Chromatin memory generated by cohesin-like interaction at looping sites: modeling an hypothesis

Andrea Papale[1] and David Holcman[1,2] [a]

[1] *Group of Computational Biology and Applied Mathematics, Ecole Normale Supérieure, IBENS, Université PSL, 75005 Paris, France.*

[2] *Churchill College, U. of Cambridge, CB30DS, UK.*

How long could a chromatin loop be stabilized inside the nucleus? Such a time is crucial for chromatin organization maintenance and stability. To answer this question, we explore several physical scenarios, where the loop maintenance is due to diffusing cross-linkers such as cohesin and CTCF that can bind and unbind at the base of chromatin loops. We find that this mechanism is sufficient to guarantee that there are always cross-linkers in place. The number of bound cross-linkers is crucial to create a positive feed-back mechanism that stabilizes the loop over long-time. The three scenarios that we explore here using first passage times of a jump process, differ in the manner cohesin is released. We find that tens to hundreds are sufficient to guarantee the ensemble of loop stability in the genome over the lifetime of a cell. To conclude, modeling and simulations suggest that long-term chromatin memory could depend on the dynamical interaction of cross-linkers at nanoregions formed by the initial interaction with transcription factors.

[a] david.holcman@ens.psl.eu

INTRODUCTION

The human genome is organized into hetero- and eu-chromatin, characterized by a high and low chromatin density [1]. Cell differentiation is often initiated by transcription factors (TFs) that bind to specific activation site [2]. This association can induce long-range chromatin loop when the promoter and enhancer are brought together transiently [3]. Interestingly, there are also proteins such as remodellers that constantly bind to open or close chromatin. If genes are likely to be expressed in open DNA regions, it remains unclear how a region can stay open after TFs leave their binding regions. Our aim here is to propose a model-hypothesis to address this question using jump processes. We recall that statistical analysis of cell nucleus based on high-resolution Hi-C experiments [4] has revealed that chromatin organization is characterized by three dimensional structures called Topologically Associating Domains (TADs). TADs constitute regulatory neighborhoods that isolate genes and their associated regulatory elements from the rest of the genome: the correct establishment of TAD structures prevent the formation of inappropriate gene-enhancer contacts between neighboring domains and TAD disruption can induce carcinogenesis [5, 6]. These TADs are formed by the chromatin itself with highly enriched intra-domain interactions [7]. They can also be seen as a specific ensemble of short and long-range loops. These loops are controlled by nanodomains enriched in specific binders [8, 9] such as CTCF proteins that bind to the majority of borders that demarcate TADs [10]. CTCF binding sites are among the most frequent mutation/deletion hotspots in multiple cancer types and these mutations/deletions at TAD borders can activate oncogenes [6].

In that context, using a Markov modeling approach with absorbing boundary conditions, introduced in [11, 12] (reviewed in [13–15]) to compute the mean time a given number of crosslinkers reach a given threshold [13], we aim to explore how the local transient chromatin organization initially induced by several promoter-enhancer contacts thus making various loops could be stabilized permanently. It is indeed conceivable that such open chromatin regions must be maintained to express specific genes and repress others.

We hypothesize here that due to the local proximity of chromatin induced by TFs and remodelers, this situation facilitates the binding of looping factors such as diffusing cohesin molecules, which can attach and detach DNA dynamically. We thus investigate whether such a dynamical binding and unbinding can lead to a preferential stabilization of a region where promoter-activator have initially generated a transient loop. This situation falls into a positive feedback process, where a loop is stabilized by a region enrich in cohesin. Since cohesins are trapped in this nanoregion, these

molecules have less chance to diffuse away. These molecules enter and leave this nanoregion and their presence maintains loops. This elementary mechanism keeps the cohesin present at the base of a loop, because cohesin that left are replaced by others entering this local region. This dynamical mechanism keeps the initial TF-remodeler chromatin changes and thus stabilizes loops that have been induced initially during a specific gene activation. This mechanism is complementary to the loop extrusion hypothesis [16], where loops are supposed to be generated by an active process, increasing the length until cohesin stops at the boundary made by two CTCF sites.

In the first part of the manuscript, we introduce the feedback model of binding and unbinding cross-linkers to a nanoregion. In the second part, we analyze three possible scenarios of linker stabilization. In the third part, we find that the mechanism of binding preferentially at the center of a looping region by a dynamical equilibrium, can generate a long-term stability, which can explain the long-term memory of genomic organization pre-established by TF binding.

A MARKOV MODEL OF CROSS-LINKER BINDING TO A NANODOMAIN LOOPING REGION

Case 1: Uniform binding and unbinding rates of cross-linkers

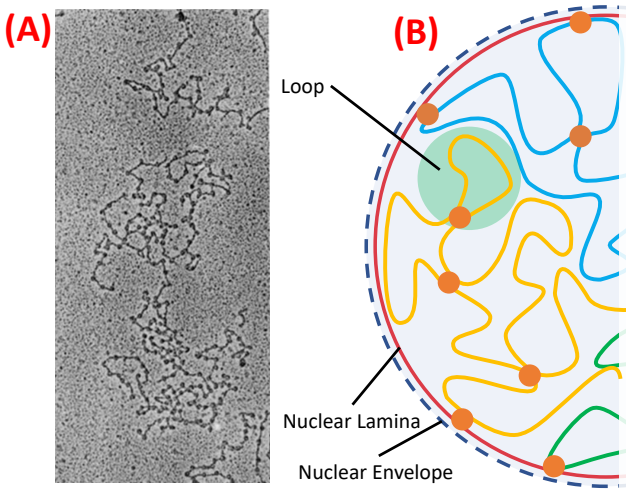

FIG. 1. **Illustration of chromatin Loops** (A) Electron-microscopy images of DNA showing many loops. (B) schematic representation of chromatin inside the nucleus ith its envelop (red and dashed blue lines), where cohesin is marked (orange dots) and often present at the base of a loop. Image reproduced by [17].

We start by describing the physical model we will use to study the interactions of diffusing cross-linkers with a nanoregion at the base of a formed loop. We use a Markov chain model to

coarse-grain the motion of N diffusing cross-linkers inside the nucleus, modeled as a ball. Cross-linkers diffuse until they enter a nanodomain region Ω (orange in Fig.1), which represents a binding region at the base of a loop. As long as there are enough cross-linkers, we consider that this region permits to stabilize the loop. Outside Ω, the cross-linker molecules diffuse freely. We coarse-grain the diffusion flux into Ω as an inward rate λ. We do not limit the number of cross-linkers in such region. Once a cross-linker enters Ω, it can stay there for a random time, which is Poissonian with rate μ.

We shall evaluate three scenarios of escape when the cross-linkers: 1- can escape independently (Fig.2A), 2- only at the two extremities (when the cross-linkers are organized linearly) can escape (Fig.2B), 3- escape with a rate which is rescaled by a factor $p \in [0, 1]$ except for the two extremities (Fig.2C).

Our goal is to estimate the number of cross-linkers inside Ω. As long as Ω contains at least

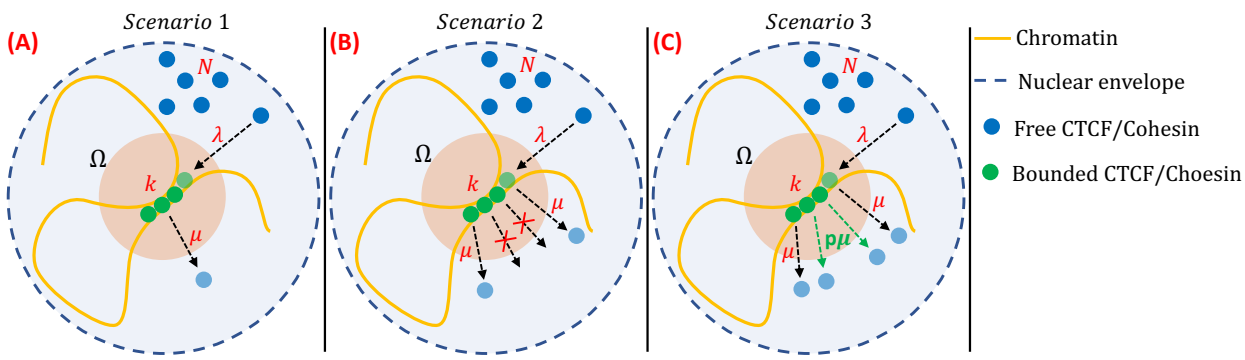

FIG. 2. **Three scenario for cross-linker bindings in a nanoregion to maintain chromatin loops.** (A) The escape rates λ (binding) and μ (binding) do not dependent on the number of cross-linkers; (B) the cross-linkers are organized linearly and only the two extremes can escape; (C) cross-linkers are still organized linearly but the escape rate λ is rescaled by a factor p ($0 \leq p \leq 1$) except for the two extremes case.

one cross-linker, the loop persists. When the last one escapes, the loop is lost and the region Ω disappears as well as the memory of the local chromatin organization. Thus to measure the loop stability, we compute the mean first passage time for the last cross-linkers to escape. To compute such time, we follow the number $n(t)$ of particles trapped within Ω at time t (only one particle at a time can enter) [12]. The probability density function (pdf) that there are $n(t) = k$ cross-linkers at time t inside Ω is for $0 \leq k \leq N$,

$$p_k(t) = P\{n(t) = k | n(0) = q\},$$

when there are initially $n(0) = q$ cross-linkers inside Ω. The transitions to the state where there are exactly k cross-linkers, is computed by considering the changes during time t and $t + \Delta t$:

$$\begin{aligned} p_k(t+\Delta t) =& p_{k-1}(t)\,\lambda\,[N-(k-1)]\,\Delta t \\ &+ p_k(t)\,[1-\mu k\Delta t-\lambda(N-k)\,\Delta t] \\ &+ p_{k+1}(t)\,(k+1)\mu\Delta t. \end{aligned}$$

For $k=0$ and $k=N$, we get

$$p_0(t+\Delta t) = p_0(t)\,(1-\lambda N\Delta t) + p_1(t)\,\mu\Delta t$$

and

$$p_N(t+\Delta t) = p_{N-1}(t)\,\lambda\Delta t + p_N(t)\,(1-\lambda N\Delta t)\,.$$

Thus, we obtain a Markov system of $N+1$ equations

$$\begin{cases} \dot{p}_0 & = -\lambda N p_0 + \mu p_1 \\ \dot{p}_k & = \lambda(N-k+1)\,p_{k-1} - [\mu k + \lambda(N-k)]\,p_k + \mu(k+1)\,p_{k+1} \\ \dot{p}_N & = \lambda p_{N-1} - \mu N p_N. \end{cases} \tag{1}$$

To compute the first time t that all cross-linkers escape the region Ω, the state $k=0$ should be absorbing so that the Markov chain is now modified to

$$\begin{cases} \dot{p}_0 & = \mu p_1 \\ \dot{p}_1 & = -[\mu + \lambda(N-1)]\,p_1 + 2\mu p_2 \\ \dot{p}_k & = \lambda(N-k+1)\,p_{k-1} - [\mu k + \lambda(N-k)]\,p_k + \mu(k+1)\,p_{k+1} \\ \dot{p}_N & = \lambda p_{N-1} - \mu N p_N \end{cases}. \tag{2}$$

The *mean time to lose a loop* $\bar{\tau}_{LL}$ is the first time when all cross-linkers have left the nanoregion. We refine the definition by considering th time $\bar{\tau}^q_{LL}$ when there are initially q bound cross-linkers:

$$\tau^q_{LL} = \inf\{t>0 | n(t)=0, n(0)=q\}\,.$$

Note that the probability of the time random time τ^q_{LL} is link to the Markov chain by the relation

$$\begin{aligned} \bar{\tau}^q_{LL} &= \int_0^\infty P\left\{\tau^q_{LL} > t | n(0)=q\right\} dt \\ &= \int_0^\infty [1-p_0(t)]\,dt = \int_0^\infty \sum_{k=1}^N p_k(t)\,dt = \sum_{k=1}^N a_k, \end{aligned}$$

where

$$a_k = \int_0^\infty p_k(t)\,dt \tag{3}$$

and the normalization condition is $\sum_{k=0}^{N} p_k(t) = 1$. We thus need to compute the a_k coefficients (see Appendix). We obtain using an induction procedure that

$$\begin{aligned}\bar{\tau}_{LL}^{q} =& \frac{1}{\mu}\sum_{n=0}^{q-1}\left(\frac{\lambda}{\mu}\right)^{n}\sum_{k=n+1}^{q}\frac{(k-n-1)!}{k!}\frac{(N-k+n)!}{(N-k)!}\\ &+\frac{1}{\mu}\sum_{n=0}^{q-1}\sum_{k=q+1}^{N}\left(\frac{\lambda}{\mu}\right)^{n+k-q}\frac{(N-q+n)!}{(N-k)!}\frac{(q-n-1)!}{k!}.\end{aligned}$$

In the limit case where $\lambda/\mu \ll 1$, this expression can be approximated at first order using $\sum_{k=1}^{N}\frac{1}{k} \approx \log(N)$ and $\sum_{n=2}^{p}\frac{1}{n(n-1)} = \frac{p-1}{p}$ by

$$\begin{aligned}\bar{\tau}_{LL}^{q} \approx& \frac{1}{\mu}\left(1-\frac{\lambda}{\mu}\right)\sum_{k=1}^{q}\frac{1}{k}+\frac{1}{\mu}\frac{\lambda}{\mu}(N+1)\frac{q}{q+1}\\ \approx& \frac{1}{\mu}\left(1-\frac{\lambda}{\mu}\right)\log(q)+\frac{1}{\mu}\frac{\lambda}{\mu}(N+1)\frac{q}{q+1}.\end{aligned} \tag{4}$$

In the opposite limit $\lambda/\mu \gg 1$, we obtain using the leading order term,

$$\bar{\tau}_{LL}^{q} \approx \frac{1}{\mu}\left(\frac{\lambda}{\mu}\right)^{N-1}\frac{1}{N}. \tag{5}$$

We plotted the time $\bar{\tau}_{LL}^{q}$ with respect to various parameters in Figs 3-6 and we will discuss the results below in comparison with the two other models.

Case 2: Unbinding rate depends on 2 cross-linker and the binding rate is uniform

We consider here the case where the unbinding can occur through the two external cross-linkers organized in line (Fig. 2B). The two extreme cross-linkers can escape with a rate λ, while the others have to wait until their turn. The probability to find k particles is now

$$\begin{aligned}p_k(t+\Delta t) =& p_{k-1}(t)\,\lambda\,[N-(k-1)]\,\Delta t\\ &+ p_k(t)\,[1-\mu k\Delta t-\lambda(N-k)\,\Delta t]+p_{k+1}(t)\,2\mu\Delta t,\end{aligned}$$

and the new equations are (state 0 is absorbing):

$$\begin{cases}\dot{p}_0 &= \mu p_1\\ \dot{p}_1 &= -[\mu+\lambda(N-1)]\,p_1+2\mu p_2\\ \dot{p}_k &= \lambda(N-k+1)\,p_{k-1}-[\mu k+\lambda(N-k)]\,p_k+2\mu p_{k+1}\\ \dot{p}_N &= \lambda p_{N-1}-\mu N p_N.\end{cases} \tag{6}$$

Following the method of induction as described in Appendix (see Appendix), we obtain the general expression:

$$\begin{aligned}\overline{\tau}_{LL}^{q} =& \frac{1}{\mu} + \frac{1}{\mu}\sum_{n=1}^{q-1}\left(\frac{\lambda}{\mu}\right)^{n}\frac{(N-1)!}{(N-n-1)!}\frac{1}{2^{n}} \\ &+ \frac{1}{\mu}\sum_{n=0}^{q-1}\left(\frac{\lambda}{\mu}\right)^{n}\sum_{k=n+2}^{q}\frac{(N-k+n)!}{(N-k)!}\frac{1}{2^{n+1}} \\ &+ \frac{1}{\mu}\sum_{n=0}^{q-2}\sum_{k=q+1}^{N}\left(\frac{\lambda}{\mu}\right)^{n+k-q}\frac{(N-q+n)!}{(N-k)!}\frac{1}{2^{k-q+n+1}} \\ &+ \frac{1}{\mu}\sum_{k=q+1}^{N}\left(\frac{\lambda}{\mu}\right)^{k-1}\frac{(N-1)!}{(N-k)!}\frac{1}{2^{k-1}}.\end{aligned}$$

In the limit $\lambda/\mu \ll 1$,we approximated this sum using the first order in μ/λ. With $\sum_{k=1}^{N}\frac{1}{k} \approx \log(N)$ and $\sum_{n=2}^{p}\frac{1}{n(n-1)} = \frac{p-1}{p}$ we get:

$$\overline{\tau}_{LL}^{q} \approx \frac{1}{\mu}\frac{q+1}{2} + \frac{1}{\mu}\frac{\lambda}{\mu}\frac{N\left(q+1\right)}{4}. \tag{7}$$

In the opposite limit $\lambda/\mu \gg 1$, this sum can be approximated at the last order of λ/μ, leading to

$$\overline{\tau}_{LL}^{q} \approx \frac{1}{\mu}\left(\frac{\lambda}{\mu}\right)^{N-1}\frac{(N-1)!}{2^{N-1}}. \tag{8}$$

Case 3: Non-uniform unbinding rate λ and uniform rate μ

In this case, we consider here that the two external cross-linkers can escape with a rate λ while the others located inside (Fig. 5) can detache and escape with a modified rate $p\lambda$ where $p \in [0,1]$ is a free parameter. The probability to find k particles in the domain Ω, for $k > 2$ is now

$$\begin{aligned}p_k\left(t+\Delta t\right) =& \, p_{k-1}\left(t\right)\lambda\left[N-(k-1)\right]\Delta t \\ &+ p_k\left(t\right)\left[1-\mu\left[2+p\left(k-2\right)\right]\Delta t - \lambda\left(N-k\right)\Delta t\right] \\ &+ p_{k+1}\left(t\right)\left[2+p\left(k-1\right)\right]\mu\Delta t.\end{aligned}$$

The Markov chain where the 0-state remains absorbing is

$$\begin{cases}\dot{p}_0 &= \mu p_1 \\ \dot{p}_1 &= -\left[\mu+\lambda\left(N-1\right)\right]p_1 + 2\mu p_2 \\ \dot{p}_k &= \lambda\left(N-k+1\right)p_{k-1} - \left[\mu\left[2+p\left(k-2\right)\right]\right. \\ & \quad \left.+\lambda\left(N-k\right)\right]p_k + \mu\left[2+p\left(k-1\right)\right]p_{k+1} \\ \dot{p}_N &= \lambda p_{N-1} - \mu\left[2+p\left(N-2\right)\right]p_N.\end{cases} \tag{9}$$

Finally, we obtain the general expression using the method of induction as described in Appendix (see Appendix):

$$\begin{aligned}\overline{\tau}_{LL}^{q} &= \frac{1}{\mu} + \frac{1}{\mu}\sum_{n=1}^{q-1}\left(\frac{\lambda}{\mu}\right)^{n}\frac{(N-1)!}{(N-n-1)!}\prod_{\tau=0}^{n-1}\left[2+p\left(n-1-\tau\right)\right]^{-1} \\ &+ \frac{1}{\mu}\sum_{n=0}^{q-1}\left(\frac{\lambda}{\mu}\right)^{n}\sum_{k=n+2}^{q}\frac{(N-k+n)!}{(N-k)!}\prod_{\tau=0}^{n}\left[2+p\left(k-2-\tau\right)\right]^{-1} \\ &+ \frac{1}{\mu}\sum_{n=0}^{q-2}\sum_{k=q+1}^{N}\left(\frac{\lambda}{\mu}\right)^{n+k-q}\frac{(N-q+n)!}{(N-k)!}\prod_{\tau=0}^{k-q+n}\left[2+p\left(k-2-\tau\right)\right]^{-1} \\ &+ \frac{1}{\mu}\sum_{k=q+1}^{N}\left(\frac{\lambda}{\mu}\right)^{k-1}\frac{(N-1)!}{(N-k)!}\prod_{\tau=0}^{k-2}\left[2+p\left(k-2-\tau\right)\right]^{-1}.\end{aligned}$$

In the limit $\lambda/\mu \ll 1$, this sum can be approximated at the first order of μ/λ. Using $\sum_{k=1}^{N}\frac{1}{k} \approx \log(N)$ and $\sum_{n=2}^{p}\frac{1}{n(n-1)} = \frac{p-1}{p}$ we get:

$$\begin{aligned}\overline{\tau}_{LL}^{q} &\approx \frac{1}{\mu} + \frac{1}{\mu}\frac{\lambda}{\mu}\frac{(N-1)}{2} + \frac{1}{\mu}\sum_{k=2}^{q}\frac{1}{\left[2+p\left(k-2\right)\right]} \\ &+ \frac{1}{\mu}\frac{\lambda}{\mu}\sum_{k=3}^{q}\frac{(N-k+1)}{\left[2+p\left(k-2\right)\right]\left[2+p\left(k-2-1\right)\right]} + \frac{1}{\mu}\frac{\lambda}{\mu}\frac{(N-q)}{\left[2+p\left(q-1\right)\right]}\end{aligned}$$

By expanding $\frac{1}{2+p(k-2)}$ as $\frac{1}{2}(1-p(k-2)/2)$ and $\frac{1}{[2+p(k-2)]2+p(k-3)]}$ as $\frac{1}{4}[1-p(k-2)/2-p(k-3)/2]$, we get

$$\overline{\tau}_{LL}^{q} \approx \frac{1}{2\mu}\left[(1+q) - \frac{p}{2}\left(q-2\right)\left(q-1\right)\right] + \frac{1}{\mu}\frac{\lambda}{\mu}\frac{N}{2}\left(1+\frac{q}{2}-\frac{pq^{2}}{4}\right). \tag{10}$$

In the limit $\lambda/\mu \gg 1$, $\overline{\tau}_{LL}^{q}$ can be approximated at the last order of λ/μ:

$$\overline{\tau}_{LL}^{q} \approx \frac{1}{\mu}\left(\frac{\lambda}{\mu}\right)^{N-1}\frac{(N-1)!}{\prod_{\tau=0}^{N-2}\left[2+p\left(N-2-\tau\right)\right]}. \tag{11}$$

Sensitivity analysis of the residence time $\overline{\tau}_{LL}^{q}$ with respect to the main parameters

We now compare the results for the residence time $\overline{\tau}_{LL}^{q}$ for the three scenario. As we shall see, the time of loop maintenance by the dynamics binding is quite robust. Since the unbinding rate μ is not precisely known, we will explore a large range of values. However, the forward rate λ in a diffusing model of CTCF or cohesin molecules moving freely in the nucleus can be estimated from the measured diffusion coefficient: $D_{free}^{CTCF} \approx 2.5\ \mu m^{2}/s$ and $D_{free}^{cohesin} \approx 1.5\ \mu m^{2}/s$ [18]. Given the typical size of a TAD ($\sim 100\ kbp$) and the radius $a \approx 0.1\ \mu m$ of spherical Ω region is around , we shall use the rate theory or Narrow Escape [19]. The reciprocal of the first mean time of Brownian particle to arrive to the region Ω is $\frac{1}{\lambda} \approx \frac{V}{4\pi a D}$, where $V \approx 1\ \mu m^{3}$ is the volume of the nucleus, then $\frac{1}{\lambda} \approx 1/\pi\ s^{-1}$.

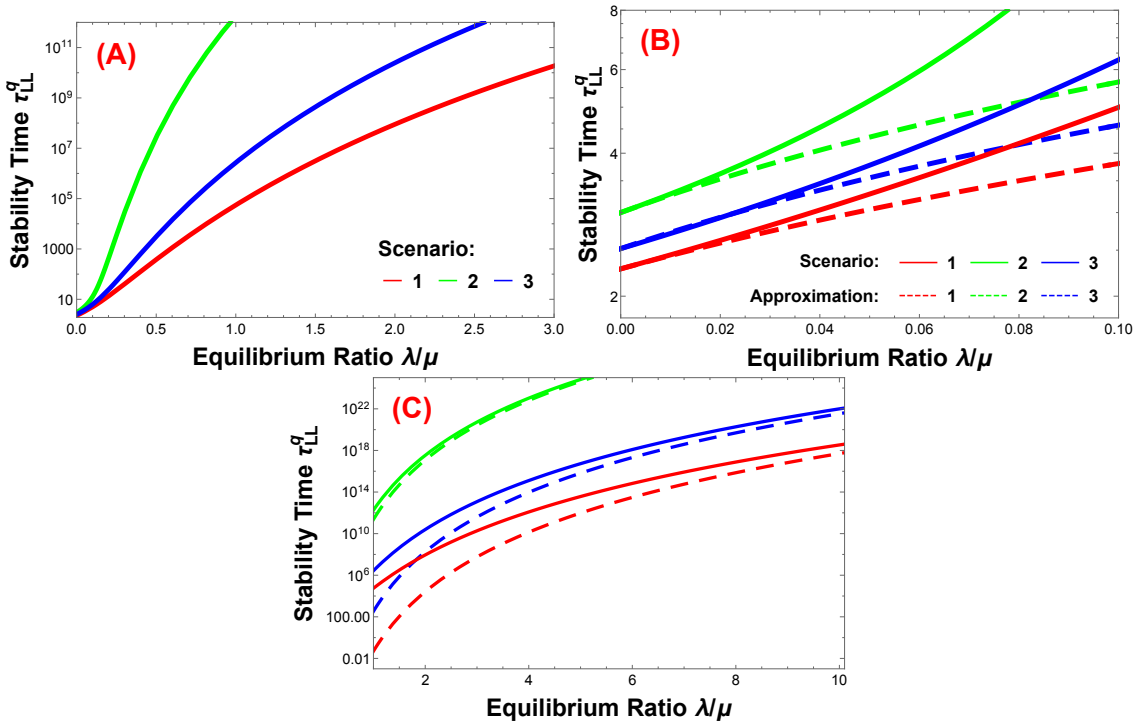

FIG. 3. **Stability time $\overline{\tau}^q_{LL}$ vs equilibrium ratio λ/μ for the three scenario.** (A) Analytic expressions in the three scenario considered. (B) Limit $\lambda/\mu \ll 1$, comparison of the exact relations (continuous lines) with the approximated equations 4, 7 and 10 (dashed). (C) Limit $\lambda/\mu \gg 1$, comparison (continuous lines) with the approximated equations 5, 8 and 11 (dashed). Total number of $N = 20$ when there are $q = 5$ molecules initially. The forward rate is set to $\mu = 1\ s^{-1}$. For the third scenario, $p = \frac{1}{2}$.

We first plotted the stability time $\overline{\tau}^q_{LL}$ vs equilibrium ratio λ/μ for the three scenario in Fig. 3A and found that for a small variation from 1 to 3 the time increases from order of magnitude, demonstrating th robustness of the process. We also compare the approximation to the full solution in Fig. 3B-C.

To further explore the stability of the time $\overline{\tau}^q_{LL}$, we vary the total number of cross-linkers and found again a robust time when N varied from 5 to 30 in all scenarios (Fig. 4). Finally for scenario 3, we varied the releasing parameter p ($p = 1$ recovers the first scenario while $p = 0$ corresponds to second). Finally, we vary the initial number q of bound cross-linkers (Fig. 6): except for a small λ/μ ratio, the time increases extremely quickly independent of the initial values q.

DISCUSSION AND CONCLUDING REMARKS

We presented three scenario of a rate model to study how long-term memory of chromatin loops can be maintained following an initial promoter-enhancer binding. The model assumes that

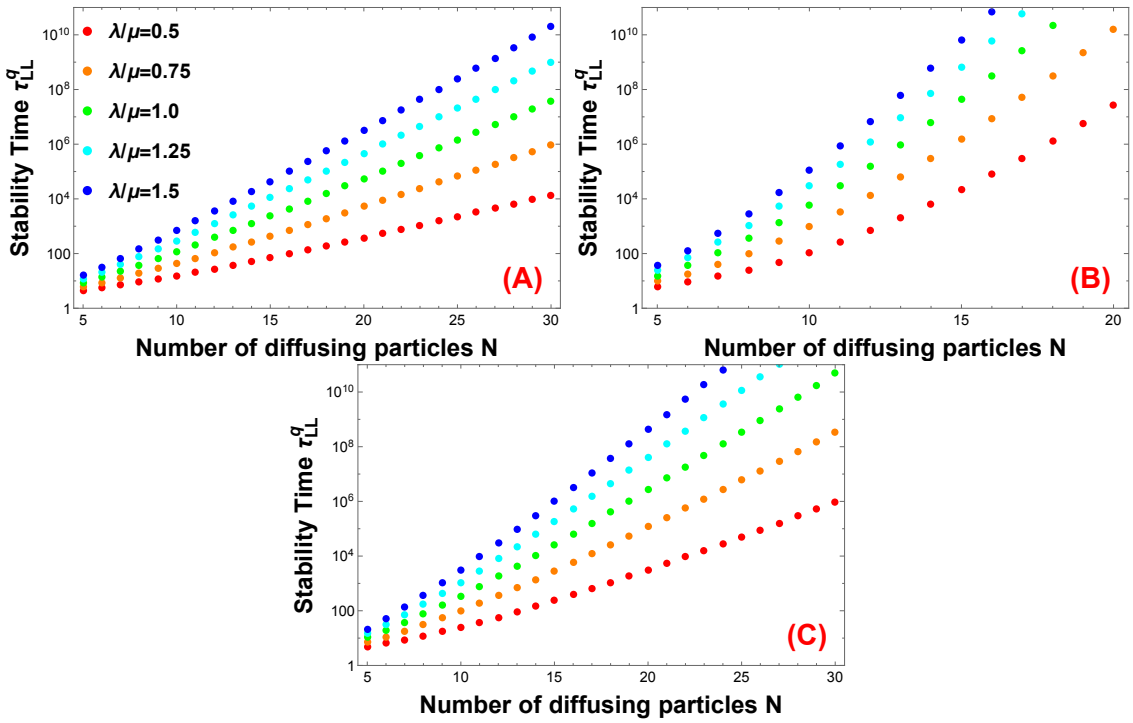

FIG. 4. **Stability time $\overline{\tau}^q_{LL}$ vs the total number of cross-linkers N for the three scenarios considered.** (A) scenario 1; (B) scenario 2; (C) scenario 3. There are $q = 5$ molecules initially. The forward rate is set to $\mu = 1\ s^{-1}$. For the third scenario, $p = \frac{1}{2}$.

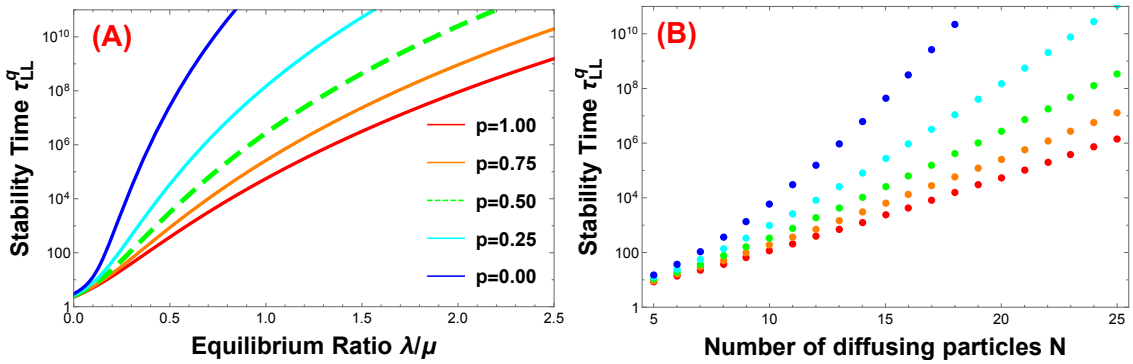

FIG. 5. **Effect of the releasing parameter p for the third scenario.** Stability time $\overline{\tau}^q_{LL}$ vs N (A) and equilibrium ratio λ/μ (panel B, where $N = 20$) for various values of the parameter p. There are $q = 5$ molecules initially and the forward rate is set to $\mu = 1\ s^{-1}$.

cross-linkers bind to a nanodomain region that keeps a loop. Although the scenario differs by the unbinding rates (uniform, unbinding from the extreme when organized linearly and partial unbinding from all cross-linkers), they all converge to show that few bound molecules in a neighborhood of the looping closing region are sufficient to guarantee that the number of cross-linkers will remain sufficient to maintain the loop for an extremely long time. With a number of less than 20 (resp

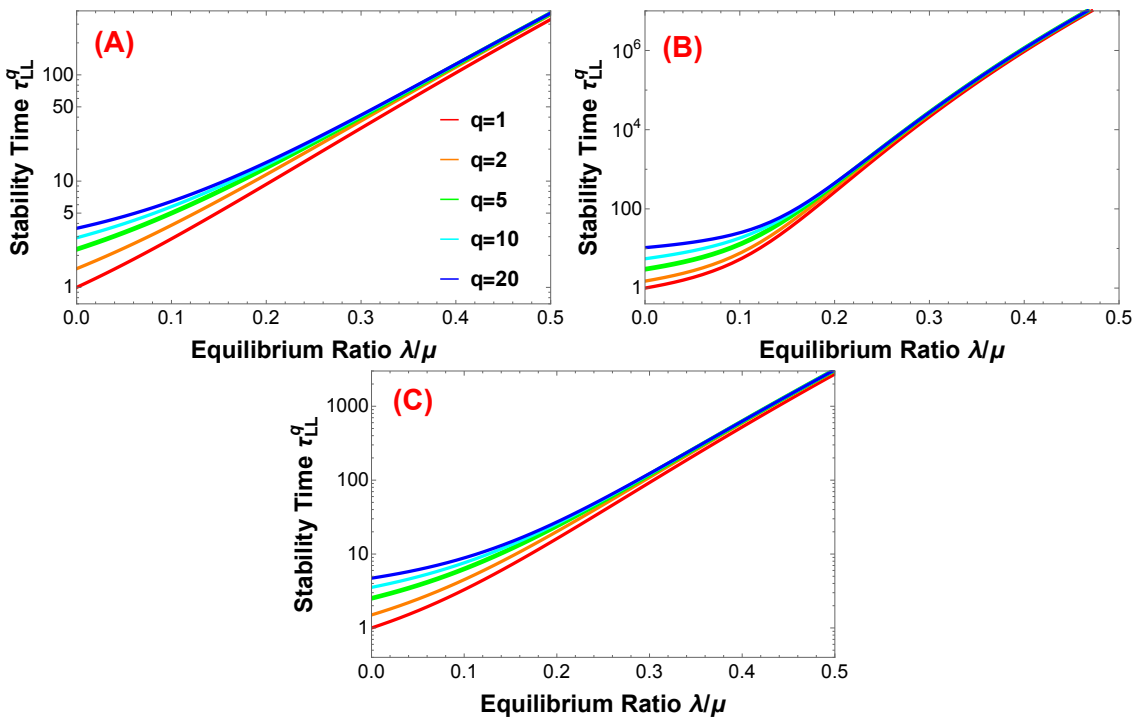

FIG. 6. **Effect of changing the initial number of bound cross-linkers on the stability time $\bar{\tau}^q_{LL}$ for the three scenario.** (A) scenario 1; (B) scenario 2; (C) scenario 3. Total number of $N = 20$ where there are $q = 5$ molecules initially. The forward rate is set to $\mu = 1\ s^{-1}$ For the third scenario, $p = \frac{1}{2}$.

30) cross-linkers, we found that a loop can be maintained for days (resp. months).

If the Ω-loop domains are sufficiently separated, the present analysis can generalized to the entire nucleus. Typically, in a nucleus there are around 10^3 TADs [7] and 10^4 loops have been identified [20], while there are around 10^5 cohesin molecules and 10^5 CTCF molecules (concentration ~ 144.3 nM) [21]. We can thus hypothesize that there are ~ 10 molecules per TAD, as single cell HiC reveal few loops per TADs [22]. Using the estimated $\lambda \sim \pi s^{-1}$, if loops are stable for days ($\sim 10^5$s) or months ($\sim 10^6$s), this would lead to the prediction that $\mu \lesssim 1.1 s^{-1}$ (for the first scenario) and $\mu \lesssim 0.51 s^{-1}$ (for the second scenario) (see figure 8). The third scenario falls between these two values, according to the parameter p. In summary, a possible scenario to keep the chromatin organization by stabilizing loops (memory hypothesis) is by a positive feedback of binding cross-linkers (fig. 7A-C): Initially a local transient chromatin loop is induced by promoter-enhancer contacts. The local proximity of chromatin induced by transcription factors and remodelers facilitates the binding of looping factors such as the diffusing cohesin molecules, which can attach and detach DNA dynamically. This dynamical binding and unbinding of cross-linker molecules lead to a preferential stabilization of loops which keeps the memory of the initial TF-remodeler chromatin change.

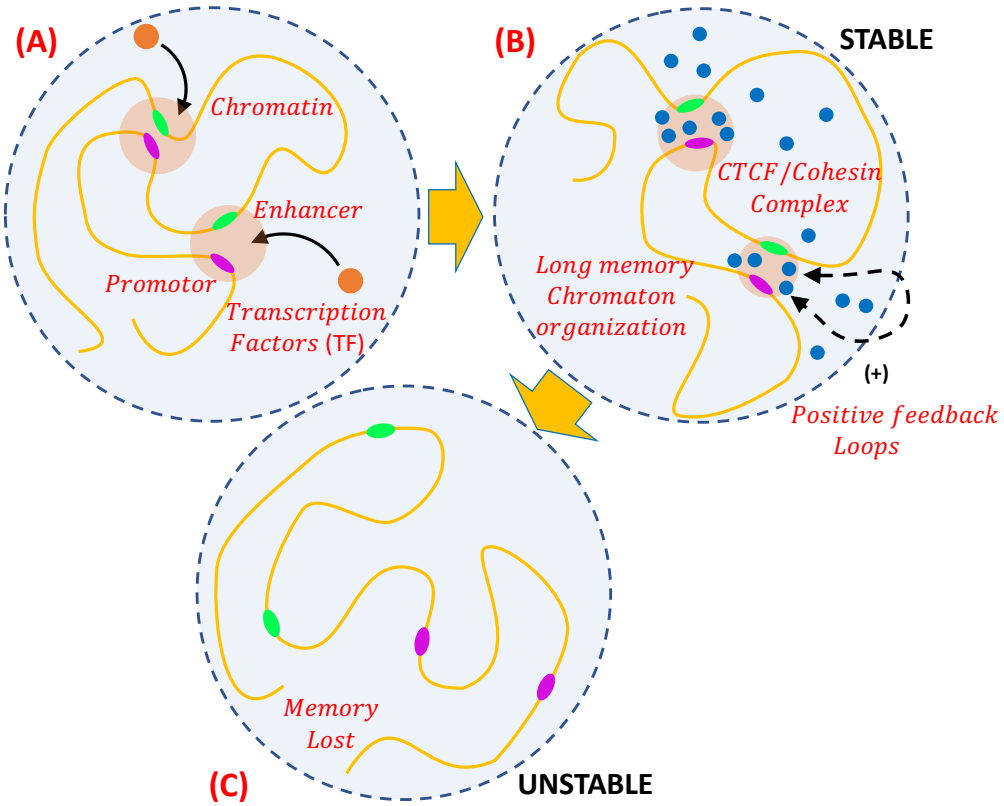

FIG. 7. **Chromatin memory hypothesis induced by positive feedback binding A** Local transient chromatin loop is induced by promoter-enhancer contacts thus making various loops. The local proximity of chromatin induced by transcription factors (TF) and remodelers facilitates the binding of looping factors such as the diffusing cohesin molecules, which can attach and detach DNA dynamically. **B** Dynamical binding and unbinding of cohesin molecules lead to a preferential stabilization of loops through a positive feedback process which keeps the memory of the initial TF-remodeler chromatin change. **C** If the dynamical balance between binding and unbinding cohesin molecules is not disrupted, the local loop organization of chromatin is lost.

However if such a dynamical balance between binding and unbinding cross-linker is disrupted, the local loop organization of chromatin is lost. This memory mechanism hypothesis should be further explored and tested experimentally.

APPENDIX

We present in the appendix the computation related to the three cases.

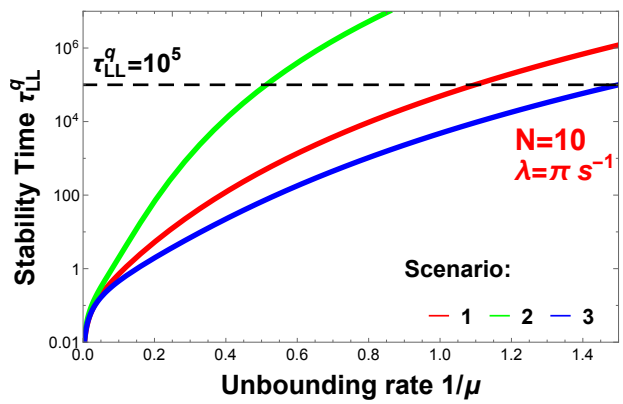

FIG. 8. **Comparison of the three scenario using parameters compatible with experimental data.** Considering around $N = 10$ cross-linkers per loop, the estimated rates are $\lambda \sim \pi s^{-1}$ and $\mu \lesssim 1.1 s^{-1}$ (for the first scenario) and $\mu \lesssim 0.51 s^{-1}$ (for the second scenario)

EXPLICIT COMPUTATION FOR MODEL 1

Using eq. 3, we rewrite the system of differential equations 2 using the initial conditions $p_k = \delta_{kq}$ and the limit of $t \to \infty$: $p_0(\infty) = 1$ and $p_k(\infty) = 0$. We get:

$$\begin{cases} 1 & = p_0(\infty) - p_0(0) = \mu a_1 \\ 0 & = p_1(\infty) - p_1(0) = -[\mu + (N-1)\lambda] a_1 + 2\mu a_2 \\ -1 & = p_q(\infty) - p_q(0) = \lambda (N-q+1) a_{q-1} \\ & \quad - [\lambda (N-q) + \mu q] a_q + (q+1) \mu a_{q+1} \\ 0 & = p_k(\infty) - p_k(0) = \lambda (N-k+1) a_{k-1} \\ & \quad - [\lambda (N-k) + \mu k] a_k + (k+1) \mu a_{k+1} \\ 0 & = p_N(\infty) - p_N(0) = \lambda a_{N-1} - \mu N a_N \end{cases}$$

We compute here the sum $\sum_{k=1}^{N} a_k$ with respect to the parameters λ, μ and N. The equations involving a_1 and a_2 resolve as

$$a_1 = \frac{1}{\mu} \quad \text{and} \quad a_2 = \frac{1}{2\mu} + \frac{\lambda}{\mu} \cdot \frac{N-1}{2\mu}. \tag{12}$$

In general, we obtain the induction relation:

$$k\mu a_k = 1 + \lambda (N-k+1) a_{k-1} \quad k \leq q$$

while

$$(q+z) \mu a_{q+z} = \lambda (N-q-z+1) a_{q+z-1} \quad z > 0$$

We express the last equation in term of a_q:

$$a_k = \left(\frac{\lambda}{\mu}\right)^{k-q} \frac{(N-q)!}{(N-k)!} \frac{q!}{(k)!} a_q.$$

where

$$a_q = \frac{1}{\mu} \sum_{n=0}^{q-1} \left(\frac{\lambda}{\mu}\right)^n \frac{(q-n-1)!}{q!} \frac{(N-q+n)!}{(N-q)!}.$$

These relations are summarized as

$$\begin{cases} a_k &= \frac{1}{k\mu} + \frac{\lambda}{\mu} \frac{(N-k+1)}{k} a_{k-1} \quad k \leq q \\ a_k &= \left(\frac{\lambda}{\mu}\right)^{k-q} \frac{(N-q)!}{(N-k)!} \frac{q!}{k!} a_q \quad k > q. \end{cases}$$

We split the sum for the expression of $\bar{\tau}_{LL}^q$ according to the initial number q of cross-linkers and we get

$$\bar{\tau}_{LL}^q = \sum_{k=1}^{N} a_k = \sum_{k=1}^{q} a_k + \sum_{k=q+1}^{N} a_k \tag{13}$$

where

$$\sum_{k=1}^{q} a_k = \frac{1}{\mu} \sum_{n=0}^{q-1} \left(\frac{\lambda}{\mu}\right)^n \sum_{k=n+1}^{q} \frac{(k-n-1)!}{k!} \frac{(N-k+n)!}{(N-k)!}$$

and

$$\begin{aligned} \sum_{k=q+1}^{N} a_k &= \frac{1}{\mu} \sum_{k=q+1}^{N} \left(\frac{\lambda}{\mu}\right)^{k-q} \frac{(N-q)!}{(N-k)!} \frac{q!}{k!} a_q \\ &= \frac{1}{\mu} \sum_{n=0}^{q-1} \sum_{k=q+1}^{N} \left(\frac{\lambda}{\mu}\right)^{n+k-q} \frac{(N-q+n)!}{(N-k)!} \frac{(q-n-1)!}{k!} \end{aligned}$$

Finally, we obtain the general expression

$$\begin{aligned} \bar{\tau}_{LL}^q =& \frac{1}{\mu} \sum_{n=0}^{q-1} \left(\frac{\lambda}{\mu}\right)^n \sum_{k=n+1}^{q} \frac{(k-n-1)!}{k!} \frac{(N-k+n)!}{(N-k)!} \\ &+ \frac{1}{\mu} \sum_{n=0}^{q-1} \sum_{k=q+1}^{N} \left(\frac{\lambda}{\mu}\right)^{n+k-q} \frac{(N-q+n)!}{(N-k)!} \frac{(q-n-1)!}{k!}. \end{aligned}$$

EXPLICIT COMPUTATION FOR MODEL 2

Following the method of induction described in the previous section of the appendix (expressions as eq. 12), we obtain the induction relation for a_k for the second scenario when the there are initially

q bound cross-linkers:

$$\begin{cases} a_k &= \frac{1}{2\mu} + \frac{\lambda}{\mu}\frac{(N-k+1)}{2} a_{k-1} \quad k \leq q \\ a_k &= \left(\frac{\lambda}{\mu}\right)^{k-q} \frac{(N-q)!}{(N-k)!}\frac{1}{2^{k-q}} a_q \quad k > q. \end{cases}$$

Using relation 13, we compute the time $\bar{\bar{\tau}}^q_{LL}$ using the three relations

$$\begin{aligned} \sum_{k=1}^{q} a_k =& \frac{1}{\mu} + \frac{1}{\mu}\sum_{n=1}^{q-1}\left(\frac{\lambda}{\mu}\right)^n \frac{(N-1)!}{(N-n-1)!}\frac{1}{2^n} \\ &+ \frac{1}{\mu}\sum_{n=0}^{q-1}\left(\frac{\lambda}{\mu}\right)^n \sum_{k=n+2}^{q} \frac{(N-k+n)!}{(N-k)!}\frac{1}{2^{n+1}}. \end{aligned}$$

$$\begin{aligned} a_q =& \frac{1}{\mu}\left(\frac{\lambda}{\mu}\right)^{q-1} \frac{(N-1)!}{(N-q)!}\frac{1}{2^n} \\ &+ \frac{1}{\mu}\sum_{n=0}^{q-2}\left(\frac{\lambda}{\mu}\right)^n \frac{(N-q+n)!}{(N-q)!}\frac{1}{2^{n+1}}. \end{aligned}$$

and the sum for $k > q$:

$$\begin{aligned} \sum_{k=q+1}^{N} a_k =& \frac{1}{\mu}\sum_{k=q+1}^{N}\left(\frac{\lambda}{\mu}\right)^{k-q} \frac{(N-q)!}{(N-k)!}\frac{1}{2^{k-q}} a_q \\ =& \frac{1}{\mu}\sum_{n=0}^{q-2}\sum_{k=q+1}^{N}\left(\frac{\lambda}{\mu}\right)^{n+k-q} \frac{(N-q+n)!}{(N-k)!}\frac{1}{2^{k-q+n+1}} \\ &+ \frac{1}{\mu}\sum_{k=q+1}^{N}\left(\frac{\lambda}{\mu}\right)^{k-1} \frac{(N-1)!}{(N-k)!}\frac{1}{2^{k-1}}, \end{aligned}$$

leading to the general expression:

$$\begin{aligned} \bar{\tau}^q_{LL} =& \frac{1}{\mu} + \frac{1}{\mu}\sum_{n=1}^{q-1}\left(\frac{\lambda}{\mu}\right)^n \frac{(N-1)!}{(N-n-1)!}\frac{1}{2^n} \\ &+ \frac{1}{\mu}\sum_{n=0}^{q-1}\left(\frac{\lambda}{\mu}\right)^n \sum_{k=n+2}^{q} \frac{(N-k+n)!}{(N-k)!}\frac{1}{2^{n+1}} \\ &+ \frac{1}{\mu}\sum_{n=0}^{q-2}\sum_{k=q+1}^{N}\left(\frac{\lambda}{\mu}\right)^{n+k-q} \frac{(N-q+n)!}{(N-k)!}\frac{1}{2^{k-q+n+1}} \\ &+ \frac{1}{\mu}\sum_{k=q+1}^{N}\left(\frac{\lambda}{\mu}\right)^{k-1} \frac{(N-1)!}{(N-k)!}\frac{1}{2^{k-1}}. \end{aligned}$$

EXPLICIT COMPUTATION FOR MODEL 3

In the third scenario, the steady-state coefficients a_1 and a_2 are again similar to the ones of equations 12 while the induction relation for a_k becomes:

$$\begin{cases} a_k &= \frac{1}{\mu[2+p(k-2)]} + \frac{\lambda}{\mu}\frac{(N-k+1)}{[2+p(k-2)]}a_{k-1} \quad k \leq q \\ a_k &= \left(\frac{\lambda}{\mu}\right)^{k-q} \frac{(N-q)!}{(N-k)!} \prod_{\tau=0}^{k-q-1} \left[2+p\left(q-1+\tau\right)\right]^{-1} a_q \quad k > q. \end{cases}$$

To evaluate the time $\overline{\tau}_{LL}^{q}$, we use the sums

$$\begin{aligned} \sum_{k=1}^{q} a_k =& \frac{1}{\mu} + \frac{1}{\mu}\sum_{n=1}^{q-1}\left(\frac{\lambda}{\mu}\right)^n \frac{(N-1)!}{(N-n-1)!}\prod_{\tau=0}^{n-1}\left[2+p\left(n-1-\tau\right)\right]^{-1} \\ &+ \frac{1}{\mu}\sum_{n=0}^{q-1}\left(\frac{\lambda}{\mu}\right)^n \sum_{k=n+2}^{q} \frac{(N-k+n)!}{(N-k)!}\prod_{\tau=0}^{n}\left[2+p\left(k-2-\tau\right)\right]^{-1}. \end{aligned}$$

Since

$$\begin{aligned} a_q =& \frac{1}{\mu}\left(\frac{\lambda}{\mu}\right)^{q-1} \frac{(N-1)!}{(N-q)!}\prod_{\tau=0}^{q-2}\left[2+p\left(q-2-\tau\right)\right]^{-1} \\ &+ \frac{1}{\mu}\sum_{n=0}^{q-2}\left(\frac{\lambda}{\mu}\right)^n \frac{(N-q+n)!}{(N-q)!}\prod_{\tau=0}^{n}\left[2+p\left(q-2-\tau\right)\right]^{-1}, \end{aligned}$$

we can obtain the sum for $k > q$:

$$\begin{aligned} \sum_{k=q+1}^{N} a_k &= \frac{1}{\mu}\sum_{k=q+1}^{N}\left(\frac{\lambda}{\mu}\right)^{k-q}\frac{(N-q)!}{(N-k)!}\prod_{\tau=0}^{k-q-1}\left[2+p\left(q-1+\tau\right)\right]^{-1} a_q \\ &= \frac{1}{\mu}\sum_{n=0}^{q-2}\sum_{k=q+1}^{N}\left(\frac{\lambda}{\mu}\right)^{n+k-q}\frac{(N-q+n)!}{(N-k)!}\prod_{\tau=0}^{k-q+n}\left[2+p\left(k-2-\tau\right)\right]^{-1} \\ &+ \frac{1}{\mu}\sum_{k=q+1}^{N}\left(\frac{\lambda}{\mu}\right)^{k-1}\frac{(N-1)!}{(N-k)!}\prod_{\tau=0}^{k-2}\left[2+p\left(k-2-\tau\right)\right]^{-1}. \end{aligned}$$

Finally, we obtain the general expression is:

$$\begin{aligned} \overline{\tau}_{LL}^{q} =& \frac{1}{\mu} + \frac{1}{\mu}\sum_{n=1}^{q-1}\left(\frac{\lambda}{\mu}\right)^n \frac{(N-1)!}{(N-n-1)!}\prod_{\tau=0}^{n-1}\left[2+p\left(n-1-\tau\right)\right]^{-1} \\ &+ \frac{1}{\mu}\sum_{n=0}^{q-1}\left(\frac{\lambda}{\mu}\right)^n \sum_{k=n+2}^{q} \frac{(N-k+n)!}{(N-k)!}\prod_{\tau=0}^{n}\left[2+p\left(k-2-\tau\right)\right]^{-1} \\ &+ \frac{1}{\mu}\sum_{n=0}^{q-2}\sum_{k=q+1}^{N}\left(\frac{\lambda}{\mu}\right)^{n+k-q}\frac{(N-q+n)!}{(N-k)!}\prod_{\tau=0}^{k-q+n}\left[2+p\left(k-2-\tau\right)\right]^{-1} \\ &+ \frac{1}{\mu}\sum_{k=q+1}^{N}\left(\frac{\lambda}{\mu}\right)^{k-1}\frac{(N-1)!}{(N-k)!}\prod_{\tau=0}^{k-2}\left[2+p\left(k-2-\tau\right)\right]^{-1}. \end{aligned}$$

ACKNOWLEDGEMENTS

This work was supported by Plan Cancer and Fondation pour la Recherche Médicale (FRM)-[SPF201909009284].

AVAILABILITY OF DATA

The data that support the findings (simulations) of this study are available from the corresponding author upon reasonable request.

Conflict of interest statement.

None declared.

[1] B. Alberts, D. Bray, J. Lewis, M. Raff, K. Roberts, and J. D. Watson, "Molecular biology of the cell," 2003.

[2] S. A. Moody, *Principles of developmental genetics.* Academic Press, 2014.

[3] M. Ptashne, "A genetic switch: Gene control and phage. lambda," 1986.

[4] 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, R. Sandstrom, B. Bernstein, M. A. Bender, M. Groudine, A. Gnirke, J. Stamatoyannopoulos, L. A. Mirny, E. S. Lander, and J. Dekker, "Comprehensive mapping of long-range interactions reveals folding principles of the human genome," *Science*, vol. 326, no. 5950, pp. 289–293, 2009.

[5] R. Katainen, K. Dave, E. Pitkänen, K. Palin, T. Kivioja, N. Välimäki, A. E. Gylfe, H. Ristolainen, U. A. Hänninen, T. Cajuso, J. Kondelin, T. Tanskanen, J.-P. Mecklin, H. Järvinen, L. Renkonen-Sinisalo, A. Lepistö, E. Kaasinen, O. Kilpivaara, S. Tuupanen, M. Enge, J. Taipale, and L. A. Aaltonen, "Ctcf/cohesin-binding sites are frequently mutated in cancer," *Nature Genetics*, vol. 47, no. 7, pp. 818–821, 2015.

[6] Y. X. See, B. Z. Wang, and M. J. Fullwood, "Chromatin interactions and regulatory elements in cancer: From bench to bedside," *Trends in Genetics*, vol. 35, no. 2, pp. 145 – 158, 2019.

[7] 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*, vol. 485, no. 7398, pp. 376–380, 2012.

[8] G. Wutz, C. Várnai, K. Nagasaka, D. A. Cisneros, R. R. Stocsits, W. Tang, S. Schoenfelder, G. Jessberger, M. Muhar, M. J. Hossain, *et al.*, "Topologically associating domains and chromatin loops

depend on cohesin and are regulated by ctcf, wapl, and pds5 proteins," *The EMBO journal*, vol. 36, no. 24, pp. 3573–3599, 2017.

[9] G. A. Busslinger, R. R. Stocsits, P. Van Der Lelij, E. Axelsson, A. Tedeschi, N. Galjart, and J.-M. Peters, "Cohesin is positioned in mammalian genomes by transcription, ctcf and wapl," *Nature*, vol. 544, no. 7651, pp. 503–507, 2017.

[10] J. R. Dixon, I. Jung, S. Selvaraj, Y. Shen, J. E. Antosiewicz-Bourget, A. Y. Lee, Z. Ye, A. Kim, N. Rajagopal, W. Xie, Y. Diao, J. Liang, H. Zhao, V. V. Lobanenkov, J. R. Ecker, J. A. Thomson, and B. Ren, "Chromatin architecture reorganization during stem cell differentiation," *Nature*, vol. 518, no. 7539, pp. 331–336, 2015.

[11] D. Holcman and Z. Schuss, "Stochastic chemical reactions in microdomains," *The Journal of chemical physics*, vol. 122, no. 11, p. 114710, 2005.

[12] K. Dao Duc and D. Holcman, "Threshold activation for stochastic chemical reactions in microdomains," *Phys. Rev. E*, vol. 81, p. 041107, Apr 2010.

[13] D. Holcman and Z. Schuss, *Stochastic Narrow Escape in Molecular and Cellular Biology.* Springer, 2015.

[14] D. Holcman and Z. Schuss, "100 years after smoluchowski: stochastic processes in cell biology," *Journal of Physics A: Mathematical and Theoretical*, vol. 50, no. 9, p. 093002, 2017.

[15] O. Shukron, U. Dobramysl, and D. Holcman, "Chemical reactions for molecular and cellular biology,"

[16] A. Goloborodko and J. F. Marko, "Chromosome compaction by active loop extrusion.," *Biophys. J.*, vol. 110, no. 10, pp. 2162–2168, 2016.

[17] J. Allan, N. Harborne, D. C. Rau, and H. Gould, "Participation of core histone "tails" in the stabilization of the chromatin solenoid.," *Journal of Cell Biology*, vol. 93, pp. 285–297, 05 1982.

[18] A. S. Hansen, I. Pustova, C. Cattoglio, R. Tjian, and X. Darzacq, "Ctcf and cohesin regulate chromatin loop stability with distinct dynamics," *eLife*, vol. 6, p. e25776, 05 2017.

[19] Z. Schuss, A. Singer, and D. Holcman, "The narrow escape problem for diffusion in cellular microdomains," *Proceedings of the National Academy of Sciences*, vol. 104, no. 41, pp. 16098–16103, 2007.

[20] 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*, vol. 159, no. 7, pp. 1665–1680, 2014.

[21] C. Cattoglio, I. Pustova, N. Walther, J. J. Ho, M. Hantsche-Grininger, C. J. Inouye, M. J. Hossain, G. M. Dailey, J. Ellenberg, X. Darzacq, R. Tjian, and A. S. Hansen, "Determining cellular ctcf and cohesin abundances to constrain 3d genome models," *eLife*, vol. 8, p. e40164, jun 2019.

[22] T. J. Stevens, D. Lando, S. Basu, L. P. Atkinson, Y. Cao, S. F. Lee, M. Leeb, K. J. Wohlfahrt, W. Boucher, A. OShaughnessy-Kirwan, *et al.*, "3d structures of individual mammalian genomes studied by single-cell hi-c," *Nature*, vol. 544, no. 7648, pp. 59–64, 2017.