# Predicting statistics of gene translocation events: Role of chromatin compaction and double-strand DNA break

Anirudh Jairam,[1,*,†] Shuvadip Dutta,[2,*,‡] Sangram Kadam,[1] Kiran Kumari,[1] and Ranjith Padinhateeri[1,3,4,§]
[1]*Department of Biosciences and Bioengineering, Indian Institute of Technology Bombay, Mumbai 400076, India*
[2]*Department of Physics, Indian Institute of Technology Bombay, Mumbai 400076, India*
[3]*Sunita Sanghi Centre of Aging and Neurodegenerative Diseases, Indian Institute of Technology Bombay, Mumbai 400076, India*
[4]*Koita Centre for Digital Health, Indian Institute of Technology Bombay, Mumbai 400076, India*

(Received 11 August 2025; revised 1 January 2026; accepted 21 January 2026; published 9 February 2026)

Chromosomal translocations, arising from unresolved double-stranded DNA breaks (DSBs), play a central role in genome instability and evolution. A prevailing hypothesis suggests that the probability of translocation between two chromatin segments depends on both their spatial proximity and the likelihood of DSB formation and rejoining. To test this, we perform Monte Carlo simulations of chromatin model polymers with varying levels of compaction and three-dimensional organization. Our results reveal that translocation probability cannot be fully explained by simple two-segment contact probabilities. Instead, it is strongly modulated by the global polymer compaction, which increases the incidence of multisegment contacts and introduces higher-order contributions. We further demonstrate that translocation probability exhibits a nontrivial functional dependence on both contact probability and DSB probability. We propose an empirical formula to compute translocation probability and provide analytical arguments for simple cases. Together, our findings highlight the critical role of chromatin organization in shaping the landscape of genome rearrangements and provide a framework to quantitatively predict translocation events from underlying polymer features and breakage statistics.

DOI: 10.1103/kyds-xmlh

## I. INTRODUCTION

The genetic code is encoded as a sequence of nucleotides and exists in the form of double-stranded DNA (dsDNA). In a eukaryotic nucleus, the DNA is further folded with the help of a large number of proteins to form a long polymer called chromatin. Every biological process, including transcription, replication, and stochastic changes in DNA sequence leading to evolution, happens in the context of chromatin [1–3]. The folding of DNA into chromatin not only aids in genome packaging [3,4] but is also crucial for gene regulation [5,6] and protecting genes from damage [7–10].

The precise nature of the three-dimensional (3D) organization of chromatin is crucial for many biological processes [11–13]. One experimental approach to determine the 3D organization of the chromatin polymer is chromatin conformation capture [14–16]. A high-throughput version of this method, known as Hi-C, measures the frequency of contact between any two regions of chromatin [17–19]. These methods have revealed that chromatin exhibits a hierarchical structure, with topologically associating domains (TADs) and compartments present at various length scales [20,21]. It has been shown that the probability $C_{ij}$ of two DNA segments $i$ and $j$, along the chromatin contour, coming into contact follows a power-law scaling relation: $C_{ij} \sim |i-j|^{-\gamma}$ [17,22]. Depending on the local nature of compaction, $\gamma$ can vary, typically ranging from 2.2 to 0.5 [23–25]. For a randomly folded polymer, $\gamma = 3/2$. For chromatin, the contact probability scaling exponent is approximately $\gamma \approx 1$ at larger genomic length scales [17], while at the scale of TADs it decreases to $\gamma \approx 0.7$ [22,26].

In cells, dsDNA is frequently damaged or broken and undergoes regular repair [21,27,28]. Some of the pathways involved in resolving dsDNA damage can lead to genetic recombination or chromosomal translocation [29,30]. Unlike in meiosis, natural recombination is uncommon in interphase cells [4]. Within the territory of a single chromosome, homologous pairs are rarely in close proximity, making it more likely that breaks induced during interphase are resolved through nonhomologous end-joining (NHEJ), which can lead to DNA repair or translocation [31–34]. This suggests that if chromatin segments with multiple DNA breaks are in close 3D proximity, the likelihood of translocation between them increases.

The role of 3D proximity between chromatin segments in facilitating recombination has been investigated using various techniques [35–38], including high-throughput genome translocation sequencing (HTGTS) [39], which provides translocation frequency information. Lee *et al.* [40] found a correlation between the spatial proximity of two genomic regions and the likelihood of homology-mediated recombination as a mechanism to repair double-stranded breaks (DSBs). Zhang *et al.* [37] used $\gamma$ irradiation to induce random DSBs throughout the genome and observed that higher contact frequencies between two regions were associated with increased translocation frequencies in mouse models lacking

*These authors contributed equally to this work.
†Contact author: anirudhjay97@gmail.com
‡Contact author: shuvadipdutta@gmail.com
§Contact author: ranjithp@iitb.ac.in

 

DNA repair proteins such as ATM. They also found that most translocations in this context were mediated through the NHEJ repair pathway. Alt *et al.* [41] proposed that the basic relationship between translocation frequency and contact probability can be expressed as

$$T_{ij} \propto d_i d_j C_{ij}, \tag{1}$$

where $T_{ij}$ and $C_{ij}$ represent the translocation probability and contact probability, respectively, between genomic regions $i$ and $j$, while $d_i$ and $d_j$ denote the DSB frequencies at locations $i$ and $j$. However, this equation does not account for the roles of proteins involved in the repair process or for biases that favor self-repair of DSBs over translocations and genomic rearrangements [41]. Additionally, the equation does not incorporate any information about the nature of chromatin organization or the surrounding environment beyond the specific locations $i$ and $j$.

Given Eq. (1), a natural way to proceed is $T_{ij} = kC_{ij}d_i d_j$ and assume that $k$ is a constant. However, if the proportionality constant depends on the local compaction and other properties, $k$ will not be a constant. To investigate this, we generated polymer configurations and simulated translocation events based on proximity of regions. From the translocation probabilities computed in the simulations, we estimated $k$. Our results show that the overall compaction of chromatin influences the parameter $k$ and thereby influences the translocation of two genomic regions.

## II. MODEL AND METHODS

Here, we describe a model to study the relationship between the 3D organization of chromatin and translocation, and a Monte Carlo method to simulate translocation events, in three parts: (i) generating an ensemble of single-cell-level interchromatin contact information from population-averaged Hi-C data, (ii) Monte Carlo moves to simulate double-stranded DNA breaks and translocation in accordance with 3D chromatin contacts, and (iii) measurements and analyses based on our simulation data.

We consider chromatin as a linear polymer made of $N$ coarse-grained beads [42–51] with each bead representing $l$ kb of chromatin. To account for additional contacts due to intrachromatin interactions, we use population-averaged contact probability matrix $C_{ij}$ either from Hi-C experiments or from theoretical descriptions of chromatinlike model polymers. We then generate an ensemble of binary contact matrices ($M_{ij}$), each representing a single chromatin polymer configuration. To achieve this we perform the following procedure. Choose a pair of beads ($i$, $j$) from a single polymer and assign a contact between the pairs if $r_n < C_{ij}$, where $r_n$ is a random number drawn from a uniform distribution of numbers in the range [0,1]. If the contact is assigned, the corresponding element in a matrix is set to $M_{ij} = 1$, and $M_{ij} = 0$ otherwise. Repeating this for every possible pair once gives us contact information of a single configuration of chromatin.

Given the $M_{ij}$ matrices representing chromatin configurations, we implement Monte Carlo moves resulting in DSBs and translocation (see Fig. 1). A DSB at location $i$ is considered as breaking of a DNA segment, belonging to the $i$th coarse-grained bead with a probability $d_i$. This results in

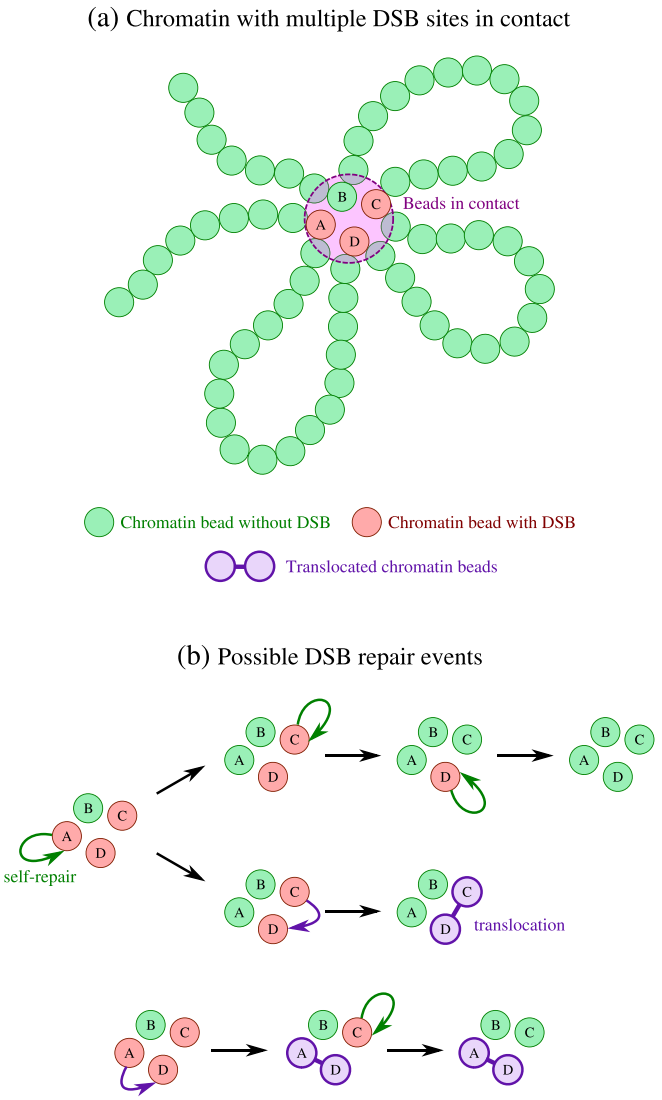


FIG. 1. (a) A polymer configuration with four segments (beads) in contact (highlighted by the circle); three among them contain DSBs (red beads). (b) Schematic representation of various possible DSB repair and translocation events. Top trajectory: segment A, segment C, and segment D undergo self-repair one after the other. Middle trajectory: after the self-repair of segment A, segments C and D get translocated with each other (violet pair). Bottom trajectory: segments A and D get translocated. Segment C repairs with itself.

a one-dimensional binary vector ($b_i = 0$ or 1) representing whether the bead is broken in a given chromatin polymer.

The broken beads are fixed as follows: we choose a broken bead ($i$) randomly (i.e., $b_i = 1$) and find out all other broken beads which are in contact with the chosen bead. That is, for the chosen bead $i$, we find all matrix elements where $M_{ij} = 1$ as well as $b_j = 1$. If there are $N_b$ broken beads in contact with $i$ (including $i$), we "repair" the breakage with one of them randomly (i.e., with probability $1/N_b$). If the broken bead $i$ chooses to repair with any other bead $j$, we call it a translocation between $i$ and $j$. If a segment chooses itself ($i$ choosing $i$), then this would be called a DSB self-repair, and we do not count this outcome as a translocation [see Fig. 1(b)]. We repeat this process until all broken segments are

either self-repaired or translocated, completing one simulation realization. When two DSBs occur, there are four broken ends. We assume that when one pair of broken ends is translocated, the remaining pair will translocate with each other. The whole process of breakage and repair is repeated $10^9$ times, ensuring that we cover the entire space of outcomes and sample enough translocations to generate good statistics to compute the probability of translocation between segments $i$ and $j$, $T_{ij}$, defined as

$$T_{ij} = \frac{\text{Total number of translocations between } i \text{ and } j}{\text{Total number of realizations}}.$$

Prompted by the experimental observations by Chiarle *et al.* [39], we assume that all translocations in our simulations follow NHEJ. Hence, there is no involvement of DNA sequence as in the case of homology-directed repair. In this work we assume that all the repair or translocation events are stochastic and any effect due to proteins will only affect the translocation timescale, which is not relevant in this study, as we are investigating only the steady-state properties. All possible scenarios of rearrangements, including deletions and circularization of small segments, are accounted for in our procedure. Once we obtain $T_{ij}$, we define the translocation fraction as

$$k_{ij} = T_{ij}/(d_i d_j C_{ij}). \tag{2}$$

Here, $k_{ij}$ represents the ratio of the actual probability of translocation to the expected translocation probability from an existing simple theory [see Eq. (1)]. We also compute the "mean translocation fraction," defined as $\bar{k} = [1/N(N-1)] \sum_{i=1}^{N} \sum_{j=1, j\neq i}^{N} k_{ij}$. Throughout this work, we assume $d_i = d_j = d$ (also see Discussion and Conclusion section). In this study we present translocation statistics computed using various polymer models, including: (i) a random walk (RW) polymer having $C_{ij} \propto |i-j|^{-3/2}$, (ii) a self-avoiding walk (SAW) polymer, (iii) polymer globules where the chromatin is considered as a bead-spring chain with every bead having a uniform attractive interaction with every other bead via a Lennard-Jones (LJ) potential, and (iv) 500-kbp domain (Chr7: 137,250 kbp - 137,750 kbp) from K562 and IMR90 cell lines, at 10-kbp resolution from simulated data in [44]. We also used the contact map of *Candida tropicalis* whole genome from published data [52].

*Polymer models.* To generate contact matrices for relevant models studied in this paper, we have simulated bead-spring polymer models consisting of $N$ beads each of size (diameter) $\sigma$. The total energy of the polymer is given by

$$E = \sum_{i=1}^{N-1} E_{\text{spring}}(|\vec{r}_{i+1} - \vec{r}_i|) + \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} E_{\text{LJ}}(|\vec{r}_j - \vec{r}_i|). \tag{3}$$

Here $\vec{r}_i$ is the position vector of $i$th bead. The first term represents the polymer connectivity through harmonic springs given as

$$E_{\text{spring}} = \frac{k}{2}(|\vec{r}_{i+1} - \vec{r}_i| - r_0)^2. \tag{4}$$

Here $r_0$ is the equilibrium distance of the spring and $k$ is the spring constant. The value of spring constant is $k = 10$ and $r_0 = 1.0$. The second term in Eq. (3) represents the interactions through Lennard-Jones potential given by

$$E_{\text{LJ}}(r) = \begin{cases} 4\epsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right] & r < r_c, \\ 0 & r \geqslant r_c. \end{cases} \tag{5}$$

Here $\epsilon$ is the energy parameter that determines the strength of the attractive interaction, $\sigma$ is the diameter of the bead, and $r_c$ is the cutoff for the potential. The Langevin dynamics simulations were performed using LAMMPS [54].

*Parameters for different polymer simulations.*

(1) *Self-avoiding walk (SAW).* In the SAW model, the Lennard-Jones (LJ) interaction strength was set to $\epsilon = 1.0$, with a cutoff radius of $r_c = 2^{1/6}\sigma$ to ensure purely repulsive interactions.

(2) *Random walk (RW).* In the RW model, we essentially simulated a Gaussian polymer where the interaction strength was set to $\epsilon = 0.0$, effectively eliminating LJ interactions.

(3) *Globules 1 and 2.* For the two globular configurations, the LJ interaction strengths were set to $\epsilon = 1.0$ (globule 1) and $\epsilon = 2.0$ (globule 2), respectively. A cutoff radius of $r_c = 2.5\sigma$ was used in both cases to include attractive interactions among all beads.

(4) *Globule 3 (compact globule).* For the compact globule (globule 3), the LJ interaction strength was set to $\epsilon = 2.0$ with a cutoff radius of $r_c = 2.5\sigma$. To increase the average contact probability per segment, LJ interactions between adjacent segments ($i$ and $i+1$) were excluded.

*Size of the polymer beads.* In generic polymer models such as the SAW or RW, a bead may represent a few kbp (e.g., 1, 5, or 10 kbp) of chromatin [42]. For the cell lines used in our study, IMR90 and K562, we adopt 10 kbp per bead, consistent with the resolution of the Hi-C datasets employed. We also note that, in principle, a larger coarse-grained bead could accommodate more than one DSB. However, for simplicity, we assume that at most one DSB can occur per bead, and we restrict our analysis to this scenario. The choice of coarse-graining scale will practically depend on the resolution of the Hi-C data and the DSB density, because we want the bead size to be larger or equal to the Hi-C resolution such that there is approximately one DSB for any bead.

## III. RESULTS

We present results from our translocation simulations and examine how DSB and contact probabilities influence the translocation probability. From Eqs. (1) and (2), the null expectation is that $k_{ij} = k$, a constant independent of $(i, j)$. We will show that this expectation does not hold and $k_{ij}$ is influenced by multiple parameters, given the statistical picture of translocation.

### A. $k_{ij}$ is not a constant but a function of chromatin compaction

To understand the role of chromatin compaction on $k_{ij}$, we simulated the following polymers ($N = 50$) with different contact probability matrices: these include random walk (RW), self-avoiding walk (SAW), different types of polymer globules, and 500-kbp domain from K562 and IMR90 cell lines, at 10-kbp resolution [see Model and Methods, Supplemental Material (SM) Fig. S1 [53]].

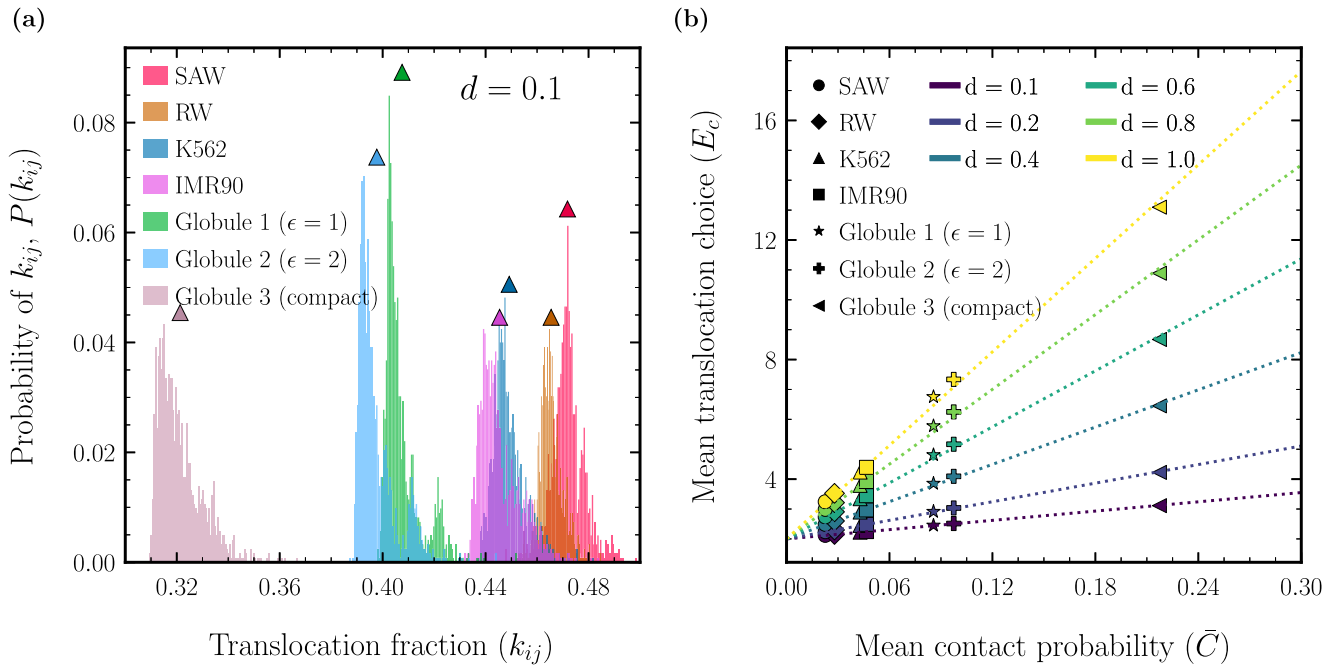


FIG. 2. Contrary to the prevalent notion, the translocation fraction ($k_{ij}$) is a function of polymer compaction: (a) Distribution of $k_{ij}$ for polymers with different compaction: SAW, RW, chromatin regions from K562 and IMR90 cell lines, three different globulelike polymers with different compaction (see Model and Methods, SM Fig. S1 [53]). Here, $d = 0.1$. The location of the mean translocation fraction for each polymer is indicated by triangle markers. (b) Mean translocation choice $E_c = 1/\bar{k}$ linearly increases with the mean contact probability ($\bar{C}$) for the same set of polymers as in (a). Each polymer type is represented by a distinct symbol corresponding to its characteristic mean contact probability.

Starting with a contact matrix $C_{ij}$ for each polymer, we introduced DSB and performed DSB repair and translocation as described in Model and Methods. We calculated $k_{ij}$ using Eq. (2). Figure 2(a) shows that the $k_{ij}$ is not a constant but has a distribution. We find that both the mean and the nature of the distribution depend on the type of polymer. Highly packed polymers such as the compact globule (beige color) show lower average $k_{ij}$ compared to open structures like the SAW (red) or RW (orange). We also observe a slight difference in the $k_{ij}$ distributions of SAW and RW polymer, consistent with the fact that RW is slightly more compact than SAW. We have simulated globules in a few different ways, and all of them have $k_{ij}$ values smaller than SAW and RW. Chromatin structures from the two cell lines (IMR90 and K562) yield $k_{ij}$ values that lie between these extremes. Notably, it is known that the chromatin region studied is slightly more compact in IMR90 than in K562 [44], and we observe a corresponding difference in the mean $k_{ij}$ values. All of these results indicate that chromatin compaction influences $k_{ij}$ and hence translocation probabilities. We define mean contact probability $\overline{C} = (1/N(N-1))\sum_{i=1}^{N}\sum_{j=1,j\neq i}^{N} C_{ij}$ as a measure of compaction and examine its relation with $k_{ij}$. Since we find an inverse dependence, we plot $E_c = 1/\bar{k}$ vs $\overline{C}$, showing a linear dependence [see Fig. 2(b)]. This quantifies how $\bar{k}$ varies for different polymers in Fig. 2(a). The intercept in Fig. 2(b) is 2; this is because, for a translocation to occur, there has to at least two choices—repairing the DSB or translocation among the two broken pairs. By fitting straight lines in Fig. 2(b), we obtain the following empirical relation:

$$\bar{k} = 1/(2 + \alpha(d)\overline{C}), \tag{6}$$

where $\alpha(d)$ is the slope, which is a function of DSB probability $d$.

### B. The translocation fraction varies with DNA double-strand break probabilities

To understand the dependence of translocation fraction ($k_{ij}$) on DSBs, we varied DSB probability ($d$) for a given contact probability matrix. Figure 3 shows how $k_{ij}$ varies for two very different polymers, SAW [Fig. 3(a)] and the compact globule [Fig. 3(b)], for various values of $d$. Assuming that each segment breaks with an equal probability $d$, we show that as $d$ increases, $k_{ij}$ decreases. Since increasing $d$ increases the number of possible choices for a segment $i$ to translocate, the fraction of translocation with any particular segment $j$ will fall. Hence the peak values of $k_{ij}$ and mean values $\bar{k}$ will depend on $d$ and compaction.

The observation that $\bar{k}$ is a function of the double-strand break probability $d$ implies that $T_{ij}$ is no longer proportional to $d^2$ but some other function of $d$. In Fig. 4 we plot the mean of $T_{ij}$ ($\overline{T}$ ) as a function of $d$, showing the deviation from the $d^2$ behavior (see also SM, Fig. S2 [53]). For a polymer like the compact globule, there is a larger deviation of $\overline{T}$ from the $d^2$ behavior as compared to a SAW polymer. As compact polymers have many contacts per segment, the limiting parameter for translocation is not the number of available contacts but the available DSBs determined by the parameter $d$. This implies that increasing $d$ will rapidly reduce $\bar{k}$ for such polymers (see Fig. 3). In the case of open polymers, the limiting parameter is not $d$ but the average number of contacts. Increasing $d$

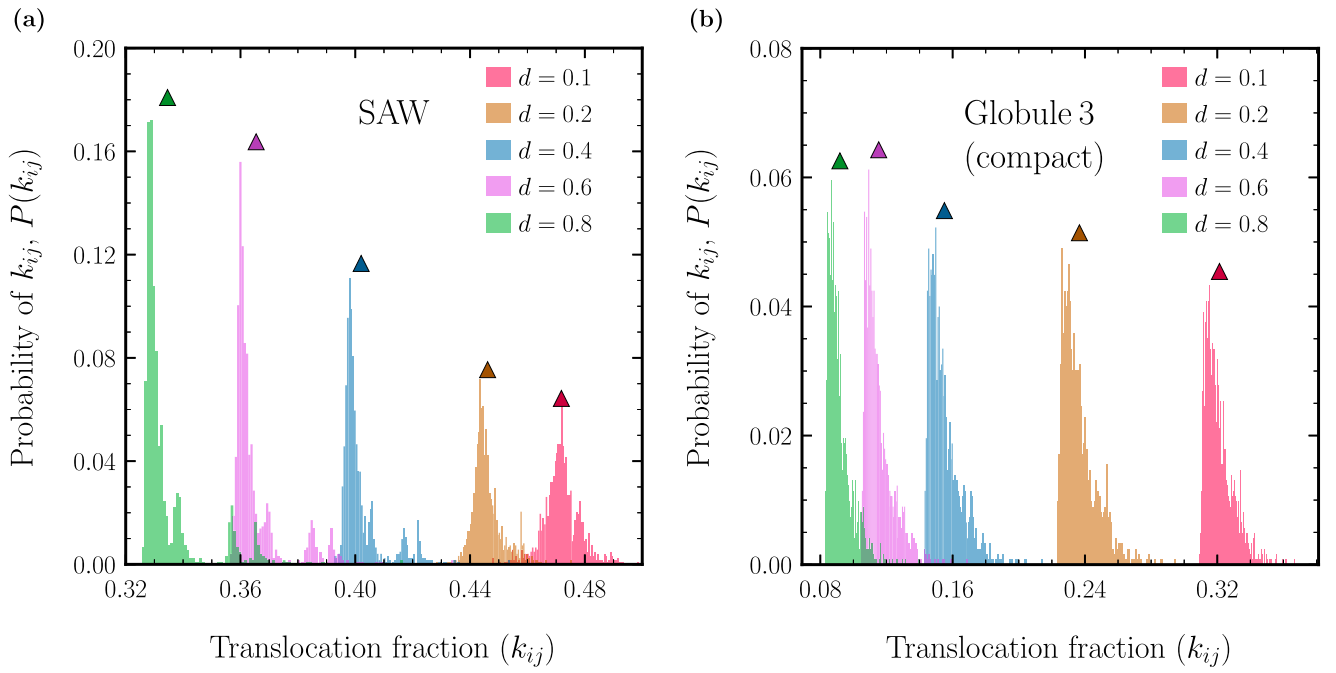


FIG. 3. Translocation fraction ($k_{ij}$) is a function of DSB probability ($d$). Distribution of $k_{ij}$ values from (a) SAW and (b) globule 3 (compact) for various $d$ values. Peak values of $k_{ij}$ reduce much less for higher $d$ in the SAW as compared to the compact globule due to its relatively more open structure. The location of the mean translocation fraction for each DSB probability ($d$) is indicated by triangle markers.

does little to increase the mean translocation choices ($E_c$) of the system. Hence, the dependence of their $1/\bar{k}$ values on $d$ is weaker. We also see that $E_c$ increases linearly with increasing $d$ (SM Fig. S3 [53]) as $E_c = (1/\bar{k}) = 2 + \beta(\bar{C})d$, suggesting

$$\bar{k} = 1/(2 + \beta(\bar{C})d), \tag{7}$$

where $\beta(\bar{C})$ is the slope for a particular polymer type.

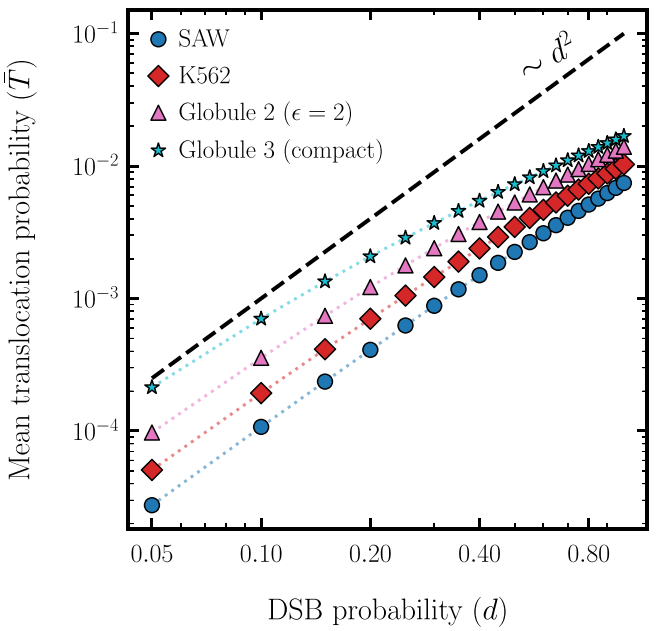


FIG. 4. Mean translocation probability ($\bar{T}$) need not be proportional to $d^2$, unlike the prevalent notion. Since $\bar{k}$ is a function of $d$, the translocation probability cannot be a function of $d^2$. The relatively open polymer SAW is closer to the $d^2$ power law, while more compact polymers deviate from $d^2$ (see also SM Fig. S2 [53]).

### C. Analytical expressions for $T_{ij}$ and $k_{ij}$

When we consider a polymer with two segments $i$ and $j$, the only translocation possible is between $i$ and $j$, and the probability is $T_{ij} = k_{ij}C_{ij}d_id_j$. As mentioned earlier, for the two-segment system $k_{ij} = \frac{1}{2}$. However, when one adds a third segment $m$, it is possible to have a breakage in segment $m$ with probability $d_m$ and contacts leading to translocations between $(i, m)$ and $(j, m)$. Accounting for all these possibilities (see SM [53]), we find that the translocation probability for a three-segment system is given by

$$T_{ij}(3) = C_{ij}d_id_j\left(\frac{d'_m}{2} + \frac{d_m}{2}C'_{im}C'_{jm} + \frac{13d_m}{36}\left(C_{im}C'_{jm} + C'_{im}C_{jm}\right) + \frac{5d_m}{18}C_{im}C_{jm}\right). \tag{8}$$

Here $C_{im}$ is the contact probability between segments $i$ and $m$, and $C_{jm}$ is the contact probability between segments $j$ and $m$; $d'_m = (1 - d_m)$ is the probability that segment $m$ is not broken. The $d'_m$ term represents the cases where only $i$ and $j$ are broken. The second term represents the case where $m$ is broken but neither $i$ nor $j$ is in contact with $m$ with $C'_{im} = (1 - C_{im})$ and $C'_{jm} = (1 - C_{jm})$. The third term represents the case where the segment $m$ is broken but only one of the pairs involving $m$ (either $im$ or $jm$) is in contact. The

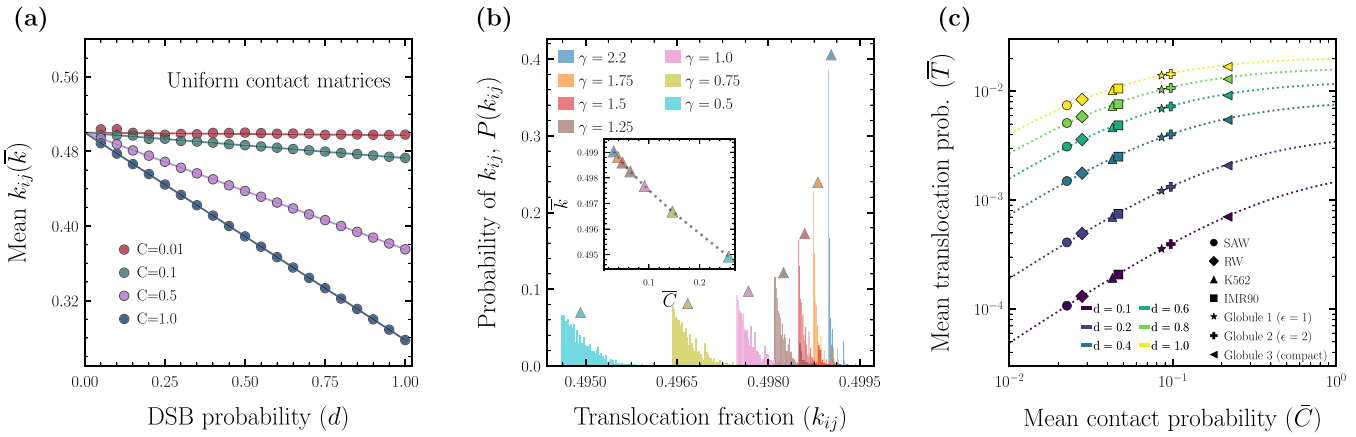


FIG. 5. (a) Mean $k_{ij}$ ($\bar{k}$) computed from the three-segment model as a function of DSB probability for a $3 \times 3$ contact matrix having all nondiagonal elements equal to 0.01, 0.1, 0.5, or 1.0. The solid lines are plotted using Eq. (10) and the points are generated from simulations. (b) Distribution of $k_{ij} = T_{ij}/(C_{ij}d_id_j)$ using Eq. (11), with $C_{ij} = |i-j|^{-\gamma}$ and $d = 3/N$, for various $\gamma$ values. Triangles mark the mean value $\bar{k}$. Inset: $\bar{k}$ vs $\bar{C}$ from the analytical theory reveals that higher chromatin compaction (larger $\bar{C}$) leads to reduced $\bar{k}$, highlighting the role of multibody contacts in $T_{ij}$ estimates. (c) Mean translocation probability ($\bar{T}$) as a function of mean contact probability ($\bar{C}$) from Eqs. (13) and (14) with $d_i = d_j = d$ for $N = 50$ (dotted lines). Points represent $\bar{T}$ calculated from the simulation for different polymers as described in Fig. 2(a).

last term represents the case where all three segments are broken and in contact with each other. The counting of all possible combinations of these gives the numerical prefactors as explained in the SM [53]. Simplifying by substituting for the primed quantities gives us

$$T_{ij}(3) = C_{ij}d_id_j\left[\frac{1}{2} - \frac{d_m}{36}(5(C_{im}+C_{jm}) - 2C_{im}C_{jm})\right]. \quad (9)$$

This is equivalent of defining

$$k_{ij} = \tfrac{1}{2} - (d_m/36)[5(C_{im}+C_{jm}) - 2C_{im}C_{jm}]. \quad (10)$$

We also find that the maximum value of $k_{ij}$ (upper limit) can be 1/2, which occurs when $d_m \to 0$ or $C_{im} \to 0$ and $C_{jm} \to 0$. We performed a translocation simulation for a three-segment system and obtained mean values of $k_{ij}$ to show that our analytical relation [Eq. (10)] matches well with the simulation results [Fig. 5(a)] for simple cases where $C_{im} = C_{jm}$= constant.

We can extend the analytical relation in Eq. (9) for the case of a polymer with $N$ segments by expanding the $T_{ij}$ in terms of the number of breaks. In principle, the expansion will have terms up to $\mathcal{O}(d^N)$. However, if we assume that four or higher breaks are improbable as well as higher-order multibead contacts are improbable, we can rewrite the equation as follows:

$$T_{ij}(N) = \left(\frac{C_{ij}d_id_j}{N-2}\right)\sum_{\substack{m=1\\ m\neq i,j}}^{N}\left(\frac{d'_m}{2} + \frac{d_m}{2}C'_{im}C'_{jm} + \frac{13d_m}{36}(C_{im}C'_{jm} + C'_{im}C_{jm}) + \frac{5d_m}{18}C_{im}C_{jm}\right). \quad (11)$$

Here two out of these three breaks should be on $i$ and $j$ so that we can compute $T_{ij}$; however, the third break can be on any of the other $N-2$ beads. We computed the $T_{ij}(N)$ from the above equation by taking $C_{ij} = |i-j|^{-\gamma}$ and $d = (3/N)$. Using Eq. (2) we get $k_{ij}$, and its distribution is plotted in Fig. 5(b) for several $\gamma$ values. As the compaction increases (as $\gamma$ decreases), the peak of the distribution shifts away from 0.5, as predicted by our simulations. Here the trend is similar to what is observed in Fig. 2(a), where full simulations without approximation produce $k_{ij}$ distributions that shift away from 0.5 as the compaction increases. We emphasize that the goal of this section is not to determine the exact functional form of $k_{ij}$ but rather to show analytically that $k_{ij}$ is not constant and indeed varies with compaction.

We can also derive a lower bound for $\bar{k}$ in a theoretical limit when all segments are broken ($d_i = 1$ for all $i$) and all segments are in contact with each other ($C_{ij} = 1$ for all $i$ and $j$). For an $N$-segment polymer, we derive a recursive relation for the lower bound of $\bar{k}$ as

$$\bar{k}_{\min}(N) = \frac{2[1 + \binom{N-2}{2}\bar{k}_{\min}(N-2)] + \binom{N-2}{1}\bar{k}_{\min}(N-1)}{N^2}, \quad (12)$$

where $\bar{k}_{\min}(N-2)$ and $\bar{k}_{\min}(N-1)$ are the lower limits of $\bar{k}$ for an $N-2$ and $N-1$ long polymers, respectively (see SM [53]). $\bar{k}_{\min}(2)$ and $\bar{k}_{\min}(3)$ are 0.5 (two choices on average) and 0.2778, respectively. It turns out that $\lim_{N\to\infty}\bar{k}_{\min} \approx \frac{1}{N}$ (see SM Fig. S4 [53]). Using the relations (6), (7), and (12), we can obtain an empirical formula for the mean value $\bar{k}$ as

$$\bar{k}\,(N, \bar{C}, d) \approx 1/(2 + (N-2)\bar{C}d) \quad (13)$$

and the corresponding translocation probability as

$$T_{ij}\,(N) \approx \bar{k}C_{ij}d_id_j. \quad (14)$$

We show that this empirical formula matches well with the data we obtained from our simulations for different polymers [Fig. 5(c) and SM Fig. S5 [53]].

## IV. DISCUSSION AND CONCLUSION

One important insight from this study is that regions with more contacts, like heterochromatin regions or TAD boundaries, will have low $k_{ij}$ values. This can be seen in SM

Fig. S6 [53] and shows that TAD boundaries have lesser $k_{ij}$ values as compared to other regions. An interesting notion emerging from the results of these simulations is that, due to excessive packaging in compact regions, each segment has many "decoy"-like segments to which it can potentially translocate. This reduces the probability of important sites (active promoters, enhancers) translocating or moving to undesirable segments. In this way, harmful translocations may be suppressed by chromatin packaging, in addition to other biological mechanisms. Whether this effect is biologically significant remains unknown.

Our work has various assumptions. We assume a uniform DSB probability, motivated by the experimental conditions, where DSBs are induced by gamma irradiation. Under such controlled induction, the spatial heterogeneity in break formation is substantially reduced, making a uniform probability a reasonable approximation [37]. However, under endogenous conditions, we assume DSBs are not uniformly distributed along the genome and that certain regions are known to exhibit elevated break frequencies ("DSB hotspots"). Another reason for neglecting heterogeneity in DSB breaks is that our primary aim in this work is to study the effect of chromatin compaction on translocation outcomes. Introducing heterogeneous, locus-dependent DSB probabilities would have added additional parameters. Incorporating realistic, nonuniform DSB landscapes would be an interesting extension.

One drawback of this model is that it does not provide information about what happens after translocations occur. That is, we likely cannot use the same contact matrix after one realization if we wish to study the time evolution of translocation probabilities. Also, the model takes into account all translocation and rearrangement events, including deletions and circularization, which may not be favorable for species and may not be seen in living populations. Equation (14) only stands when all events are purely statistical.

To conclude, we investigated how the 3D polymer organization affects the translocation probability. We simulated different types of polymers with different 3D organization and compaction. From the simulation we obtained a functional relation between translocation probability, contact probability, and double-strand DNA break probability. Naive expectation from the simplest possible model is that $T_{ij} \propto d_i d_j C_{ij}$, as shown in Eq. (1). However, we show that translocation probability depends not only on the pairwise contact probability $C_{ij}$ but also exhibits a nontrivial dependence that reflects the multibead contacts, the overall compaction of the polymer, and the DSB probability. Using our simulation data, we derived an empirical relation for the mean translocation probability that is particularly relevant under conditions of high DSB rates and high chromatin compaction.

## ACKNOWLEDGMENTS

R.P. acknowledges funding from Anusandhan National Research Foundation (Grant No. CRG/2022/007974) and Sunita Sanghi Centre of Aging and Neurodegenerative Diseases (SCAN), IIT Bombay. S.D. acknowledges fellowship support from the Prime Minister's Research Fellowship, Ministry of Education, India.

## DATA AVAILABILITY

The data that support the findings of this article are openly available [55].

[1] T. Misteli, Beyond the sequence: Cellular organization of genome function, Cell **128**, 787 (2007).
[2] G. Cavalli and T. Misteli, Functional implications of genome topology, Nat. Struct. Mol. Biol. **20**, 290 (2013).
[3] K. Maeshima, S. Iida, and S. Tamura, Physical nature of chromatin in the nucleus, Cold Spring Harbor Perspect. Biol. **13**, a040675 (2021).
[4] B. Alberts, D. Bray, J. Lewis, M. Raff, K. Roberts, and J. D. Watson, *Molecular Biology of the Cell*, 4th ed. (Garland Science, New York, 2002).
[5] D. M. Ibrahim and S. Mundlos, The role of 3D chromatin domains in gene regulation: A multi-facetted view on genome organization, Curr. Opin. Genet. Dev. **61**, 1 (2020).
[6] D. G. Lupiáñez, M. Spielmann, and S. Mundlos, Breaking TADs: How alterations of chromatin domains result in disease, Trends Genet. **32**, 225 (2016).
[7] A. Groth, W. Rocha, A. Verreault, and G. Almouzni, Chromatin challenges during DNA replication and repair, Cell **128**, 721 (2007).
[8] C. P. Spampinato, Protecting DNA from errors and damage: An overview of DNA repair mechanisms in plants compared to mammals, Cell. Mol. Life Sci. **74**, 1693 (2017).
[9] C. Arnould, V. Rocher, F. Saur, A. S. Bader, F. Muzzopappa, S. Collins, E. Lesage, B. Le Bozec, N. Puget, T. Clouaire, *et al.*, Chromatin compartmentalization regulates the response to DNA damage, Nature (London) **623**, 183 (2023).
[10] P. Caron, F. Aymard, J. S. Iacovoni, S. Briois, Y. Canitrot, B. Bugler, L. Massip, A. Losada, and G. Legube, Cohesin protects genes against $\gamma$H2AX induced by DNA double-strand breaks, PLoS Genet. **8**, e1002460 (2012).
[11] B. Bonev and G. Cavalli, Organization and function of the 3D genome, Nat. Rev. Genet. **17**, 661 (2016).
[12] A. Pombo and N. Dillon, Three-dimensional genome architecture: Players and mechanisms, Nat. Rev. Mol. Cell Biol. **16**, 245 (2015).
[13] T. Misteli, The self-organizing genome: Principles of genome architecture and function, Cell **183**, 28 (2020).
[14] J. Dekker, K. Rippe, M. Dekker, and N. Kleckner, Capturing chromosome conformation, Science **295**, 1306 (2002).
[15] S. Sati and G. Cavalli, Chromosome conformation capture technologies and their impact in understanding genome function, Chromosoma **126**, 33 (2017).
[16] J. Han, Z. Zhang, and K. Wang, 3C and 3C-based techniques: The powerful tools for spatial genome organization deciphering, Mol. Cytogenet. **11**, 21 (2018).
[17] 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).
[18] S. S. P. Rao, M. H. Huntley, N. C. Durand, E. K. Stamenova, I. D. Bochkov, J. T. Robinson, A. L. Sanborn, I. Machol, A. D. Omer, and E. S. Lander, *et al.*, A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping, Cell **159**, 1665 (2014).
[19] J.-M. Belton, R. P. McCord, J. H. Gibcus, N. Naumova, Y. Zhan, and J. Dekker, Hi–C: A comprehensive technique to capture the conformation of genomes, Methods **58**, 268 (2012).
[20] J. R. Dixon, D. U. Gorkin, and B. Ren, Chromatin domains: The unit of chromosome organization, Mol. Cell **62**, 668 (2016).
[21] E. P. Nora, B. R. Lajoie, E. G. Schulz, L. Giorgetti, I. Okamoto, N. Servant, T. Piolot, N. L. Van Berkum, J. Meisig, J. Sedat, *et al.*, Spatial partitioning of the regulatory landscape of the X-inactivation, Nature (London) **485**, 381 (2012).
[22] A. L. Sanborn, S. S. P. 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).
[23] I. M. Sokolov, J. Mai, and A. Blumen, Paradoxal diffusion in chemical space for nearest-neighbor walks over polymer chains, Phys. Rev. Lett. **79**, 857 (1997).
[24] K. E. Polovnikov, H. B. Brandão, S. Belan, B. Slavov, M. Imakaev, and L. A. Mirny, Crumpled polymer with loops recapitulates key features of chromosome organization, Phys. Rev. X **13**, 041029 (2023).
[25] N. Naumova, M. Imakaev, G. Fudenberg, Y. Zhan, B. R. Lajoie, L. A. Mirny, and J. Dekker, Organization of the mitotic chromosome, Science **342**, 948 (2013).
[26] S. Dutta, A. Rajakumar, R. Padinhateeri, and M. K. Mitra, Compaction of chromatin domains regulates target search times of proteins, PLoS Comput. Biol. **22**, e1013843 (2026).
[27] T. Clouaire and G. Legube, A snapshot on the cis chromatin response to DNA double-strand breaks, Trends Genet. **35**, 330 (2019).
[28] J. H. Yang, H. B. Brandão, and A. S. Hansen, DNA double-strand break end synapsis by DNA loop extrusion, Nat. Commun. **14**, 1913 (2023).
[29] Z. Chen and J. K. Tyler, The chromatin landscape channels DNA double-strand breaks to distinct repair pathways, Front. Cell Dev. Biol. **10**, 909696 (2022).
[30] M. Nambiar and S. C. Raghavan, Chromosomal translocations among the healthy human population: Implications in oncogenesis, Cell. Mol. Life Sci. **70**, 1381 (2013).
[31] N. R. Pannunzio, G. Watanabe, and M. R. Lieber, Nonhomologous DNA end-joining for repair of DNA double-strand breaks, J. Biol. Chem. **293**, 10512 (2018).
[32] S. Gao, S. Honey, B. Futcher, and A. P. Grollman, The non-homologous end-joining pathway of *S. cerevisiae* works effectively in G1-phase cells, and religates cognate ends correctly and non-randomly, DNA Repair **42**, 1 (2016).
[33] V. Roukos, T. C. Voss, C. K. Schmidt, S. Lee, D. Wangsa, and T. Misteli, Spatial dynamics of chromosome translocations in living cells, Science **341**, 660 (2013).
[34] J. Zagelbaum, A. Schooley, J. Zhao, B. R. Schrank, E. Callen, S. Zha, M. E. Gottesman, A. Nussenzweig, R. Rabadan, J. Dekker, *et al.*, Multiscale reorganization of the genome following DNA damage facilitates chromosome translocations via nuclear actin polymerization, Nat. Struct. Mol. Biol. **30**, 99 (2023).
[35] P. J. Wijchers and W. de Laat, Genome organization influences partner selection for chromosomal rearrangements, Trends Genet. **27**, 63 (2011).
[36] M. Schwartz and O. Hakim, 3D view of chromosomes, DNA damage, and translocations, Curr. Opin. Genet. Dev. **25**, 118 (2014).
[37] Y. Zhang, R. P. McCord, Yu.-J. Ho, B. R. Lajoie, D. G. Hildebrand, A. C. Simon, M. S. Becker, F. W. Alt, and J. Dekker, Spatial organization of the mouse genome and its role in recurrent chromosomal translocations, Cell **148**, 908 (2012).
[38] J. M. Engreitz, V. Agarwala, and L. A. Mirny, Three-dimensional genome architecture influences partner selection for chromosomal translocations in human disease, PLoS ONE **7**, e44196 (2012).
[39] R. Chiarle, Y. Zhang, R. L. Frock, S. M. Lewis, B. Molinie, Yu-J. Ho, D. R. Myers, V. W. Choi, M. Compagno, D. J. Malkin, *et al.*, Genome-wide translocation sequencing reveals mechanisms of chromosome breaks and rearrangements in B cells, Cell **147**, 107 (2011).
[40] C.-S. Lee, R. W. Wang, H.-H. Chang, D. Capurso, M. R. Segal, and J. E. Haber, Chromosome position determines the success of double-strand break repair, Proc. Natl. Acad. Sci. USA **113**, 146 (2016).
[41] F. W. Alt, Y. Zhang, F.-L. Meng, C. Guo, and B. Schwer, Mechanisms of DNA lesions and genomic instability in the immune system, Cell **152**, 417 (2013).
[42] S. Kadam, K. Kumari, V. Manivannan, S. Dutta, M. K. Mitra, and R. Padinhateeri, Predicting scale-dependent chromatin polymer properties from systematic coarse-graining, Nat. Commun. **14**, 4108 (2023).
[43] Y. Qi and B. Zhang, Chromatin network retards nucleoli coalescence, Nat. Commun. **12**, 6824 (2021).
[44] K. Kumari, J. R. Prakash, and R. Padinhateeri, Heterogeneous interactions and polymer entropy decide organization and dynamics of chromatin domains, Biophys. J. **121**, 2794 (2022).
[45] G. Shi and D. Thirumalai, From Hi-C contact map to three-dimensional organization of interphase human chromosomes, Phys. Rev. X **11**, 011051 (2021).
[46] S. Nath, S. Dutta, S. D. Sarkar, D. Sengupta, M. K. Mitra, and K. Sengupta, Liquid–liquid phase separation of lamin drives altered chromatin organization in cardiomyopathic mutations of lamin A, Nucleic Acids Res. **53**, gkaf615 (2025).
[47] N. Haddad, D. Jost, and C. Vaillant, Perspectives: Using polymer modeling to understand the formation and function of nuclear compartments, Chromosome Res. **25**, 35 (2017).
[48] M. Nicodemi and A. Pombo, Models of chromosome structure, Curr. Opin. Cell Biol. **28**, 90 (2014).
[49] G. Forte, S. Buonomo, P. R. Cook, N. Gilbert, D. Marenduzzo, and E. Orlandini, Modeling the 3D spatiotemporal organization of chromatin replication, PRX Life **2**, 033014 (2024).
[50] Y. Zhang, L. Boninsegna, M. Yang, T. Misteli, F. Alber, and J. Ma, Computational methods for multiscale 3D genome organization, Nat. Rev. Genet. **25**, 123 (2024).
[51] S. Brahmachari, V. G. Contessoto, M. D. Pierro, and J. N. Onuchic, Shaping the genome via lengthwise compaction,

phase separation, and lamina adhesion, Nucleic Acids Res. **50**, 4258 (2022).

[52] K. Guin, Y. Chen, R. Mishra, S. R. B. M. Muzaki, B. C. Thimmappa, C. E. O'Brien, G. Butler, A. Sanyal, and K. Sanyal, Spatial inter-centromeric interactions facilitated the emergence of evolutionary new centromeres, Elife **9**, e58556 (2020).

[53] See Supplemental Material at http://link.aps.org/supplemental/10.1103/kyds-xmlh for additional figures and calculations.

[54] A. P. Thompson, H. M. Aktulga, R. Berger, D. S. Bolintineanu, W. M. Brown, P. S. Crozier, P. J. in't Veld, A. Kohlmeyer, S. G. Moore, T. D. Nguyen, R. Shan, M. J. Stevens, J. Tranchida, C. Trott, and S. J. Plimpton, LAMMPS – A flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales, Comput. Phys. Commun. **271**, 108171 (2022).

[55] Simulation code and data: https://github.com/ShuvadipD/translocation.