Protein search processes mediated by chromatin topology

Shuvadip Dutta[1], Adarshkrishnan R.[1], Ranjith Padinhateeri[2],[*] and Mithun K. Mitra[1][†]

[1]*Department of Physics, Indian Institute of Technology Bombay, Mumbai 400076, India*

[2]*Department of Biosciences and Bioengineering,*

Indian Institute of Technology Bombay, Mumbai 400076, India

Abstract

We investigate the role of compaction of chromatin domains in modulating search kinetics of proteins. Collapsed conformations of chromatin, characterised by long loops which bring distant regions of the genome into contact, and manifested structurally as Topologically Associated Domains (TADs) affect search kinetics of DNA associated transcription factors and other proteins. In this study, we investigate the role of the compactness of chromatin on the dynamics of proteins using a minimal model. Using theory and simulations, we show that an optimal compaction exists for which the residence time of proteins on a chromatin-like polymer backbone is minimum. We then show that the fraction of unique sites visited also shows a similar non-monotonic behavior as a function of the polymer compaction. We extend these results to more detailed polymer models - using the Freely Rotating Chain model, a Lennard-Jones bead-spring polymer, and a soft polymer chain model, which approximates chromatin behavior. We show that our results continue to hold for these polymer models, with a minimum residence time at an optimum polymer compaction. Finally, we also analyse the dynamics of proteins on networks generated using experimental chromatin conformation data for several TADs extracted from human chromosomes. Our analysis suggests that TADs exist near this zone of optimality, indicating that chromatin conformations can play a crucial role in modulating protein search strategies.

[*] ranjithp@iitb.ac.in

[†] mkmitra@iitb.ac.in

I. INTRODUCTION

The mechanism of target search by DNA-associated proteins such as Transcription Factors (TFs) on DNA backbones remains an open question [1–7]. Simple three-dimensional diffusion to locate specific binding sites cannot account for search timescales, necessitating proposals of facilitated diffusion processes [8–16]. Four distinct search mechanisms have been proposed to explain the dynamic motion of proteins along DNA: 1D sliding, hopping or jumping, 3D diffusion, and intersegmental transfer [17]. A combination of 3D and 1D diffusion has been shown to lead to optimal search times compared to a pure 1D or a pure 3D search [7–9]. However, while some studies have investigated experimental evidence for a combination of bulk 3D diffusion and 1D sliding in prokaryotes [18, 19] – evidence in mammalian organisms remains scant [20]. Indeed, experimental evidence points to a preponderance of sliding modes of transport - for different proteins such as T7 RNA polymerases, E. coli RNA polymerase, as well as DNA repair proteins such as Rad51, hOgg1 and LacI repressor, and the eukaryotic protein p53 [12, 21–25]. In addition, intersegmental hops or "direct transfers" have been observed experimentally [26], especially in densely packed nucleoids. This assumes particular significance for protein motion in mammalian cells, where highly packed chromatin regions with high density and viscosity [27], hinder free diffusion in the solution phase, and multivalent proteins, due to nonspecific interactions, slow the mobility of other searching proteins, limiting their capacity for free 3D diffusion [28].

In this crowded milieu, intersegmental transfer from one segment of DNA to another via long loops assumes particular significance as a mechanism for effective search, which does not include free diffusion in the bulk [17]. While most studies assume that intersegmental transfer moves the protein to an uncorrelated segment, and thus has the same effect as 3D diffusion [2], this is not true for densely packed chromatin, where backbone correlations can persist over long length scales, and impact the intersegmental transfer dynamics. Numerous theoretical and experimental studies have delved into the hierarchical organization of the genome across various length scales [29–52]. In mammals, the genome is organised into compact regions, such as Topologically Associating Domains (TADs) [53–55] (Fig. 1A). TADs are structural and functional units of the genome characterised by large squares of increased contact frequency tiling the contact matrices' diagonal [54]. Concurrent with this high intra-domain contacts inside TADs, is a relative insulation among adjacent domains

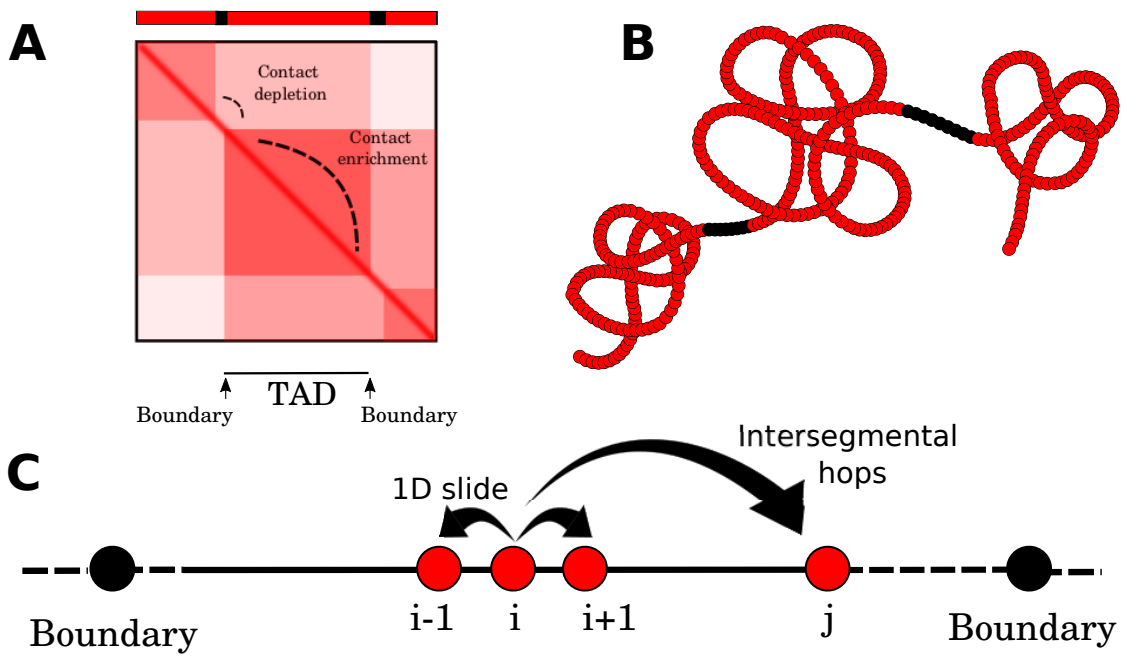

FIG. 1. **A.** A schematic of a Hi-C map, showing TADs as square blocks of increased interaction frequencies along the diagonal. Neighbouring TADs are separated by contact depleted zones. **B.** Polymer representation of chromatin with compact TAD regions (red) separated by relatively extended boundary regions (black). **C.** 1D model for TAD – each small bead is connected via bonds to its neighbour. there are long bonds which maintains contact between 3D neighbours. Black beads denote boundaries of the TAD.

through TAD boundaries (Fig. 1B) [34, 56]. TAD boundaries, marked by the enrichment of H3K4Me3 and the presence of abundant Pol2 [57], serve as hubs where the promoter regions of numerous transcriptional units are strategically located. These enhanced contacts within TADs result in a higher occurrence of long intersegmental hops, consequently influencing the kinetics of the search processes of proteins on chromatin [58]. Indeed, search processes on polymer backbones have shown surprising results for protein kinetics [13], both in presence and absence of polymer fluctuations [59, 60]. Looped structures such as rosettes have been predicted to affect search kinetics [61]. In addition, DNA conformations and dynamics, protein-crowder-DNA interactions, DNA coverage by other proteins, crowder concentration and confinement effects impact search kinetics [62–68].

In this article, we investigate the role of 3D chromatin organization, in particular for compact TAD-like regions, in the target search process by proteins. We show how the

residence times of proteins on TAD-like domains and the efficiency of the search process, described by fraction of unique sites visited, crucially depend on the topology of polymer domains characterised by long-range intersegmental contacts.

II. A UNIFORM NETWORK REPRESENTATION OF CHROMATIN

We first consider a contact space network representation of polymer. As a minimal model we assume any two non-neighbouring beads are in contact with a uniform probability p (see Fig. 1C and SI Sec. 1.1). A protein can slide along the polymer backbone and jump between non-neighbouring beads via the contacts, with equal probabilities (see SI Sec. 1.2). We compute the configuration-averaged mean residence times $\langle T_{\rm res}\rangle$, defined as the time a protein takes to reach either boundaries of the polymer for the first time, starting from the mid-point (see SI Sec. 1.3). This is equivalent to target search kinetics, with targets at the TAD boundaries. We show an interesting non-monotonic relation between $\langle T_{\rm res}\rangle$ and p (Fig. 2A). When $p = 0$, the protein performs a 1D random walk, giving $\langle T_{\rm res}\rangle = \frac{L^2}{8D} \equiv T_d$ [69]. As p increases, more paths to reach the boundaries emerge, reducing $\langle T_{\rm res}\rangle$. However, beyond a threshold connectivity, $\langle T_{\rm res}\rangle$ increases due to excess intersegmental bonds diverting the walker. In Fig. 2B, we show that this non-monotonic behavior of $\langle T_{\rm res}\rangle$ persists even when sliding along the backbone and jumps along intersegmental bonds have different rates (see SI Fig. S1). At low p, the $\langle T_{\rm res}\rangle$ remains independent of $t_{\rm jump}/t_{\rm slide}$, dominated by 1D sliding. Conversely, for higher p, $\langle T_{\rm res}\rangle$ decrease as the $t_{\rm jump}/t_{\rm slide}$ rises, favoring long hops over sliding. This non-monotonic behavior of $\langle T_{\rm res}\rangle$ remains robust even when the polymer network undergoes dynamic re-configurations (see SI Sec. 1.3 and SI Fig. S2) at a fixed connection probability (Fig. 2A (red symbols) and Fig. 2C). Dynamic rewiring of network connections reduces $\langle T_{\rm res}\rangle$, with faster configuration changes yielding lower $\langle T_{\rm res}\rangle$ at low p values. However, at high p values, network dynamicity offers little advantage, as most nodes already have numerous long-range connections. The optimal behavior of $\langle T_{\rm res}\rangle$ is consistent for all polymer lengths, as shown SI Fig. S3A.

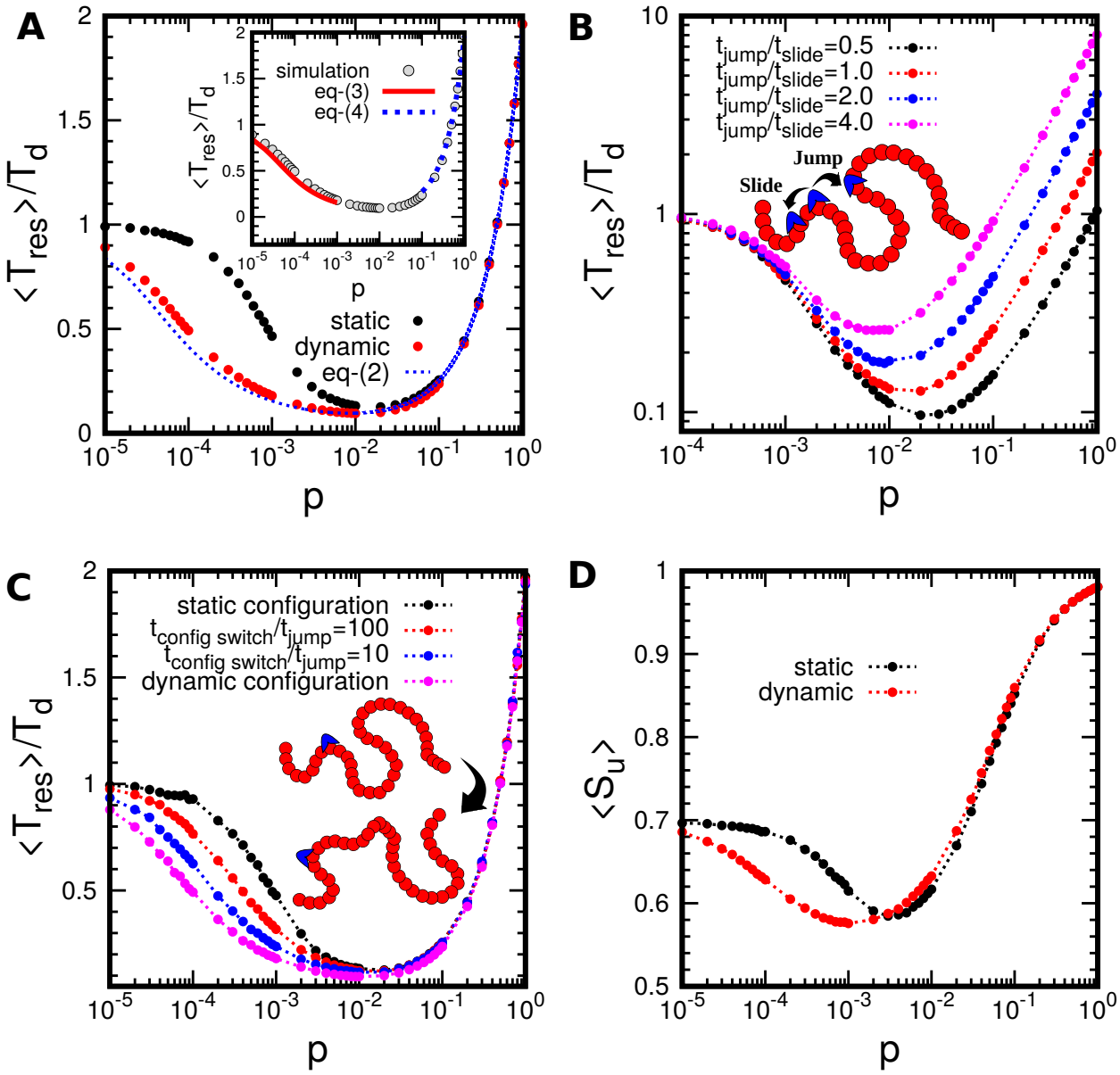

FIG. 2. **A.** Scaled mean residence time ($\frac{\langle T_{\rm res}\rangle}{T_d}$) in uniformly connected network domain vs probability of contact (p) between any two non-neighbouring beads for $L = 100$. Static (black) and dynamic configuration (red) show different behaviour in small p limit. In high p limit, both curves coincide with each other. The solid line is the analytical solution of Eq. 2. The inset shows the lwo-p and high-p approximations of Eqs. 3 and 4. **B.** $\frac{\langle T_{\rm res}\rangle}{T_d}$ in uniformly connected network domain vs p for different ratios of $t_{\rm jump}/t_{\rm slide}$ for $L = 100$. **C.** $\frac{\langle T_{\rm res}\rangle}{T_d}$ in uniformly connected network domain vs p for different ratios of $t_{\rm config-switch}/t_{\rm jump}$ for $L = 100$. **D.** Fraction of unique sites visited before exit in uniformly connected network domain ($\langle S_u\rangle$) vs p for $L = 100$ in static configuration (black) and dynamic configuration (red).

For the case of the dynamic polymer, where the polymer undergoes rewiring at a fixed connection probability p, we can solve for the mean residence times analytically. The mean residence times($T_k(p)$) for a protein starting from the kth site for obeys,

$$T_k = \tau + r(p)\left(T_{k+1} + T_{k-1}\right) + \sum_{j \neq nn} s(p) T_j \tag{1}$$

where, τ is the time taken for a single step,; $r(p) = \dfrac{1}{2 + p(L-3)}$ and $s(p) = p\ r(p)$ (see SI Sec. 2.1) are slide and jump probabilities, respectively. This can be solved exactly, subject to the absorbing boundary conditions at the two ends to yield (see SI Sec. 2.1),

$$\langle T_{\text{res}} \rangle = \frac{\beta}{\alpha} + 2e^{\frac{L}{2}\sqrt{\alpha}} \left(\frac{\phi_1 \tau + \frac{\phi_2 \beta}{\alpha}}{\phi_3} \right) \tag{2}$$

where

$$\alpha = \frac{1 + 3s - 2r}{r - s}$$

$$\beta = \frac{\tau \left(1 + 2s(1+s) + \frac{2s\phi_1\psi(1+2s)}{\phi_3} + 2s(r-s)\frac{\phi_1\phi_4}{\phi_3} \right)}{r - s - \frac{s(L-3)(1+2s)}{\alpha} - \frac{2s\phi_2\psi(1+2s)}{\alpha\phi_3} - \frac{2s(r-s)}{\alpha}\left(1 + \frac{\phi_2\phi_4}{\phi_3}\right)}$$

$$\phi_1 = 1 + (r+s)(1+s)$$

$$\phi_2 = (1 + r + s)s(L-3) + r^2 + r - (s+1)^2$$

$$\phi_3 = \left[1 + s - r^2 + s^2\right]\left(e^{2\sqrt{\alpha}} + e^{(L-2)\sqrt{\alpha}}\right) - 2\psi s(1 + r + s) - (r - s)\left(e^{3\sqrt{\alpha}} + e^{(L-3)\sqrt{\alpha}}\right)$$

$$\phi_4 = e^{2\sqrt{\alpha}} + e^{(L-2)\sqrt{\alpha}}$$

$$\psi = \left[e^{2\sqrt{\alpha}} \left(\frac{e^{(L-3)\sqrt{\alpha}} - 1}{e^{\sqrt{\alpha}} - 1} \right) \right]$$

This matches exactly with the simulation results (blue dashed line in Fig 2A). For low values of p, the governing equation can be simplified to obtain a solution (see SI Sec. 2.2),

$$\begin{aligned} \frac{\langle T_{\text{res}} \rangle}{T_d} &\overset{p \to 0}{=} \frac{4\tau}{L^2(\alpha(r-s) - s\gamma)} \left[1 - \frac{2e^{\frac{L}{2}\sqrt{\alpha}}}{1 + e^{L\sqrt{\alpha}}} \right] \\ &\approx 1 - \frac{(L^3 - 5L^2 - 20L + 48)}{48} p \quad \text{for } p \ll \frac{1}{L^2} \end{aligned} \tag{3}$$

a decreasing function of p. At high connection probability, as $p \to 1$, all sites become equivalent, and a simplified governing equation can be solved to yield (see SI Sec. 2.3),

$$\langle T_{\text{res}} \rangle \overset{p\to 1}{=} 2 + (L-3)p + \frac{(2+(L-4)p)^2}{2p} \tag{4}$$

which implies a linear increase with p as $p \to 1$, also consistent with a solution using the effective medium approach [70] (see SI Sec. 2.4 and SI Fig. S5). As can be seen from Fig 2A inset, the analytic expression at both small and high p limits matches the simulation results.

If a protein enters a domain and wants to search for a specific target sequence, a relevant quantity is the fraction of distinct sites $\langle S_u \rangle$ visited by the protein, before exiting the domain through the boundaries. At $p = 0$ (1D walk), we have $\langle S_u \rangle \simeq 0.698$ [71]. With increasing p, the intersegmental bonds increase, and the protein can escape the domain using long-range hops, leading to a decrease of $\langle S_u \rangle$. As $p \to 1$, the protein remains trapped due to an excess number of bonds and $\langle S_u \rangle \to 1$ [71]. This predicts a non-monotonic behaviour for $\langle S_u \rangle$ for both static and dynamic networks, as shown in Fig. 2D. Interestingly, the connection probability at which $\langle S_u \rangle$ is minimum is different from the optimal connectivity for $\langle T_{\text{res}} \rangle$. Indeed, around the region of minimum $\langle T_{\text{res}} \rangle$, $\langle S_u \rangle$ is high, indicating that the search process is relatively efficient around this connectivity. The non-monotonic behavior of $\langle S_u \rangle$ is consistent regardless of the polymer length, as shown in the contour plot in SI Fig. S3B.

III. POLYMER REPRESENTATIONS OF CHROMATIN

We now examine whether these effects persist when canonical polymer models are used to model chromatin structure. Thus the allowed intersegmental jumps arise directly from polymer topology, and hence random walks on these backbones have direct relevance to protein search processes.

A. Freely Rotating Chain Model

We first study the Freely Rotating Chain (FRC) model which has been used to investigate chromatin behaviour [72–75]. Neighbouring bonds in an FRC polymer are constrained to lie on a cone subtending an angle θ [76]. The angle θ controls the compaction of the

polymer, with large θ leading to a collapsed state (see SI Fig. S6A). As the polymer collapses, the number of non-neighbouring bonds N_b increases monotonically (see SI Fig. S6B), indicating a greater probability of intersegmental jumps. We simulated search on these FRC backbones (see SI Sec. 3.1), and the mean residence times $\langle T_{\text{res}} \rangle$ shows a non-monotonic behaviour with increasing bond angle (Fig. 3A). This can be also interpreted as a function of polymer compactness or collapse. A dimensionless quantity L/R_g indicates compactness of the polymer domain, and residence times show a non-monotonic behaviour with increasing polymer compactness (Fig. 3A inset). Thus in line with the expectations from the network model of polymer, there exists an optimal compaction where residence times are minimum. The fraction of unique sites visited $\langle S_u \rangle$ also exhibits a non-monotonic behaviour with increasing θ, or equivalently, decreasing R_g (see SI Fig. S10A). At the polymer conformations where $\langle T_{\text{res}} \rangle$ is minimum, $\langle S_u \rangle \sim 0.75$, similar to results from the artificial network.

B. Lennard-Jones polymer model

We next use a bead-spring polymer model of chromatin with attractive Lennard-Jones (LJ) potential [35, 36, 77]. This introduces self-avoidance effects, and bond length fluctuations absent in the FRC model. Any pair of beads of diameter σ, separated by distance r, interacts via the potential,

$$E_{LJ}(r) = \begin{cases} 4\epsilon \left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6} \right] & r < 2.5\sigma \\ 0 & r \geq 2.5\sigma \end{cases} \tag{5}$$

where ϵ is the LJ interaction strength (see SI Sec. 3.2). A LJ polymer undergoes a second-order phase transition from a coiled to a compact state (See SI Fig. S7A), accompanied by an increase in the number of intersegmental bonds (see SI Fig. S7B), as ϵ is increased. The residence times $\langle T_{\text{res}} \rangle$ on these LJ polymer backbones exhibited a non-monotonic pattern as the LJ interaction strength is increased (Fig. 3B), or equivalently as the polymer compaction increases as shown in Fig. 3B inset. There was a distinct minimum observed in the residence times, occurring close to the collapse transition. The behavior of the fraction of unique sites visited, $\langle S_u \rangle$, also followed a non-monotonic nature with increase in ϵ (see SI Fig. S10B). Thus a LJ polymer qualitatively mirrors the results obtained from the uniform network and FRC polymer models.

However, unlike the significant rise seen in the FRC and the network model, the extent of $\langle T_{\rm res}\rangle$ increase beyond a critical ϵ is less pronounced in this case. This limited increase is due to the inherent hard-core repulsion in the LJ potential, which imposes a lower limit on polymer compaction, and on the number of intersegmental bonds. This interplay between LJ interactions and polymer structure thus introduces a unique feature that distinguishes it from both FRC and the network model.

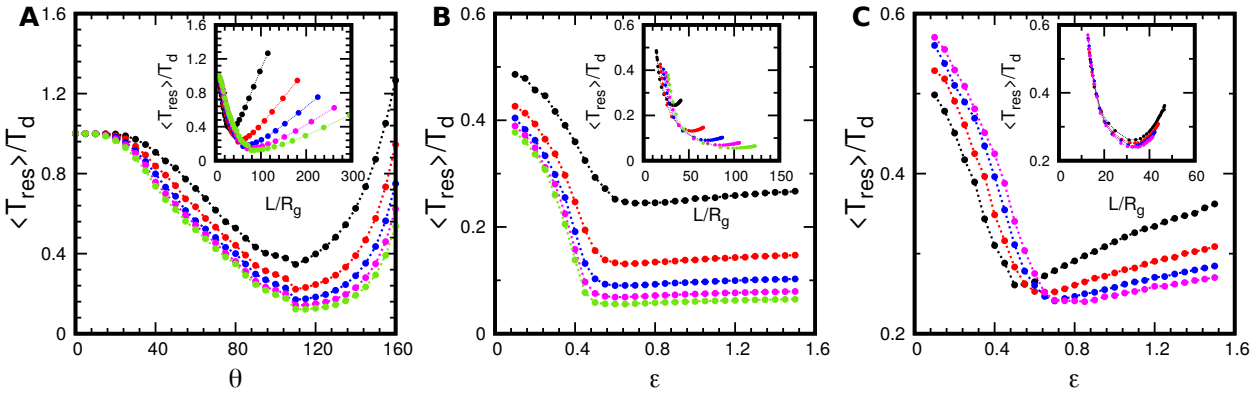

FIG. 3. **A.** Scaled mean residence time ($\frac{\langle T_{\rm res}\rangle}{T_d}$) in FRC polymer domain as a function of angle (θ) between any two bond vectors for $L = 100 - 500$. Inset: Variation of $\frac{\langle T_{\rm res}\rangle}{T_d}$ in FRC polymer domain with radius of gyration (R_g). **B.** $\frac{\langle T_{\rm res}\rangle}{T_d}$ in LJ polymer domain as a function of LJ interaction strength (ϵ) between any two non-bonded beads for $L = 100 - 500$. Inset: Variation of $\frac{\langle T_{\rm res}\rangle}{T_d}$ in LJ polymer domain with radius of gyration (R_g). **C.** $\frac{\langle T_{\rm res}\rangle}{T_d}$ in soft LJ polymer domain as a function of soft LJ interaction strength (ϵ) between any two non-bonded beads for $L = 100$ with varying softness characterised by η_1. Inset: Variation of ($\frac{\langle T_{\rm res}\rangle}{T_d}$) in soft LJ polymer domain with radius of gyration (R_g).

C. Soft Lennard-Jones polymer model

Is this weak non-monotonic rise in $\langle T_{\rm res}\rangle$ then the correct physical description of residence times? The answer hinges on whether hardcore LJ potential best represents coarse-grained chromatin conformations. Experimental and theoretical studies of chromatin configurations and 3D distances suggest that a soft inter-bead potential that allows for overlap is a more realistic description of coarse-grained chromatin [52, 78]. We study a polymer model with following soft inter-bead potential,

$$E_{soft}(r) = \begin{cases} V_0[1 - (\frac{r}{r_m})^{\eta_1}]^{\eta_2} - \epsilon & 0 \leq r < r_m \\ 4\epsilon \left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6} \right] & r_m \leq r < 2.5\sigma \\ 0 & r \geq 2.5\sigma \end{cases} \tag{6}$$

where $V_0 - \epsilon$ is the energy penalty at complete overlap, and η_1 and η_2 tune the softness of the potential (see SI Sec. 3.3 and SI Fig. S8). We investigate $\langle T_{\text{res}} \rangle$ in polymers with varying degrees of softness by tuning η_1. A lower value of η_1 indicates a softer polymer, whereas high η_1 resembles the classical LJ potential. For a very soft polymer ($\eta_1 = 0.03$), $\langle T_{\text{res}} \rangle$ shows a strong non-monotonicity, analogous to the FRC and the network results. This is because a soft polymer core cannot effectively enforce volume exclusion, resulting in a more tightly packed polymer compared to the conventional LJ model (see SI Fig. S9A). This then leads to a higher number of non-neighboring contacts within the polymer domain (see SI Fig. S9B). As we decrease the softness (increasing η_1), the degree of non-monotonicity in the residence times decreases and approaches the LJ-like behaviour. Thus non-monotonic behaviour of residence times remains a crucial feature of protein diffusion on coarse-grained chromatin-like soft polymer topologies. The degree of non-monotonicity $\langle T_{\text{res}} \rangle$, as well as the efficiency of search, $\langle S_u \rangle$ (see SI Fig. S10C) are controlled by the softness of the inter-bead potential.

IV. POLYMERS FROM TAD DATA

Is this non-monotonic behaviour of residence times relevant for real TADs? Where do TAD conformations lie with respect to this optimal compaction state? To answer these, we utilized publicly accessible experimental Hi-C contact probability data [79] to determine the spatially neighbouring contacts for TAD domains. Specifically, we characterised TADs within Chromosomes 1, 17, and 18 of the human genome, with a resolution of 1 kbp (see SI Fig. S11A). The distribution of TAD lengths revealed that TADs spanning 300 kbp were the most probable within these chromosomes [77] (see SI Fig. S11B), and these were chosen for subsequent analysis. We create a contact space network representation of chromatin, with bonds between any two beads created with the direct Hi-C probabilities (see SI Sec. 4.1). We then computed $\langle T_{\text{res}} \rangle$ on these networks (Fig 4A). We observed a substantial correlation

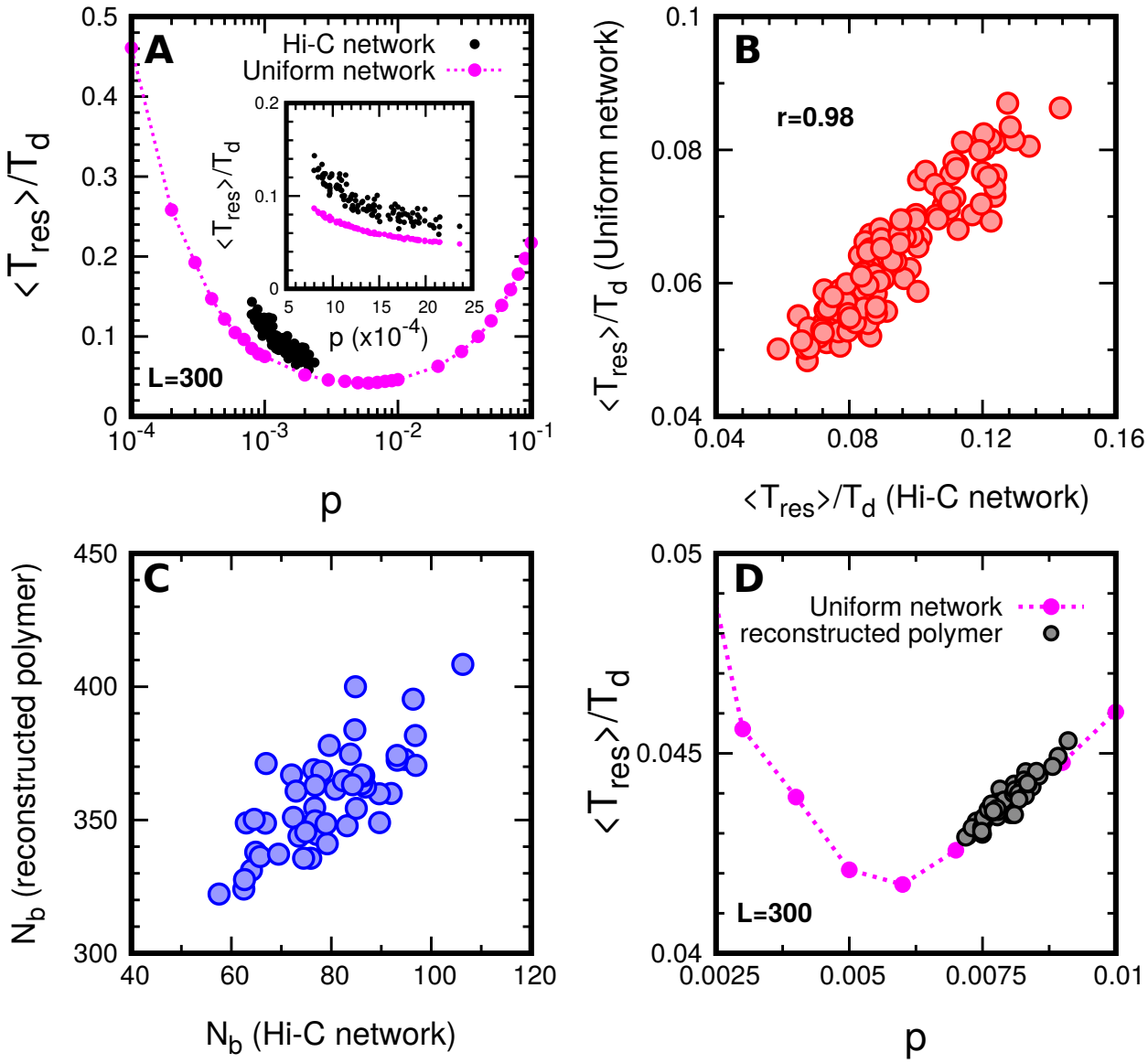

FIG. 4. **A.** Scaled mean residence time ($\frac{\langle T_{\rm res}\rangle}{T_d}$) in Hi-C network domain as a function of average contact probability (p) between any two non-bonded beads. Inset: Zoomed in plot to show the variation of $\frac{\langle T_{\rm res}\rangle}{T_d}$. **B.** Correlation between ($\frac{\langle T_{\rm res}\rangle}{T_d}$) values in Hi-C network domain and uniformly connected domain with same average p value. **C.** Number of non-neighbouring bonds (N_b) in Hi-C network domain and in a reconstructed polymer domain from using Hi-C matrices. **D.** $\frac{\langle T_{\rm res}\rangle}{T_d}$ in reconstructed polymer domain as a function of average contact probability (p) between any two non-bonded beads. Inset: Zoomed in plot to show the variation of $\frac{\langle T_{\rm res}\rangle}{T_d}$ near the minima.

between these residence times and those obtained from randomly generated uniform networks (Fig 4B). This finding suggests that the non-monotonic variation in residence times

is likely to persist in networks resembling TAD connectivities.

However, the number of intersegmental bonds obtained in these networks reconstructed from direct Hi-C probabilities were quite low, a consequence of the known issue of sparseness of Hi-C matrices, which can often fail to capture many of the actual genomic interactions [80–82]. To situate the spatial structure of these TADs accurately, we then proceed to reconstruct 3D polymer structures of TADs based on the Hi-C data (see SI Sec. 4.2 and SI Figs. S11C, S11D) using Langevin dynamics simulations [52]. Unlike the non-polymer network representation in the preceding approach, this results in a physical polymer structure. As expected, the number of intersegmental bonds in these reconstructed polymers were an order of magnitude higher than that was obtained in the network representation (Fig. 4C) [83]. Remarkably, when residence times were computed using the contact information from these reconstructed polymers, these were found to lie very close to the minimum residence times (Fig. 4D). This suggests that the organization of real TADs, particularly their compaction, is orchestrated in such a manner that facilitates relatively efficient protein search kinetics within these domains.

V. DISCUSSION

We propose that facilitated diffusion which takes into account correlated intersegmental jumps can leverage the compaction state of chromatin, characterised by the presence of TADs and specific polymer conformations, to access distant genomic regions for more efficient search. We show that there exists an optimal level of chromatin compaction where the residence time of TFs on a polymer domain is minimized, leading to more efficient search kinetics. This is in line with direct evidence from single-molecule experiments which have reported that search process by a site-specific restriction enzyme on DNA becomes almost doubly efficient when the DNA configuration is collapsed instead of a fully extended configuration [84]. Because of enhanced DNA looping in collapsed configurations, a protein that would typically rebind rapidly to the segment it just visited can instead be captured by another spatially close (but genomically distant) segment, thereby altering search kinetics. The optimal compaction state of polymers which minimises residence time predicted by our study was a feature in all the polymer models studied. In particular, motivated by studies showing chromatin is most aptly characterised as a soft polymer, we show how the

compaction of chromatin allows an interplay between chromatin structure and protein kinetics thereby influences search processes. Tantalisingly, an analysis of connection probabilities from experimentally determined Hi-C data suggests that TADs reside near a zone of optimal compaction where residence times are minimised. While the role of TADs in genome organisation and gene regulation has been extensively studied, we show a hitherto unsuspected role of TAD organisation in optimising protein search strategies. The structural organization of the genome within TADs may thus have evolved to facilitate rapid and precise dynamics of DNA-associated proteins, which is crucial for gene regulation and cellular function.

VI. ACKNOWLEDGEMENT

SD acknowledges fellowship support from the PMRF, MoE, India. MKM and RP acknowledge funding from Core Research Grant by Science and Engineering Research Board, India (Grant number: CRG/2022/008142).

---

[1] A. D. Riggs, S. Bourgeois, and M. Cohn, Journal of molecular biology **53**, 401 (1970).

[2] O. G. Berg, R. B. Winter, and P. H. Von Hippel, Biochemistry **20**, 6929 (1981).

[3] R. B. Winter, O. G. Berg, and P. H. Von Hippel, Biochemistry **20**, 6961 (1981).

[4] S. Halford, "Hopping, jumping and looping by restriction enzymes," (2001).

[5] S. E. Halford and J. F. Marko, Nucleic acids research **32**, 3040 (2004).

[6] J. Widom, Proceedings of the National Academy of Sciences **102**, 16909 (2005).

[7] P. H. von Hippel and O. G. Berg, Journal of Biological Chemistry **264**, 675 (1989).

[8] O. G. Berg and P. H. von Hippel, Annual review of biophysics and biophysical chemistry **14**, 131 (1985).

[9] P. H. von Hippel and O. G. Berg, Journal of Biological Chemistry **264**, 675 (1989).

[10] A. Veksler and A. B. Kolomeisky, The journal of physical chemistry B **117**, 12695 (2013).

[11] A. B. Kolomeisky, Physical Chemistry Chemical Physics **13**, 2088 (2011).

[12] L. Mirny, M. Slutsky, Z. Wunderlich, A. Tafvizi, J. Leith, and A. Kosmrlj, Journal of Physics A: Mathematical and Theoretical **42**, 434013 (2009).

[13] O. Bénichou, C. Chevalier, B. Meyer, and R. Voituriez, Physical review letters **106**, 038102 (2011).

[14] C. Loverdo, O. Benichou, R. Voituriez, A. Biebricher, I. Bonnet, and P. Desbiolles, Physical review letters **102**, 188101 (2009).

[15] M. C. DeSantis, J.-L. Li, and Y. Wang, Physical Review E **83**, 021907 (2011).

[16] N. P. Stanford, M. D. Szczelkun, J. F. Marko, and S. E. Halford, The EMBO journal **19**, 6546 (2000).

[17] J. Gorman and E. C. Greene, Nature structural & molecular biology **15**, 768 (2008).

[18] I. Bonnet, A. Biebricher, P.-L. Porté, C. Loverdo, O. Bénichou, R. Voituriez, C. Escudé, W. Wende, A. Pingoud, and P. Desbiolles, Nucleic Acids Research **36**, 4118 (2008), https://academic.oup.com/nar/article-pdf/36/12/4118/16749358/gkn376.pdf.

[19] J. Elf, G.-W. Li, and X. S. Xie, Science **316**, 1191 (2007).

[20] M. Castellanos, N. Mothi, and V. Muñoz, Nature Communications **11**, 540 (2020).

[21] J. H. Kim and R. G. Larson, Nucleic Acids Research **35**, 3848 (2007), https://academic.oup.com/nar/article-pdf/35/11/3848/16762701/gkm332.pdf.

[22] H. Kabata, O. Kurosawa, I. Arai, M. Washizu, S. A. Margarson, R. E. Glass, and N. Shimamoto, Science **262**, 1561 (1993).

[23] A. Granéli, C. C. Yeykal, R. B. Robertson, and E. C. Greene, Proceedings of the National Academy of Sciences **103**, 1221 (2006).

[24] P. C. Blainey, A. M. van Oijen, A. Banerjee, G. L. Verdine, and X. S. Xie, Proceedings of the National Academy of Sciences **103**, 5752 (2006).

[25] Y. Wang, R. H. Austin, and E. C. Cox, Physical review letters **97**, 048302 (2006).

[26] R. D. Giuntoli, N. B. Linzer, E. J. Banigan, C. E. Sing, M. O. de la Cruz, J. S. Graham, R. C. Johnson, and J. F. Marko, Journal of Molecular Biology **427**, 3123 (2015).

[27] C. M. Caragine, S. C. Haley, and A. Zidovska, Phys. Rev. Lett. **121**, 148101 (2018).

[28] K. Kamagata, R. Kusano, S. Kanbayashi, T. Banerjee, and H. Takahashi, Nucleic Acids Research **51**, 6654 (2023), https://academic.oup.com/nar/article-pdf/51/13/6654/50903029/gkad471.pdf.

[29] G. Nir, I. Farabella, C. Pérez Estrada, C. G. Ebeling, B. J. Beliveau, H. M. Sasaki, S. D. Lee, S. C. Nguyen, R. B. McCole, S. Chattoraj, *et al.*, PLoS genetics **14**, e1007872 (2018).

[30] Y. Qi and B. Zhang, PLoS computational biology **15**, e1007024 (2019).

[31] S. Bianco, A. M. Chiariello, M. Conte, A. Esposito, L. Fiorillo, F. Musella, and M. Nicodemi, Current opinion in cell biology **64**, 10 (2020).

[32] J. J. Parmar, M. Woringer, and C. Zimmer, Annual review of biophysics **48**, 231 (2019).

[33] M. Barbieri, M. Chotalia, J. Fraser, L.-M. Lavitas, J. Dostie, A. Pombo, and M. Nicodemi, Proceedings of the National Academy of Sciences **109**, 16173 (2012).

[34] D. Jost, P. Carrivain, G. Cavalli, and C. Vaillant, Nucleic acids research **42**, 9553 (2014).

[35] C. A. Brackley, J. Johnson, S. Kelly, P. R. Cook, and D. Marenduzzo, Nucleic acids research **44**, 3503 (2016).

[36] A. Buckle, C. A. Brackley, S. Boyle, D. Marenduzzo, and N. Gilbert, Molecular cell **72**, 786 (2018).

[37] A. Rosa and R. Everaers, PLoS computational biology **4**, e1000153 (2008).

[38] G. Shi and D. Thirumalai, Physical Review X **11**, 011051 (2021).

[39] S. Brahmachari, V. G. Contessoto, M. Di Pierro, and J. N. Onuchic, Nucleic Acids Research **50**, 4258 (2022).

[40] M. Di Pierro, B. Zhang, E. L. Aiden, P. G. Wolynes, and J. N. Onuchic, Proceedings of the National Academy of Sciences **113**, 12168 (2016).

[41] S. Brahmachari and J. F. Marko, Proceedings of the National Academy of Sciences **116**, 24956 (2019).

[42] G. Bajpai, D. A. Pavlov, D. Lorber, T. Volk, and S. Safran, Biophysical Journal **118**, 549a (2020).

[43] J. O. Davies, A. M. Oudelaar, D. R. Higgs, and J. R. Hughes, Nature methods **14**, 125 (2017).

[44] T.-H. S. Hsieh, C. Cattoglio, E. Slobodyanyuk, A. S. Hansen, O. J. Rando, R. Tjian, and X. Darzacq, Molecular cell **78**, 539 (2020).

[45] S. Wang, J.-H. Su, B. J. Beliveau, B. Bintu, J. R. Moffitt, C.-t. Wu, and X. Zhuang, Biophysical Journal **112**, 217a (2017).

[46] R. P. Ghosh and B. J. Meyer, Annual review of cell and developmental biology **37**, 199 (2021).

[47] T. Misteli, Cell **183**, 28 (2020).

[48] J. Dekker and L. Mirny, Cell **164**, 1110 (2016).

[49] A. Pombo and N. Dillon, Nature reviews Molecular cell biology **16**, 245 (2015).

[50] S. Guha and M. K. Mitra, Soft Matter **19**, 153 (2023).

[51] D. Amiad-Pavlov, D. Lorber, G. Bajpai, A. Reuveny, F. Roncato, R. Alon, S. Safran, and T. Volk, Science advances **7**, eabf6251 (2021).

[52] S. Kadam, K. Kumari, V. Manivannan, S. Dutta, M. K. Mitra, and R. Padinhateeri, Nature Communications **14**, 4108 (2023).

[53] 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.*, Nature **485**, 381 (2012).

[54] J. R. Dixon, S. Selvaraj, F. Yue, A. Kim, Y. Li, Y. Shen, M. Hu, J. S. Liu, and B. Ren, Nature **485**, 376 (2012).

[55] Q. Szabo, F. Bantignies, and G. Cavalli, Science advances **5**, eaaw1668 (2019).

[56] A. L. Sanborn, S. S. Rao, S.-C. Huang, N. C. Durand, M. H. Huntley, A. I. Jewett, I. D. Bochkov, D. Chinnappan, A. Cutkosky, J. Li, *et al.*, Proceedings of the National Academy of Sciences **112**, E6456 (2015).

[57] T.-H. Hsieh, A. Weiner, B. Lajoie, J. Dekker, N. Friedman, and O. Rando, Cell **162**, 108 (2015).

[58] M. Mazzocca, A. Loffreda, E. Colombo, T. Fillot, D. Gnani, P. Falletta, E. Monteleone, S. Capozi, E. Bertrand, G. Legube, *et al.*, Nature Communications **14**, 6433 (2023).

[59] I. M. Sokolov, J. Mai, and A. Blumen, Phys. Rev. Lett. **79**, 857 (1997).

[60] T. Schötz, R. A. Neher, and U. Gerland, Phys. Rev. E **84**, 051911 (2011).

[61] C. A. Brackley, M. E. Cates, and D. Marenduzzo, Physical review letters **109**, 168103 (2012).

[62] C. Brackley, M. Cates, and D. Marenduzzo, Physical review letters **111**, 108101 (2013).

[63] A. Mondal and A. Bhattacherjee, Nucleic Acids Research **43**, 9176 (2015).

[64] A. Mondal and A. Bhattacherjee, The Journal of Physical Chemistry B **121**, 9372 (2017).

[65] B. Punia and S. Chaudhury, Journal of Statistical Mechanics: Theory and Experiment **2021**, 073502 (2021).

[66] B. Punia and S. Chaudhury, Physical Chemistry Chemical Physics **24**, 29074 (2022).

[67] P. Dey and A. Bhattacherjee, Scientific Reports **8**, 844 (2018).

[68] D. Krepel and Y. Levy, Physical Chemistry Chemical Physics **19**, 30562 (2017).

[69] S. Redner, *A Guide to First-Passage Processes*, A Guide to First-passage Processes (Cambridge University Press, 2001).

[70] V. Sood, S. Redner, and D. ben Avraham, Journal of Physics A: Mathematical and General **38**, 109 (2004).

[71] J. Klinger, R. Voituriez, and O. Bénichou, Physical Review E **103**, 032107 (2021).

[72] H. Schiessel, W. M. Gelbart, and R. Bruinsma, Biophysical Journal **80**, 1940 (2001).

[73] G. Bajpai, I. Jain, M. M. Inamdar, D. Das, and R. Padinhateeri, PLoS computational biology **13**, e1005365 (2017).

[74] G. Bajpai and R. Padinhateeri, Biophysical Journal **118**, 207 (2020).

[75] A. J. Spakowitz, The Journal of Chemical Physics **151** (2019).

[76] M. Rubinsten, *Polymer physics* (United States of America, 2003).

[77] S. Shin, G. Shi, and D. Thirumalai, PRX Life **1**, 013010 (2023).

[78] S. Fujishiro and M. Sasai, Proceedings of the National Academy of Sciences **119**, e2109838119 (2022).

[79] 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.*, Cell **159**, 1665 (2014).

[80] V. Y. Goel, M. K. Huseyin, and A. S. Hansen, Nature Genetics , 1 (2023).

[81] A. A. Galitsyna and M. S. Gelfand, Briefings in bioinformatics **22**, bbab316 (2021).

[82] R. A. Beagrie, C. J. Thieme, C. Annunziatella, C. Baugher, Y. Zhang, M. Schueler, A. Kukalev, R. Kempfer, A. M. Chiariello, S. Bianco, *et al.*, Nature Methods , 1 (2023).

[83] K. Kumari, B. Duenweg, R. Padinhateeri, and J. R. Prakash, Biophysical Journal **118**, 2193 (2020).

[84] B. van den Broek, M. A. Lomholt, S.-M. Kalisch, R. Metzler, and G. J. Wuite, Proceedings of the National Academy of Sciences **105**, 15738 (2008).