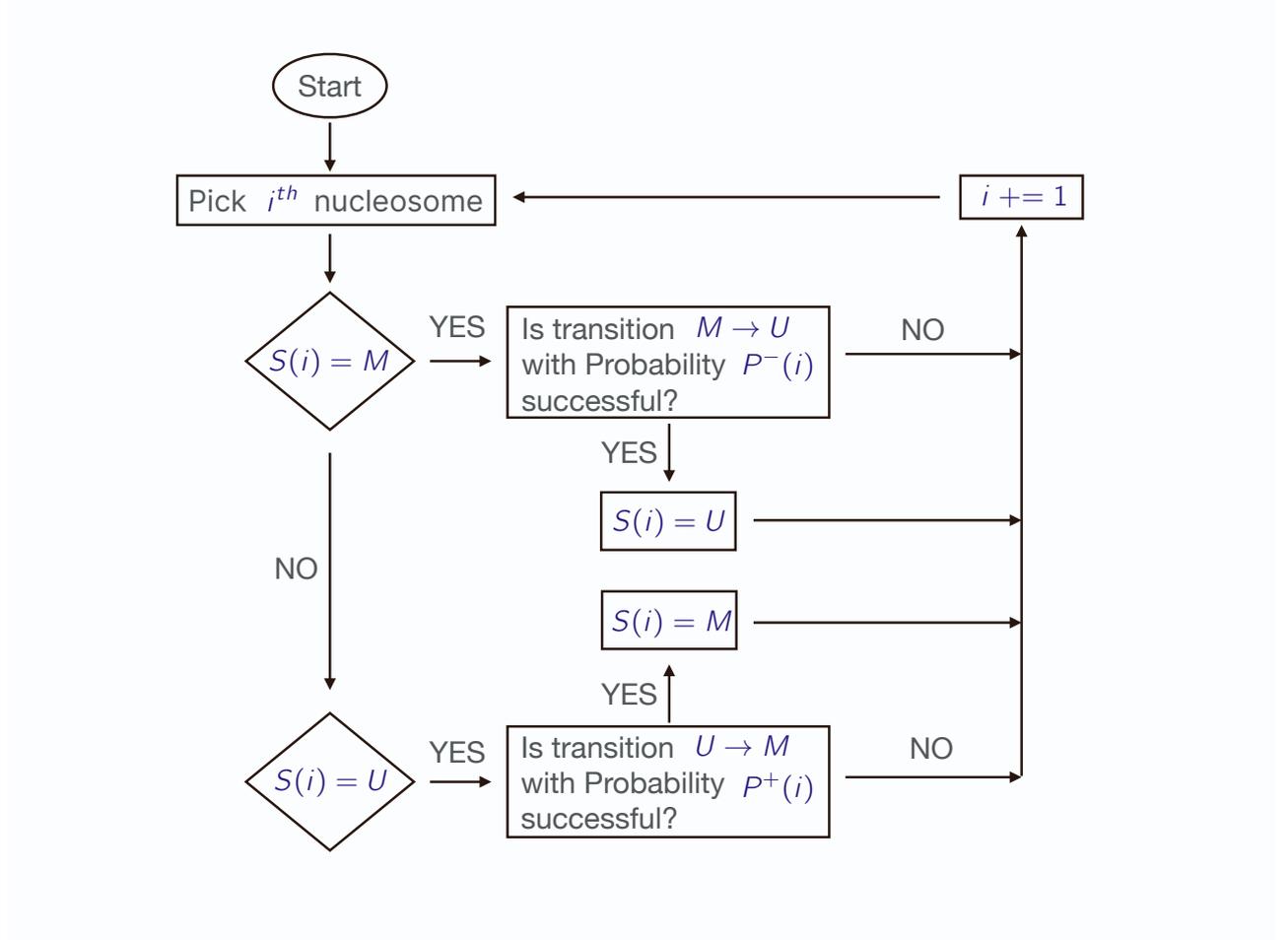


Figure S2: Flowchart of the algorithm for epigenetic modifications at each time step. The probability values are given in Figure S1.

5 Structural relaxation time

We calculated the structural relaxation time, τ_r , from the decay of the structure factor,

$$F(q, t) = \frac{1}{N} \left\langle \sum_{j=1}^N \exp^{-iq \cdot (r_j(t) - r_j(0))} \right\rangle, \quad (9)$$

with $q = \frac{2\pi}{r_c}$. In the above equation, $r_j(t) - r_j(0)$ is the displacement of nucleosome r_j . From the decay of $F(q, t)$ (Figure S3) we estimated the characteristic time, τ_r . The time τ_r , which is an estimate for the relaxation of the chromatin polymer, is assumed to set the over all time scale. All other rates that are relevant to epigenetic spreading are set relative to τ_r . It is natural to use τ_r , especially for 3D spreading to monitor the modification process. Note that τ_r is a function of N , l_p , as well as solvent quality.

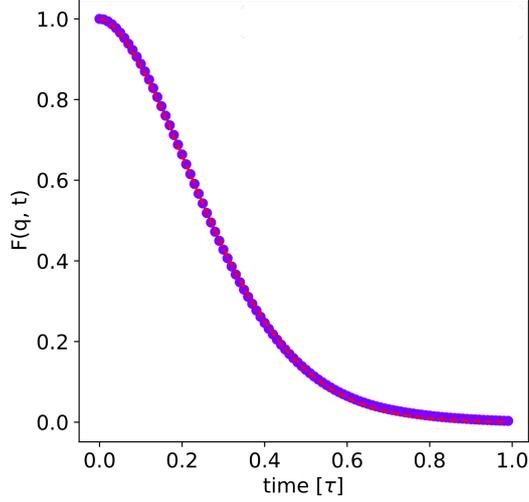


Figure S3: Intermediate scattering function, evaluated at the wave vector $q = \frac{2\pi}{r_c}$ ($r_c = 1.122\sigma$), as a function of time. $F(q, t)$ was averaged over 15 independent trajectories by removing the overall rotational and translational degrees of freedom. The value of τ_r is extracted using the fit, $F(q, t) = \exp^{-(t/\tau_r)^\beta}$ with $\beta=1.7$, yielding $\tau_r \approx 0.3\tau$ where τ is the natural time governing Eq. 8. For our purposes the precise value of the τ_r is irrelevant. What matters are the spreading probabilities given in Figure S1.

6 Solvent quality

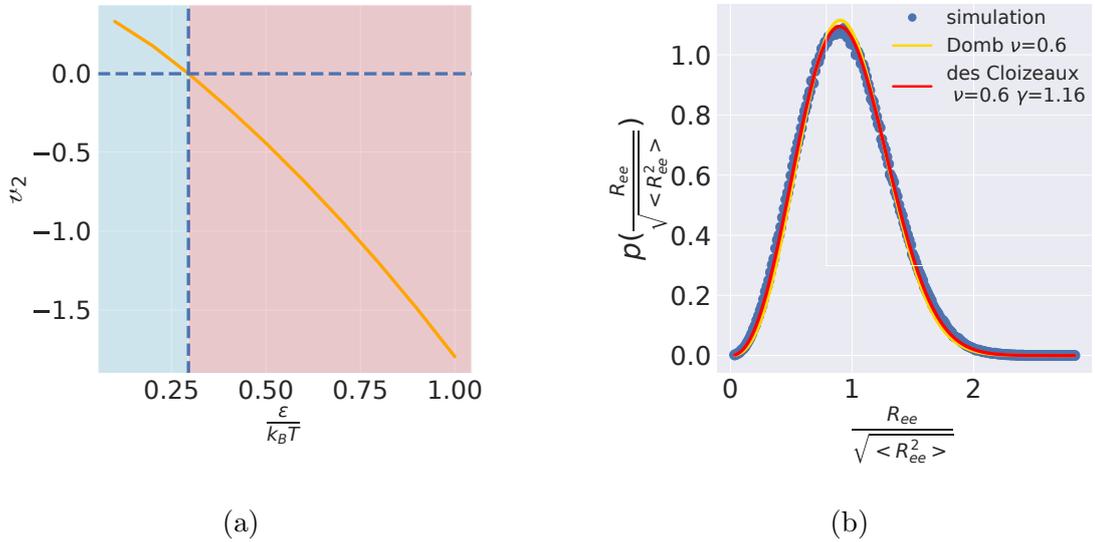


Figure S4: (a) Second virial coefficient as a function of the Lennard-Jones parameter ϵ , characterizing the interaction between the nucleosomes. The blue (red) region corresponds to “good solvent” (“poor solvent”). (b) Comparison of the end-to-end distance distributions computed from simulations ($\epsilon = 0.1k_B T$) with the rigorous theoretical prediction for a polymer in a good solvent [2].

The non-bonded interactions are determined by the Lennard-Jones interaction strength, ϵ (Eq. 3 in the SI), between the nucleosomes that are separated by at least two bonds from each other.

The ϵ parameter is an effective interaction strength between the loci. The dimension of the chain depends on the second virial coefficient,

$$v_2 = 2\pi \int_0^\infty r^2 dr [1 - \exp^{-\beta U_{LJ}(r)}], \quad (10)$$

where $U_{LJ}(r)$ is given in Eq (3) in the main text, $\beta = 1/k_B T$. If $v_2 > 0 (< 0)$, then the chromatin could be extended (random coil). In Figure S4, we show v_2 as a function of $\beta\epsilon$. The θ -point at which $v_2 = 0$ corresponds to $\epsilon = 0.3k_B T$ (Figure S4). We varied ϵ in our simulations to assess the impact of solvent quality on the nature of epigenetic spreading. For $\epsilon = 0.1k_B T$, which is in the good solvent region, chromatin behaves as a Flory random coil. We expect that the distribution $P(x)$, with $x = R_{ee}/\langle R_{ee}^2 \rangle^{1/2}$, should follow the universal behavior with $P(x) \sim x^\delta \exp^{-x^\delta}$, where $\delta \approx 1/(1 - \nu)$, where ν is the Flory exponent (≈ 0.6 in 3D). The excellent agreement between theory and simulation confirms that the chromatin polymer is indeed a random coil. We also show results for $\epsilon \geq 0.3k_B T$, see Figure ?? in main text.

7 Contact times and τ_r

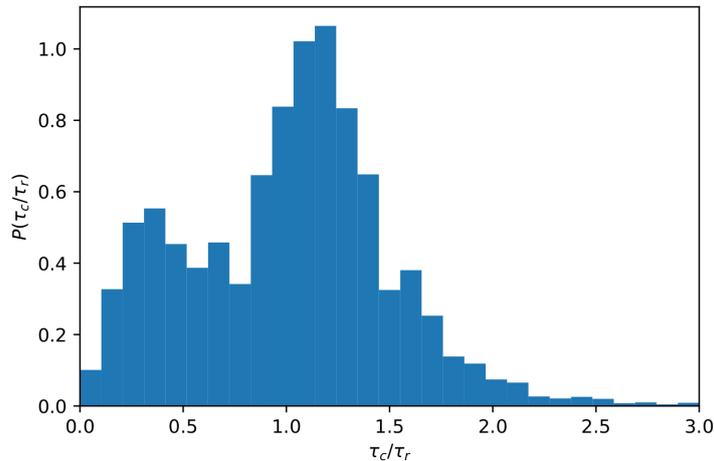


Figure S5: Distribution of contact duration times τ_c . A contact forms if two nucleosomes are within $r_c = 1.122\sigma$. The mean value is $\langle \tau_c \rangle = 0.84\tau_r$.

There are three time scales that characterize our epigenetic polymer model (i) The rate of the forward reaction rate k^+ (Eq 4 in the main text), (ii) The second is k^- , the backward reaction rate. Both k^+ and k^- are defined in the main text. (iii) The third is, τ_r , the chromatin relaxation time. Below we describe how these timescales are chosen, and assigned physical meaning.

As shown in Figure ??, epigenetic spreading could occur either linearly (1D) (i and $i \pm 1$) or by non-bonded nucleosomes that come into proximity through loop formation (3D). The looping time could be substantial, making the modification probability through the 3D mechanism less efficient than by 1D. To account for the separation in time scales, we allow for 1D modifications to occur on time scale $t = \gamma\tau_r$. In other words, the probabilities of modification through 1D mechanism are computed using equations in Figure S1 (b). Spreading in 3D occurs if two loci,

separated by at least 2 bonds, come into contact. There is a spectrum of looping times that depend on the separation $|i - j|$ between the nucleosomes. A relevant time for modification is the contact life time (τ_c) during which 3D spreading could occur. The distribution of $P(\tau_c/\tau_r)$ contact life times, expressed in units of τ_r , (Figure S5) shows that the average $\langle \tau_c \rangle \approx 0.84\tau_r$. To simplify the computations, we calculated the probabilities of 3D spreading using equations shown in Figure S1 (b). The use of $\langle \tau_c \rangle$ as a proxy for looping, independent of the genomic separation between the loci, simplifies the computations without qualitatively altering the results. We assume that spreading is slow, which implies that the polymer relaxes much faster than the time scale in which the enzyme reaction ($\frac{1}{k^+}$) occurs. This is implemented by choosing $k^+ = \frac{0.01}{\tau_r}$.

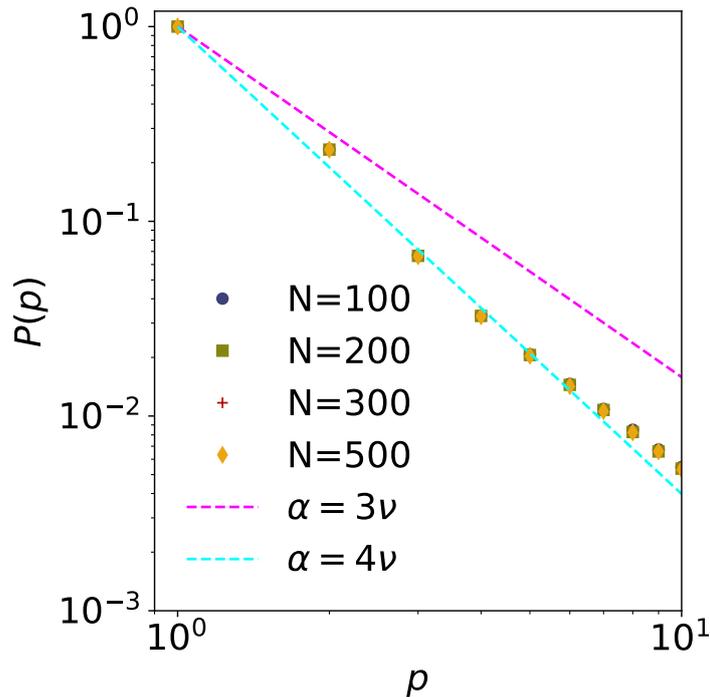


Figure S6: Dependence of the contact probability $P(p)$ as a function of the number of bonds that separate two nucleosomes, p . The dashed lines represent theoretical scaling $P(p) \sim p^{-\alpha}$, where α is a multiplicative value of the Flory exponent $\nu = 3/5$.

8 Epigenetic ergodicity

In order to assess if the time scales governing spreading result in a non-equilibrium state, we introduce the epigenetic ergodicity measure by generalizing a measure introduced in the context of glasses [3]. Multiple pairs of independent simulations were used to calculate the epigenetic ergodicity, defined as,

$$d(t) = \frac{1}{N} \sum_{j=1}^N [s_{a,j}(t) - s_{b,j}(t)]^2, \quad (11)$$

where $s_{a,j}(t)$ is the time-average of the epigenetic state of nucleosome j in trajectory a over time t , $s_{b,j}(t)$ is the corresponding quantity in trajectory b . If the system is ergodic, each trajectory

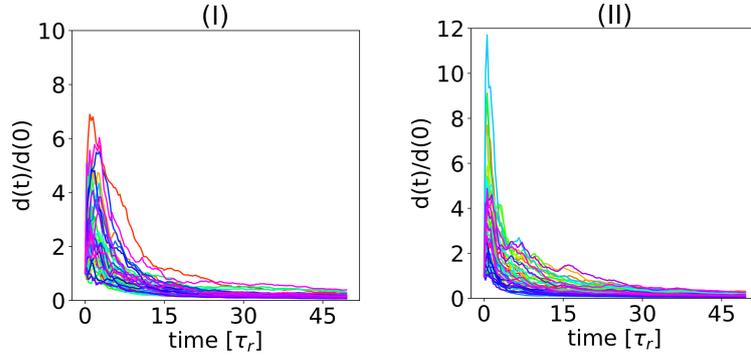


Figure S7: Epigenetic spreading is ergodic. We compared 10 independent trajectories for both biological mechanisms. The epigenetic ergodic measure vanishes at long times. The parameter values in simulations are $k^+/k^- = 50$, $k^+ = 100/\tau_r$, with total simulation time $15,000\tau_r$.

would explore the entire epigenetic phase space, and the time-average for both the trajectories should converge at long times. Thus, in an ergodic, or quasi ergodic system, the quantity $d(t)$ should vanish at long t . Figure S7 shows that $d(t)$ vanishes at long times for all the spreading processes on time scales that are far less than the time needed for steady-state spreading to be established. Thus, for the range of time scales considered here, and for $N = 300$, epigenetic ergodicity is established. It is conceivable that epigenetic ergodicity could be broken, resulting in glass-like epigenetic states for different N , and solvent quality governing the spreading dynamics. It is unclear if this could confer any biological advantage for epigenetic memory.

9 Influence of chromatin persistence length on epigenetic spreading

We anticipate two extreme scenarios for the influence of the persistence length, l_p , on the spreading mechanisms. If $L/l_p \gg 1$ (L is the contour length of the polymer) then 3D transient loop-driven spreading, as found here and elsewhere [4, 5], is likely. In the opposite stiff chain limit $L/l_p \ll 1$, we expect that spreading would be predominantly determined by the 1D mechanism. To verify these expectations, we performed simulations by varying the bending stiffness (l_k) in the bond angle potential. For different values of l_k , we extracted the persistence length using,

$$\langle \cos\theta(p) \rangle = \exp^{-\frac{p}{l_p}}, \quad (12)$$

where θ is the angle between two bond vectors separated by a distance p along the contour of the polymer. The exponential fit to simulations, with l_k ranging from $(2 - 200)\sigma$, is shown in Figure S8. A crossover between the flexible chain ($L \gg l_p$) to a rigid chain is expected when $L \ll l_p$. Figure S8 shows that the best fit is obtained when l_p spans several bond distances, while the rigid and flexible limit exhibit stronger deviations from the nominal estimate, $l_p = l_k/2$. The exponential decay $\langle \cos\theta(p) \rangle$ (Eq. (12)), is valid strictly for a polymer without excluded volume interactions, and the deviations of the exponential scaling have previously been reported for large

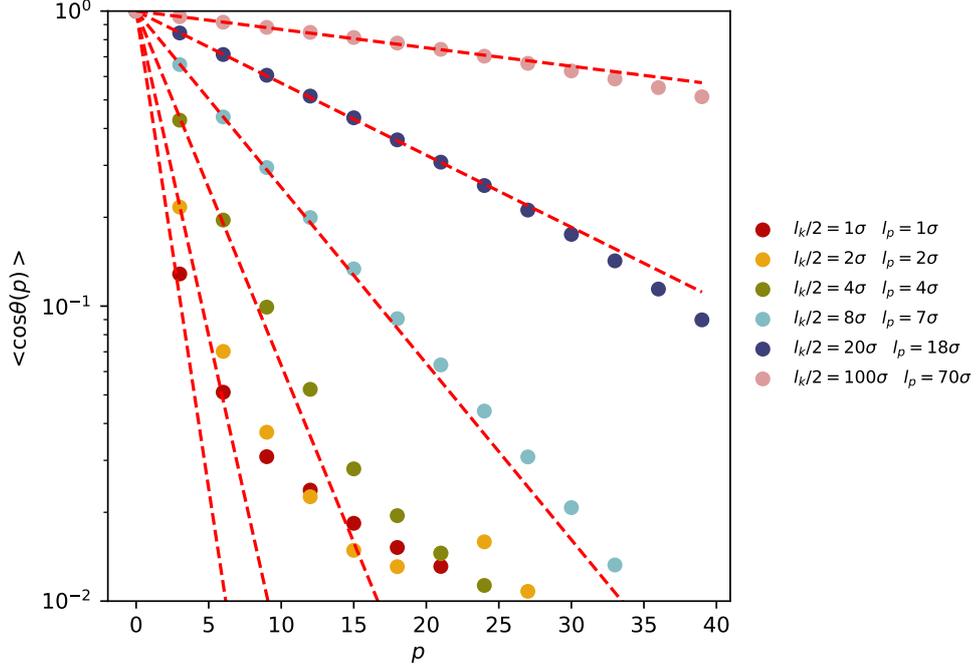


Figure S8: Bond vector correlation function $\langle \cos\theta(p) \rangle$ as a function of the bond distance p along the polymer for different intrinsic stiffness values. Red dashed lines are fits to the curves. The values of the persistence length l_p (listed on the right) are extracted using Eq 12. The fits are increasingly more accurate as l_p increases.

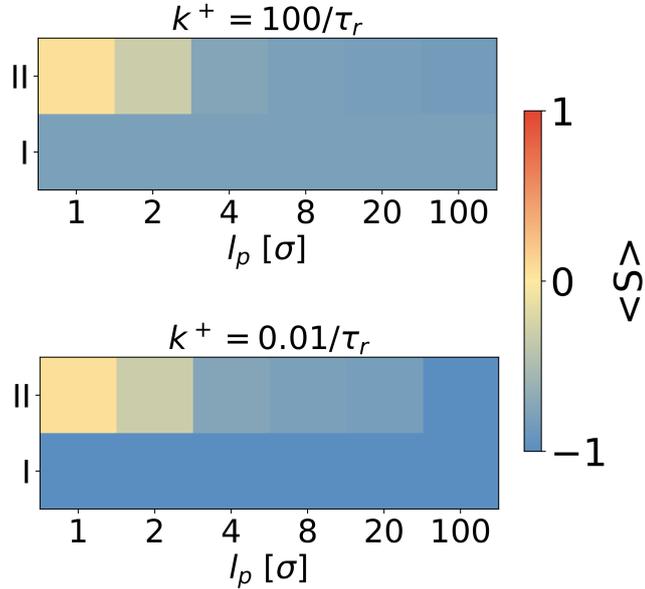


Figure S9: Heat maps showing the average value of the epigenetic order parameter $\langle S \rangle = \frac{n_m - n_u}{N}$ as a function of the persistence length. It is clear that the 3D result converges to the 1D result as the persistence length increases.

p when excluded volume conditions are taken into account [6, 7].

We characterize the global epigenetic state using $\langle S \rangle$ for two values of k^+ as a function of l_p . As anticipated, the 3D model converges to the 1D limit (Figure S9) once the chain stiffness

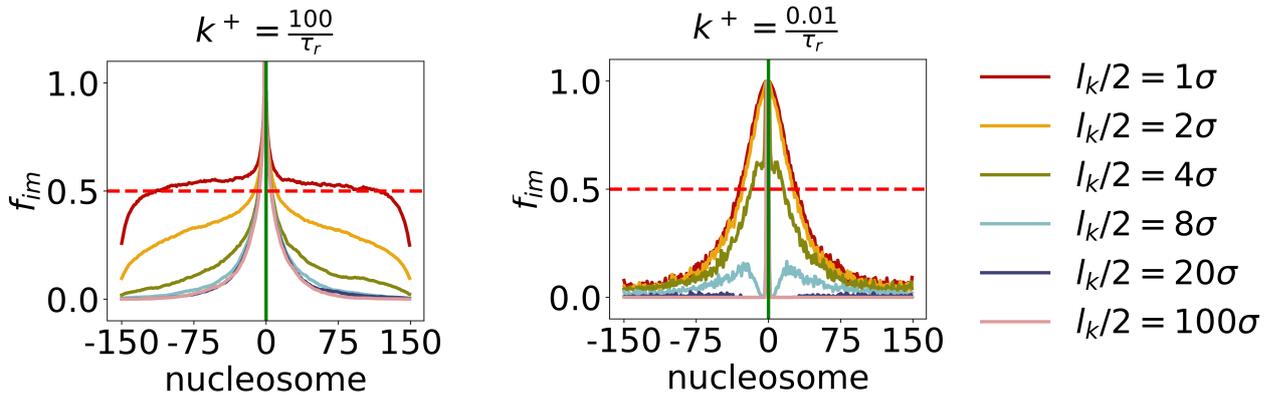


Figure S10: The effect of persistence length, $l_p = l_k/2$, on spreading for Scheme II depicted as the fraction f_{im} of modified state of nucleosome i . Scheme II is tested under fast, $k^+ = \frac{100}{\tau_r}$, and slow, $k^+ = \frac{0.01}{\tau_r}$, spreading conditions. We used $k^+/k^- = 50$ for fast spreading, and $k^+/k^- = 10,000$ for slow spreading, which are the same values used in the SI and the main text.

increases substantially ($\frac{l_p}{\sigma} \geq 8$). For instance, when spreading is fast ($k^+ = 100/\tau_r$) or slow ($k^+ = 0.01/\tau_r$), the value of $\langle S \rangle$ for I and II mechanisms are roughly the same at $\frac{l_p}{\sigma} = 8$, as shown in Figure S9. Thus, regardless of the enzyme rates for modifying a nucleosome, the 1D and 3D mechanisms converge when l_p is sufficiently large. The reason is that in the stiff chain limit, the energy penalty to bend the polymer is high, thus effectively preventing the formation of looping contacts, which is needed for 3D spreading.

Upon closer observation of 3D spreading, the simulated domains in Figure S10 reveal that enhanced flexibility improves the propensity for stable domain formation. In particular, we observe a stabilizing effect of enhanced flexibility on the epigenetic pattern around the nucleation site. These results show that there is an interplay between 3D and 1D spreading, which is determined by L/l_p .

10 Finite size effects

Polymer length N affects relaxation timescales, which in turn changes polymer looping kinetics [8], with possible effects on the 3D spreading mechanism. We assessed the effects of changing N from (100-1000) for the fast-spreading mechanism. In the slow-spreading regime, the 1D spreading in Scheme I (left panel in Fig S11) shows very little local spreading due to low spreading probability at each time step. However, Fig S11 Scheme II reveals that the inactivation domain profiles behave identically at equal distances from the nucleation site, independent of the length of the polymer. This is because the nucleation site is a major contributor to domain formation, and 3D contact formation of the nucleation site with other residues determines the domain shape. The fraction of modification decreases with genomic distance from the nucleation site, underlined by a lower probability of contact formation at larger distances (Figure S6). Thus, similar domain shapes for chains of different lengths reflect a given probability contact scaling with genomic distance, which is constant in these simulations.

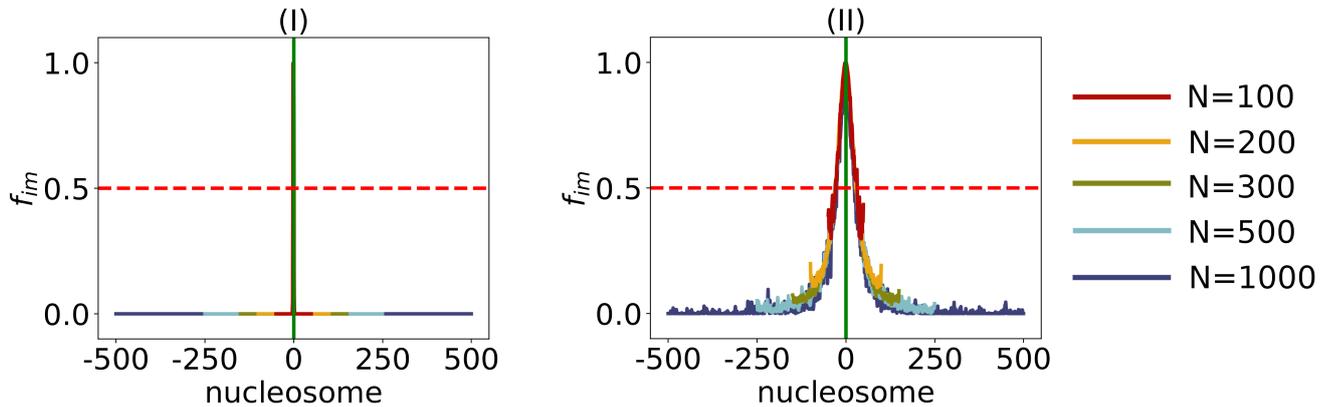


Figure S11: Tests for biological models as a function of locus length, N , with parameters $k^+ = \frac{0.01}{\tau_r}$, $k^+/k^- = 10,000$. The left (right) panel shows modification obtained using Scheme **I** (Scheme **II**). The modification probabilities for the Schemes are given in Figure S1.

10.1 The epigenetic switch

The figure Figure S12 shows that the interplay between the three parameters, k^+ , k^- , and α with $\gamma = 0.03$, determines the global epigenetic state. For spreading events driven by the nucleation site, $\alpha = 1$, while for those triggered by modified nucleosomes (M_s), $\alpha < 1$. We consider the fast spreading case here in order to test parameter sensitivity, as it converges faster, and a wider range of parameters could be tested in a reasonable time. In this case, the spreading rate (k^+) is much faster than the polymer relaxation rate (τ_r^{-1}), $k^+ = 100/\tau_r$. The k^- value is a fraction of the forward rate, which means the reversal of M occurs at a slower rate compared to modification of the U state. We follow the modification status of all the nucleosomes over the course of the simulation time in order to assess if the chromatin is predominantly in the global active (unmarked with $\langle S \rangle \approx -1$) or inactive (marked with $\langle S \rangle \approx 1$) state.

Figure S12 (A) shows the global epigenetic state $\langle S \rangle$ as an average over 10 trajectories with different initial states, as a function of the rescaling parameter α , which modifies the forward rate distally of nucleation site (see main text). A range of α and k^+/k^- are probed, and the global epigenetic state $\langle S \rangle$ is computed for each parameter set. The data shown in subfigure (A) is then rearranged and shown in subfigure (B) as follows: (i) For a given k^+/k^- , α values are rescaled to $\alpha k^+/k^-$. This yields a different set of $\alpha k^+/k^-$ values for each k^+/k^- . (ii) With the new scale, $\alpha k^+/k^-$, only $\langle S \rangle$ data that falls within the range of $\alpha k^+/k^- \in [0.5, 3.0]$ is taken into account. For each k^+/k^- ratio separately, the data is divided into bins with width 0.5. (iii) $\langle S \rangle$ values are associated to $\alpha k^+/k^-$ bins. If multiple $\langle S \rangle$ values are in the same, their average is computed. (iv) Each $\alpha k^+/k^-$ bin is associated with a single $\langle S \rangle$ value, as shown Figure S12 (B).

Plots of the global epigenetic state and the nucleosome-dependent state in Figure S12 (B) allow us to draw a few pertinent conclusions. (i) Modifications are highly improbable if k^+/k^- is less than a certain value (≈ 10) regardless of the value of α . The chromatin remains in the global U state with $\langle S \rangle \approx -1$, regardless of whether the modifications occurs by Scheme **I** or Scheme **II** (see the left panels in Figure S12). In general, αk^+ has to exceed k^- for spreading

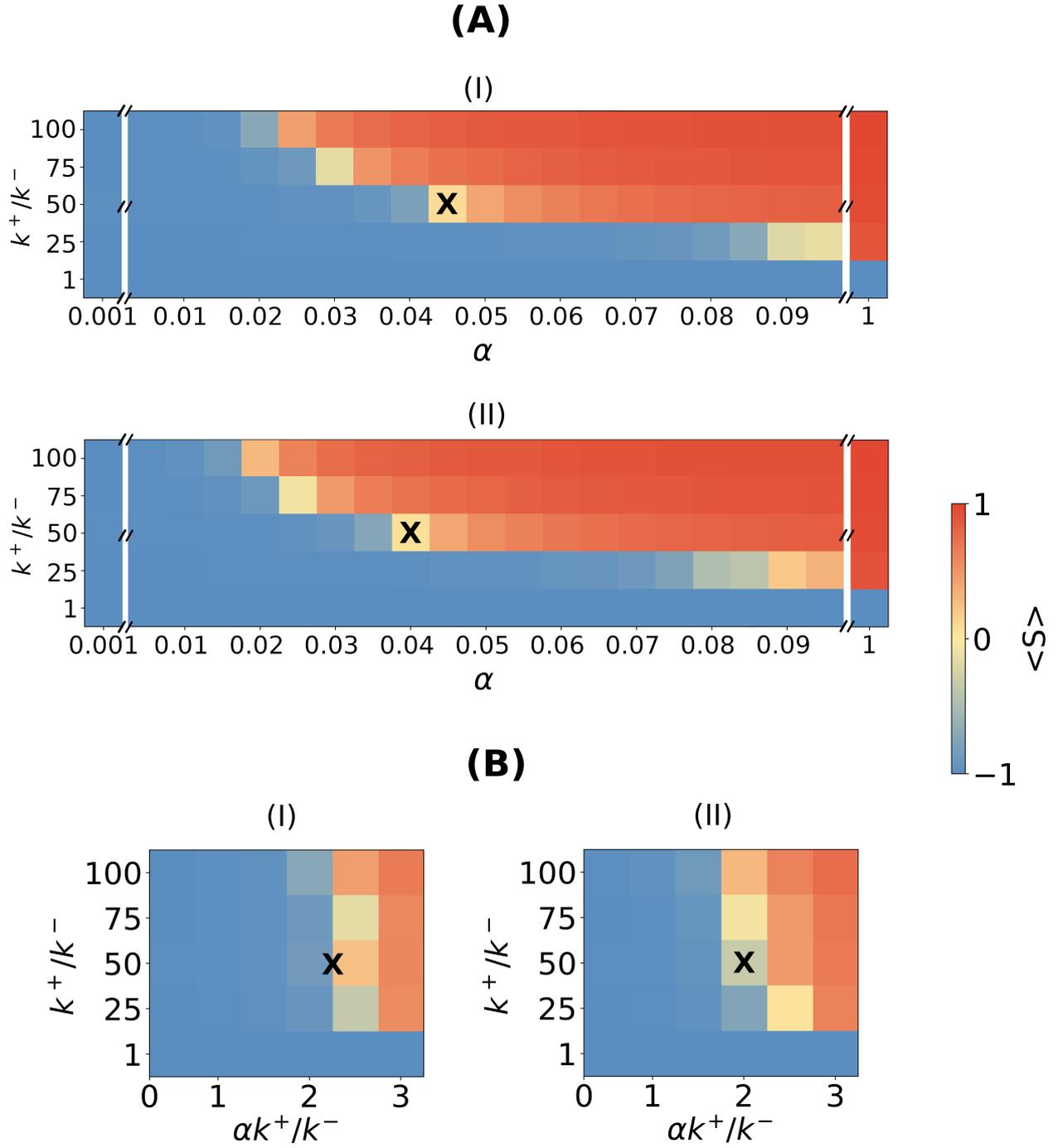


Figure S12: Epigenetic switching depends on modulation of the forward vs back rate distally from the nucleation site. Heatmaps show the overall epigenetic state, characterized by $\langle S \rangle$. Scheme **I** and Scheme **II** parameters used in main text are marked by **X**. (A) Epigenetic switching shown as a function of α , quantified using the mean value, $\langle S \rangle$. (B) The global epigenetic state is a delicate balance between forward and backward rates, and the domain acts as an epigenetic switch (see the main text). The parameters are chosen to ensure that the global average $\langle S \rangle \approx 0$. In all the panels, $k^+ = 100/\tau_r$ (fast spreading), the total length of the simulations is $15,000\tau_r$, and $k^- = k^+/50$.

to occur. At higher values of k^+/k^- there is a global $U \rightarrow M$ transition as α is increased. (ii) The transition from a predominantly active to predominantly inactive state occurs over a narrow range of k^+/k^- , and α . For instance, for both mechanisms, **I** and **II**, the switch from $\langle S \rangle < 0$ to $\langle S \rangle > 0$ occurs over a narrow range of α , as is evident from Figure S12 (B). The critical point in the switch from active to inactive state is determined by k^+/k^- . The transition hinges on the competition between the forward (favors spreading) and the reverse reaction (inhibits spreading) distally from the NS, and is controlled by $\alpha k^+ \approx k^-$. Under this condition, the modification results in the concurrent formation of similar-sized patches containing U and M nucleosomes, resulting in $\langle S \rangle \approx 0$. This is most evident in spreading by mechanism **II** (Figure S12 (B)), showing a switch from active state (yellow color, $\langle S \rangle < 0$) to the inactive state (red color, $\langle S \rangle > 0$) through the mixed state (white squares, $\langle S \rangle \approx 0$). The delicate state of the domain, depending on a narrow $\alpha(k^+/k^-)$ range produces an epigenetic switch, reminiscent of gene control expression in cells, whereby cellular fate depends on the relative concentrations of the epigenetic regulators. Shifting the balance in one direction or the other could result in dramatically different outcomes in terms of gene expression patterns [9]. (iii) If $\alpha k^+/k^- < 1$ then the reverse reaction occurs predominantly by random histone turnover (‘noise’ in the list of transitions) that does not depend on the epigenetic identity of neighbors. On the other hand, if $\alpha k^+/k^- > 1$ the reverse reaction is facilitated due to positive feedback, which pushes the domain state towards the global U state. To mitigate this effect, stronger enhancement of the forward reaction is needed to achieve silencing levels comparable to the reverse noise dominated mechanism.

11 NS location

We investigated the effect of changing the location of the NS on the spreading process. Given that mechanistically spreading occurs bi-directionally from the NS along the chromatin chain, we expected similar results if the NS location is changed. The simulations confirm that it is the NS that drives spreading both in Schemes **I** and **II**. The results are similar to the ones in the main text except that the spreading profiles are centered around the NS (Figure S13).

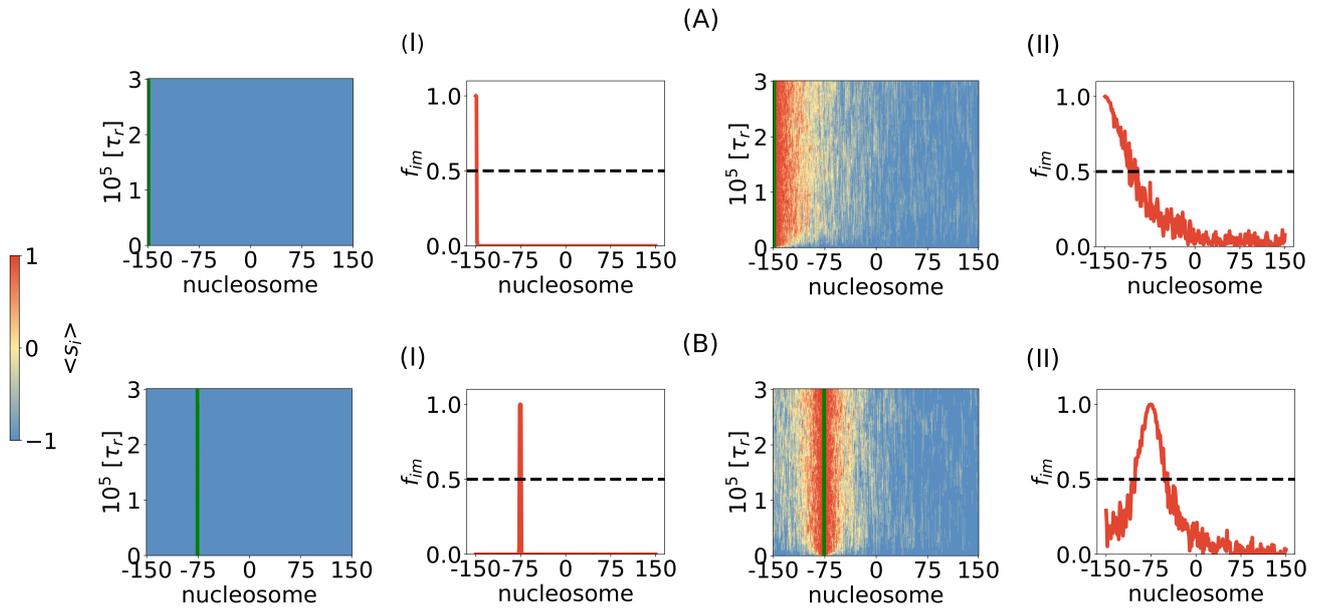


Figure S13: In (A) , the NS is in position $i = -150$, while in (B) it is in position $i = -75$. The relevant parameter values are $k^+ = \frac{0.01}{\tau_r}$, $k^+/k^- = 10,000$.

References

- [1] William C Swope, Hans C Andersen, Peter H Berens, and Kent R Wilson. A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: Application to small water clusters. *The Journal of Chemical Physics*, 76(1):637–649, 1982.
- [2] Michael E Fisher. Shape of a self-avoiding walk or polymer chain. *The Journal of Chemical Physics*, 44(2):616–622, 1966.
- [3] D Thirumalai, Raymond D Mountain, and TR Kirkpatrick. Ergodic behavior in supercooled liquids and in glasses. *Physical Review A*, 39(7):3563, 1989.
- [4] Ian B Dodd, Mille A Micheelsen, Kim Sneppen, and Geneviève Thon. Theoretical analysis of epigenetic cell memory by nucleosome modification. *Cell*, 129(4):813–822, 2007.
- [5] Fabian Erdel and Eric C Greene. Generalized nucleation and looping model for epigenetic memory of histone modifications. *Proceedings of the National Academy of Sciences*, 113(29):E4180–E4189, 2016.
- [6] J P Wittmer, H Meyer, J Baschnagel, A Johner, S Obukhov, L Mattioni, M Müller, and A N Semenov. Long range bond-bond correlations in dense polymer solutions. *Phys. Rev. Lett.*, 93(14):147801, October 2004.
- [7] Hsiao-Ping Hsu, Wolfgang Paul, and Kurt Binder. Polymer chain stiffness versus excluded volume: A monte carlo study of the crossover towards the wormlike chain model. *Euro Phys. Lett.*, page 28003, 2010.
- [8] Ngo Minh Toan, Greg Morrison, Changbong Hyeon, and D Thirumalai. Kinetics of loop formation in polymer chains. *J. Phys. Chem. B*, 112(19):6094–6106, May 2008.
- [9] Rudolf Jaenisch and Adrian Bird. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nature genetics*, 33(3):245–254, 2003.