Biophysical *Journal*
Biophysical Letter

Biophysical Society

Chromatin Configuration Affects the Dynamics and Distribution of a Transiently Interacting Protein

Assaf Amitai[1,2,*]
[1]Department of Chemical Engineering, Institute for Medical Engineering and Science, Massachusetts Institute of Technology, Cambridge, Massachusetts and [2]Ragon Institute of MGH, MIT, and Harvard, Cambridge, Massachusetts

We present a theoretical study of the interaction between a protein (diffusing particle) and chromatin (polymer chain). Each monomer is a trap where a particle can transiently bind. We derive, to our knowledge, novel formulas for the transition rate between monomer sites, given a specific polymer configuration, and find that a particle is likely to rapidly rebind many times to its release site before moving to another site. The reattachment probability is larger when the local density around the release site is smaller. Interestingly, for an equilibrated polymer, the transition probability decays as a power law for close monomer-to-monomer distances and reaches an asymptotic value for long distances. By computing the transition rate between monomers, we show that the problem of facilitated search by a protein can be mapped to a continuous-time Markov chain, which we solve. Our findings suggest that proteins may be locally trapped for a time much longer than their dissociation time, whereas their overall motion is ergodic. Our results are corroborated by Brownian simulations.

The interaction of proteins with chromatin regulates many cellular functions. Most DNA-binding proteins interact both non-specifically and transiently (1) with many chromatin sites, as well as specifically and more stably with cognate binding sites. These interactions and chromatin structure are important in governing protein dynamics (2,3). However, the effect of these transient interactions on protein motion and distribution has not yet been shown from a first principle.

Some aspects of protein interactions with DNA have been studied in the context of the search of a gene promoter site by a transcription factor (TF) (4). It was first noted (5) that the search for a promoter site by a TF would be faster if it involves three-dimensional excursions, as well as sliding of the protein along DNA (6), as was shown in prokaryotes (7). These different types of motion were observed experimentally, leading to massive interest in models of facilitated diffusion (8–15), and in the impact of a regulating site's position on transcription (16,17). In current microscopy experiments, it is impossible to examine the search process to its fullest (18). Thus, we concentrate here on modeling the experimentally observed dynamics of the protein as a diffusing and interacting particle.

We will show that proteins are likely to stay in the proximity of a site for a time much longer than their dissociation time from DNA due to reattachment. Moreover, we find that reattachment depends on the local density around the release site, and that the precise configuration of the polymer impacts the interacting particle's dynamics. We further find the rates associated with the transition between different monomer sites. Finally, we show that the process as a whole is ergodic; it has no long-time power-law distribution of the residence time at a site as has been previously suggested (19).

We consider a point particle (protein) placed at a distance a from monomer n (locus), part of a long flexible polymer (20,21) (chromatin) (Fig. 1 *a*). The particle diffuses until it encounters monomer l. Absorption occurs at capture radius $\epsilon < a$ from l (Fig. 1 *b*), where the particle remains bound for characteristic time T. Upon dissociation, the particle is placed at a distance a from the monomer l position (Fig. 1 *b*), with a uniform angular distribution, and starts diffusing again. Although we postulate here two radii, release (a) and capture (ϵ), the effective behavior is equivalent to a model with one release radius and a partially reflecting boundary condition on it. In a partially reflecting model, the particle has a probability of re-absorbing immediately after release. Small re-absorption probability corresponds to $\epsilon \ll a$. Thus, when the release and capture radii are comparable, the particle spends much time around one monomer cluster, then jumps to another, only to come back to the first (Fig. 1 *c*).

Submitted September 5, 2017, and accepted for publication December 27, 2017.

*Correspondence: amitai@mit.edu

Editor: Andrew Spakowitz.
https://doi.org/10.1016/j.bpj.2017.12.037

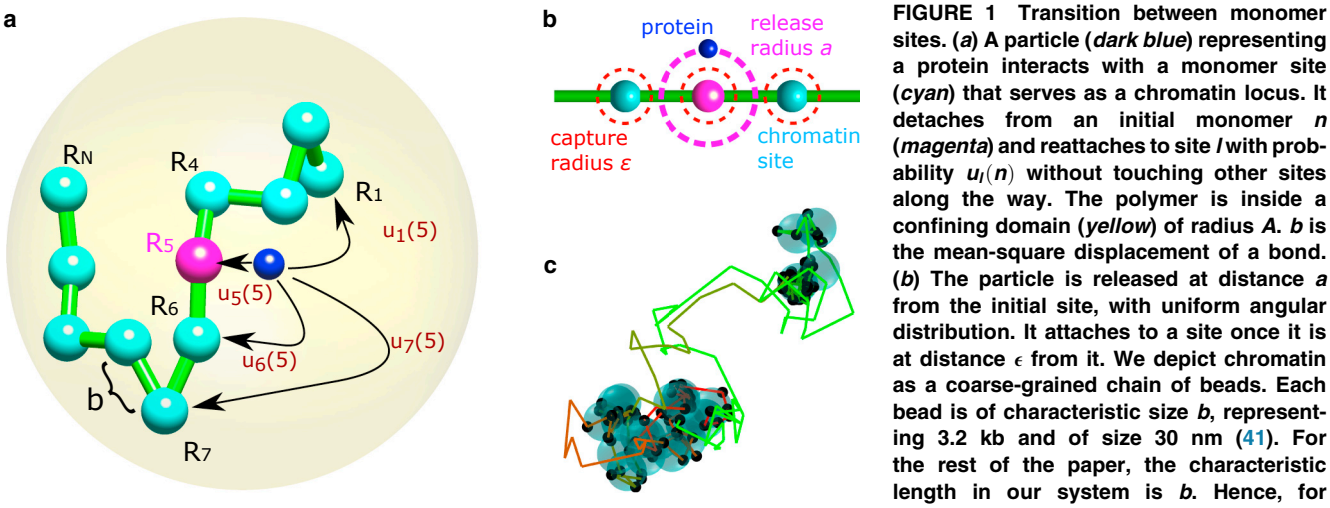

FIGURE 1 Transition between monomer sites. (*a*) A particle (*dark blue*) representing a protein interacts with a monomer site (*cyan*) that serves as a chromatin locus. It detaches from an initial monomer *n* (*magenta*) and reattaches to site *l* with probability $u_l(n)$ without touching other sites along the way. The polymer is inside a confining domain (*yellow*) of radius *A*. *b* is the mean-square displacement of a bond. (*b*) The particle is released at distance *a* from the initial site, with uniform angular distribution. It attaches to a site once it is at distance ϵ from it. We depict chromatin as a coarse-grained chain of beads. Each bead is of characteristic size *b*, representing 3.2 kb and of size 30 nm (41). For the rest of the paper, the characteristic length in our system is *b*. Hence, for $\epsilon = 0.03b = 0.9$ and $a = 0.3b = 9$nm, the release and capture radii are of the order of the size of a protein and the interaction distance. (*c*) Trajectory of the particle interacting with the monomer sites (*cyan*). The trajectory color changes with time from red to green. The black dots are attachment points at the monomer site ($a = 0.5b, \epsilon = 0.49b$, $N = 100$, and $A = 6b$). To see this figure in color, go online.

The probability of the particle arriving at a certain monomer site before another via three-dimensional diffusion depends on their respective initial distance from each other (22). The probability $u_l(\boldsymbol{x})$ that a particle starting from $\boldsymbol{x}$ arrives at monomer l before encountering any other monomer (see Supporting Material) is

$$u_l(\boldsymbol{x}) = \frac{1}{N} + 4\pi\epsilon\left[G(\boldsymbol{x},\boldsymbol{R}_l) - \frac{\sum_{j=1}^{N} G(\boldsymbol{x},\boldsymbol{R}_j)}{N}\right] + \epsilon\chi_l + \mathcal{O}(\epsilon^2), \tag{1}$$

where N is the polymer length, $\boldsymbol{R}_i$ is the position of monomer i, $G(\boldsymbol{x},\boldsymbol{y})$ is the Neumann Green's function of the Laplacian in a sphere of radius A, χ_l is a constant that depends on the monomers' positions but not on the initial position, $\boldsymbol{x}$. In writing Eq. 1, we assume that the trapping monomers are well separated. When the ϵ neighborhoods of two monomers merge, the equation can be modified (23).

Proteins move much faster in the nucleus than chromatin. The diffusion coefficient of a chromatin locus can be estimated by inserting a fluorescent tag and following its trajectory (24). Assuming a coarse-grained model of chromatin, it was found that a locus of size $3kb$ has a diffusion coefficient of about $D = 10^{-2}\ \mu\text{m}^2/\text{s}$ (21), whereas proteins move three orders of magnitude faster. for example, 13.5 $\mu\text{m}^2/\text{s}$ for the c-Myc protein (25). We thus assume that the polymer is fixed at an equilibrium configuration inside the domain while the particle is diffusing.

Assuming that the polymer is equilibrated in the domain, we found the transition probability $u_l(n)$, that a particle starting in the proximity of site n finds site l first before encountering other monomers, by averaging Eq. 1 with the equilibrium distribution of the polymer in bulk (see the Supporting Material):

$$\begin{aligned} u_l(n) &\approx \frac{1}{N}\left(1 - \frac{\epsilon}{a}\right) + \frac{4\pi\epsilon 3^{1/2}}{(2\pi)^{3/2}b}\left[|n-l|^{-1/2} - \frac{2}{N}\left(-\frac{35}{24}\right.\right. \\ &\left.\left. \sum_{k=l,n}\left((N-k)^{1/2} + (k-1)^{1/2}\right) - \frac{4(N-1)^{3/2}}{3N}\right)\right], l \neq n, \\ u_n(n) &\approx \frac{1}{N} + \frac{\epsilon}{a}\left(1 - \frac{1}{N}\right) - \left(\frac{3}{2\pi}\right)^{1/2}\frac{8\epsilon}{bN}\left((N-n)^{1/2}\right. \\ &\left. + (n-1)^{1/2} - \frac{2(N-1)^{3/2}}{3N} - \frac{35}{48}\right) \end{aligned} \tag{2}$$

where $2 < n, l < N$, b is the standard deviation of a bond length (Fig. 1 *a*), and we assume $\epsilon(N^{1/2}/b + Ca^{-1}) \ll 1$, with C a numerical factor of order 1. The expression for the end monomers is different (see Supporting Material, Section S4). Although we did not explicitly include a sliding state of the protein along DNA (6), it could be included by modifying the nearest-neighbors transition probability, $(u_{n-1}(n), u_{n+1}(n))$, and will not qualitatively change our results.

The reattachment probability $u_n(n)$ at the release site is larger than the probability of attaching to faraway sites (Fig. 2 *a*; Eq. 2), suggesting that once released, the particle is likely to rebind at the same site. For a longer polymer strand, the ratio between the reattachment probability and probability to attachment to a faraway site is larger than for a short polymer (Fig. 2 *b*). This result is contrary to previous studies that assumed that the transition probability and the k_{on} for rebinding were equal among all monomers (13,26–32), or were related to a Lévy type diffusion of the particle (30).

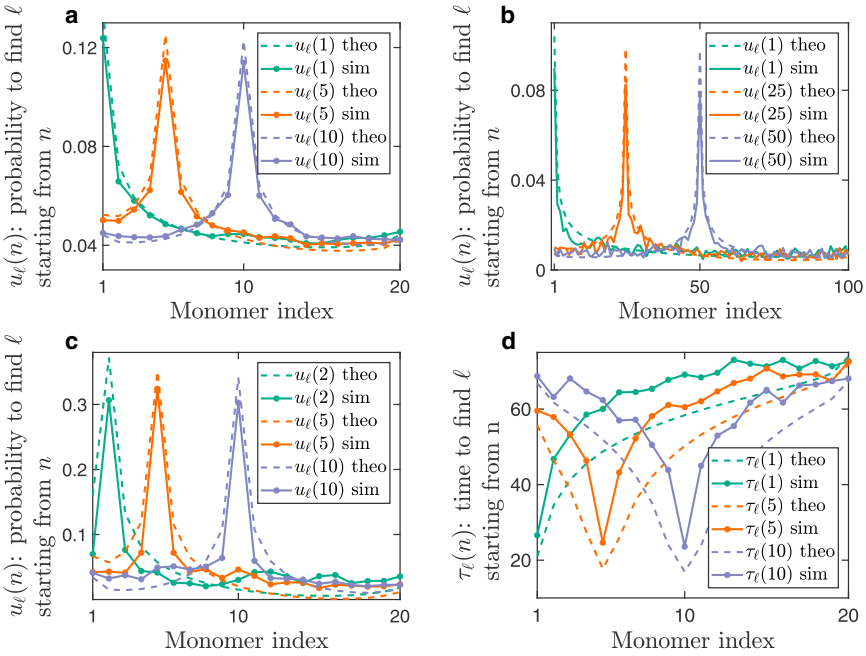

FIGURE 2 Transition between monomer sites. (*a*) The transition probability starting from sites 1 (*turquoise*), 5 (*orange*), or 10 (*purple*). The polymer has 20 monomer sites ($N = 20$), and $a = 0.3b$, $\epsilon = 0.03b$, and $A = 10b$. The solid lines are the result of Brownian simulations, whereas the dashed lines are computed using the analytical formula (Eq. 2). (*b*) The site transition probability when the polymer is longer ($N = 100$ sites). (*c*) The polymer is of length $N = 20$. It is crowded into a small domain, $A = 2b$, and the capture radius is $\epsilon = 0.12b$. (*d*) The mean first-transition time, $\tau_l(n)$, from site *n* to site *l* without interacting with any other site along the way. The full lines are the result of Brownian simulation and the dashed lines are computed using Eq. 5 with the same parameters as in (*a*). A single time unit is equal to1 t.u $= 10b^2/D$, where *D* is the diffusion coefficient of the particle. To see this figure in color, go online.

Interestingly, we find that the reattachment probability is minimal for the middle monomer (Eq. 2). To clarify this effect, we estimated the expected number of monomers in a ball around the release monomer, for different monomers along the chain (see Fig. S1). This quantity is akin to the local density around the release site. The monomer density is highest around the middle monomer, which is closest to the polymer center of mass. Thus, the reattachment probability is sensitive to the local density around the release site and not just to the average monomer density in the domain.

Segregation of chromosomes in the nucleus may be the result of self-avoiding interaction (SAI) between them (33). To study how SAI between monomers would modify the behavior of $u_l(n)$, we performed Brownian simulation (BS) where the monomers interacted through the Lennard-Jones potential (34) in addition to engaging in nearest-neighbors spring interactions (see Supporting Material). The separation of monomers due to SAI decreases the local monomer density around the release point, leading to a larger reattachment probability, $u_n(n)$, compared to that of a phantom chain (see Fig. S2 *a*).

The root mean-square distance between monomers scales with their distance along the chain (rmsd $= \sqrt{\langle (\boldsymbol{R}_n - \boldsymbol{R}_l)^2 \rangle} \sim |n-l|^{\nu}$). This results in the power-law scaling $u_l(n) \sim |n-l|^{-1/2}$ for proximal sites of a phantom polymer (Eq. 2), for which $\nu = 1/2$. To study the behavior for an SA polymer, we computed $u_l(n)$ between proximal sites from the BS, averaged over all release sites n, and fitted it: $\langle u_{|l-n|}(n) \rangle \sim A|l-n|^{-\alpha} + C$. We found the fitted exponent to be $\alpha = 0.59$ (see Fig. S2 *b*), as would be expected from the rmsd of an SA polymer, for which $\nu = 0.6$ is the Flory exponent (35). When a and ϵ are larger, α increases (see Fig. S2 *c*). This behavior may be explained by studying high-order expansions of $u_l(\boldsymbol{x})$ in the parameter ϵ.

Computing $u_l(n)$, we assumed that the polymer configuration is not much affected by confinement (the ratio of gyration radius to domain radius: $R_g/A = b\sqrt{N/6}/A = 0.18$ and $R_g/A = 0.4$ for Fig. 2, *a* and *b*, respectively). Thus, its equilibrium distribution was well approximated by the equilibrium distribution of a flexible chain in bulk. Since one end of the polymer is anchored at the origin of the domain, when the end-to-end distance ($R_{e2e} = b\sqrt{N}$) is of the order of the domain radius, the polymer feels the effect of confinement.

In the nucleus, chromosomes are tightly packed and are not in the dilute regime. We thus simulated a polymer in a smaller domain for which $R_g/A = 0.91$ and $R_{e2e}/A = 2.23$ and computed $u_l(n)$ (Fig. 2 *c*). Interestingly, Eq. 2 still matches the transition probability, even for this moderately crowded polymer. Indeed, when the target monomer l is close by, its distance distribution from the release monomer is not affected much by the presence of confinement. At the same time, the transition probability to faraway monomers weakly depends on the distance between the release site and the target. Thus, it is not much affected by confinement. The validity of Eq. 2 is expected to break for extreme packing ($A \ll R_g$).

To understand the rates involved in the encounters between proteins and chromatin, we computed the conditional mean first-passage time (MFPT), $\tau_l(\boldsymbol{x})$, of a particle from position $\boldsymbol{x}$ to site l without encountering other monomers

along the way. The conditional dynamics of the particle obeys the Langevin equation (36),

$$d\boldsymbol{x} = \boldsymbol{a}(\boldsymbol{x})dt + \sqrt{2D}d\boldsymbol{\omega}, \tag{3}$$

where $d\boldsymbol{\omega}$ is a white Gaussian noise, the drift is $\boldsymbol{a}(\boldsymbol{x}) = 2D\nabla u_l(\boldsymbol{x})/u_l(\boldsymbol{x})$, and $u_l(\boldsymbol{x})$ is given by Eq. 2. Since $u(\boldsymbol{x})$ approaches zero when $\boldsymbol{x}$ approaches any monomer other than l (see Eq. S5 in the Supporting Material), the drift $a(\boldsymbol{x})$ will diverge for $\boldsymbol{x}\rightarrow\boldsymbol{R}_i$ $(i\neq l)$. Thus, moving according to Eq. 3, the particle experiences a drift pushing it away from all monomers except for its final destination, l.

$\tau_l(\boldsymbol{x})$ can be found by solving a boundary value problem (see Eq. S73). It has the approximate solution

$$\tau_l(\boldsymbol{x}) \approx \frac{N^{-1}+\epsilon\chi_l}{D\lambda_{0,\epsilon}}\left[\frac{1}{u_l(\boldsymbol{x})} - 4\pi C\epsilon\sum_{j=1}^{N}\frac{G(\boldsymbol{x},\boldsymbol{R}_j)}{u_l(\boldsymbol{x})}\right], \tag{4}$$

where $\lambda_{0,\epsilon}$ is the eigenvalue of an associated eigenvalue problem (see Supporting Material, Section S5).

Averaging Eq. 4 with the equilibrium polymer configuration, we find an asymptotic formula for the mean conditional transition time starting from monomer n,

$$\tau_l(n) \approx \frac{N^{-1}+\epsilon\chi_l}{D\lambda_{0,\epsilon}}\Big[\frac{1}{u_l(n)} - 4\pi\epsilon N\Bigg(\sum_j \tilde{G}_{nj} - 4\pi\epsilon N\tilde{G}_{nn}\left[\tilde{G}_{nl}+\chi_l-\frac{1}{N}\sum_j\tilde{G}_{nj}\right]\Bigg)\Big], \tag{5}$$

where $\tilde{G}_{nl}$ is the Green's function between monomer n and j positions, averaged over the equilibrium distribution.

Compared to BS, this formula (Eq. 5) matches the simulation, where the difference is up to $\sim$20% of the MFPT (see Fig. 2 *d*). Interestingly, $\tau_l(n)$ to the release site $(l=n)$ is much faster than to other sites, and the MFPT converges asymptotically when the target monomer is far from the release site ($|l-n| \gg 0$). When $\epsilon\rightarrow a$, the recapture time is faster with respect to transition time to another site. Indeed, when an interacting protein starts from the boundary layer of the release site, the characteristic recapture rate is significantly higher than the travel time to other sites. Hence, we suggest that when a protein is released from a chromatin site, it would quickly rebind to it with high probability.

We find that the addition of SAI reduces $\tau_l(n)$ compared to a phantom polymer (Fig. S2 *d*). The addition of SAI increases the average distance between monomers. Consequently, the eigenvalue $\lambda_{0,\epsilon}$ is larger (see Eq. S87), resulting in a smaller MFPT (Eq. 5). Therefore, the monomers of an SA polymer fill the domain more optimally than those of a phantom polymer, resulting in a rapid capture.

Since the reattachment rate at a site is larger, we expect that when SAIs are dominant, proteins will spend a larger fraction of their time bound at monomer sites. Heterochromatin is considered to be denser than euchromatin (37). However, the nature of interaction of proteins with heterochromatin is still unclear. We would expect $u_n(n)$ to be smaller in heterochromatin domains and that a protein would "forget" faster its release position. Alternatively, if reattachment occurs with higher probability $(\epsilon\rightarrow a)$ in heterochromatin, $u_n(n)$ will be larger in these domains.

Proteins in the nucleus can be either bound to chromatin or other nuclear compartments, or can stochastically move between association events. We denote the unbound state as "in transit" from one monomer site to another. We assume that at site n, the particle remains bound for a characteristic time, T_n, and leaves with a Poissonian rate. We thus formulated the transition between the different states using a continuous-time Markov chain (CTMC) (Fig. 3 *a*). We constructed the rate matrix $\boldsymbol{Q}$ between bound states and transit states that depend on $u_l(n)$, $\tau_l(n)$, and T_n (see the Supporting Material). We used $\boldsymbol{Q}$ to find the time evolution of the probability distribution function of the particle,

$$p(\boldsymbol{S},t) = e^{\boldsymbol{Q}t}p(\boldsymbol{S},0), \tag{6}$$

where $\boldsymbol{S}$ is the state vector of the CTMC and $p(\boldsymbol{S},0)$ is the initial particle distribution.

We performed BS and estimated numerically the $u_l(n)$ and $\tau_l(n)$ values, taking $T_n = T$ $(n = 1\ldots N)$; we estimated $p(\boldsymbol{S},t)$ starting from site n using Eq. 6. When the reattachment probability is high, the particle remains in proximity to its initial site for a long time (Fig. 3, *b* and *c*) compared with T. Thus, the equilibration time of the protein in the domain is much longer than its disassociation rate from a site ($k_{\text{off}} = 1/T$). When the reattachment probability is smaller, the particle diffuses farther from its initial site (Fig. 3 *c*). Interestingly, the residence probability at the original site is not uniform along the chain. Since the middle monomer resides where monomer density is highest, its reattachment probability, $u_{N/2}(N/2)$, is minimal (Eq. 2).

Using the long-time behavior of $p(\boldsymbol{S},t)$, we estimated fraction f of bound particles. For high reattachment probability, $f^{\epsilon=0.49b,a=0.5b}\approx 0.93$, whereas the rest of the probability is in the unbound (in-transit) states. For smaller reattachment probability, $f^{\epsilon=0.3b,a=0.5b}\approx 0.6$. In the Supporting Material, we plotted f for different values of ϵ (see Fig. S3). f estimated using Eq. 6 corresponds to the bound fraction estimated directly with BS.

A naive estimate for the bound fraction, which does not take into account reattachment, can be found through the ratio of the off rate, k_{off}, and on rate starting from the bulk ($k_{\text{on}} = 4\pi\epsilon ND/|\Omega|$ (38)): $f_{\text{bulk}} = k_{\text{on}}/(k_{\text{on}}+k_{\text{off}})$. We found that the bulk estimate greatly underestimates the bound fraction found from the BS or using the CTMC formalism (see Fig. S3). Thus, starting at the boundary layer of the initial

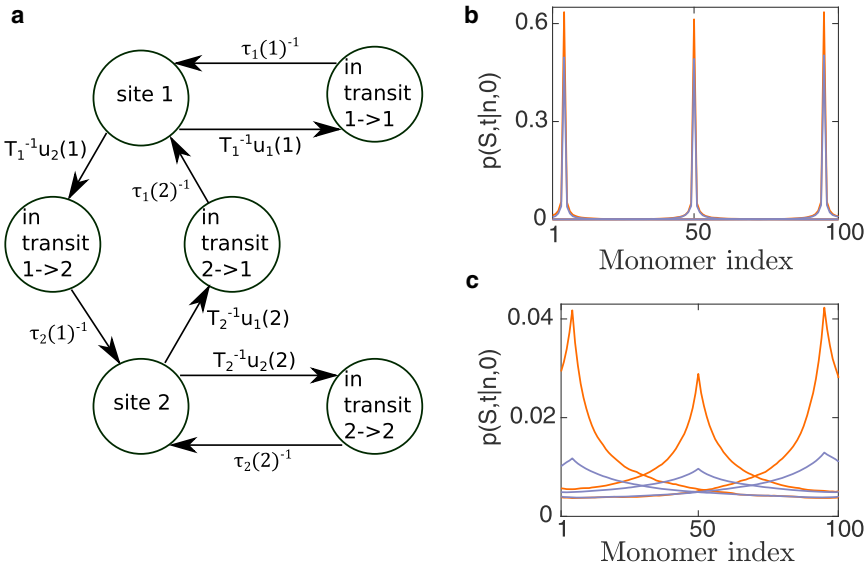

FIGURE 3 Site-site transition time and particle distribution in the domain. (*a*) Formulating the behavior of a particle (protein) that interacts with the polymer (chromatin), inside a domain (nucleus) as a continuous-time Markov chain. We illustrate it for the case of two binding monomer sites. The particle can be bound at site 1 or 2 and is released with Poissonian dissociation rates T_1^{-1} and T_2^{-1}. While diffusing to bind to a site, it is "in transit" and arrives at its destination with probability 1 and rate $\tau_I(n)^{-1}$. (*b* and *c*) The monomer is bound at a site for a characteristic time $T = 1$ t.u. The probability, *p*, of the particle being in a bound state at the monomer site after $1T$ (*b*) or $10T$ (*c*) is computed by solving numerically Eq. 6. The monomer starts at a distance *a* from either monomer 2, 50, or 99. $u_I(n)$ and $\tau_I(n)$ were estimated from Brownian simulation: $N = 100$, $a = 0.5b$, $A = 10b$, and $\epsilon = 0.49$ (*orange*) or $\epsilon = 0.3b$ (*purple*). For $N = 100$, there are 100 bound states and 10^4 in-transit states. To see this figure in color, go online.

monomer site can impact significantly the fraction of bound proteins. This may be the origin of the observed non-uniform distribution (39) and protein clusters (40) in the nucleus.

To conclude, there is a finite probability, after each release, that the protein diffuses away to another remote site, given that there is a trajectory between them. Thus, the process is ergodic. When the ϵ-neighborhoods of many traps overlap around the released site, such that the particle cannot find a path out, it will be quenched in this area.

Our findings can explain the long residence-time distributions that are observed experimentally (19) without the need for a power-law waiting-time distribution as assumed in a continuous-time random-walk model. As we have shown, escaping a binding site involves several dissociation and association events, with different characteristic rates. Hence, the localization time distribution may not appear to have exponential distribution in experiments.

Since the three-dimensional organization of chromatin guides search of TFs through transient interactions, the number of proteins and their interaction strength is not sufficient to understand their collective behavior. To fully model protein behavior at chromatin loci, one has to study the nature of the local interactions around the site of interest. Based on our model, we can extract directly from microscopy data the interaction parameters of proteins at specific chromatin domains. Thus, we can understand how different proteins "see" chromatin differently.

SUPPORTING MATERIAL

Supporting Materials and Methods and three figures are available at http://www.biophysj.org/biophysj/supplemental/S0006-3495(18)30057-2.

AUTHOR CONTRIBUTIONS

A.A. performed the research and wrote the article.

ACKNOWLEDGMENTS

The author thanks M. Kardar, D. Holcman, J. Reingruber, and A. K. Chakraborty for their helpful discussions and comments.

I acknowledge financial support from Massachusetts General Hospital (internal fund 214931).

REFERENCES

1. Halford, S. E., and J. F. Marko. 2004. How do site-specific DNA-binding proteins find their targets? *Nucleic Acids Res.* 32:3040–3052.
2. Misteli, T. 2001. Protein dynamics: implications for nuclear architecture and gene expression. *Science.* 291:843–847.
3. Hansen, A. S., I. Pustova, …, X. Darzacq. 2017. CTCF and cohesin regulate chromatin loop stability with distinct dynamics. *eLife.* 6:e25776.
4. Ptashne, M. 1986. Gene regulation by proteins acting nearby and at a distance. *Nature.* 322:697–701.
5. Berg, O. G., R. B. Winter, and P. H. von Hippel. 1981. Diffusion-driven mechanisms of protein translocation on nucleic acids. 1. Models and theory. *Biochemistry.* 20:6929–6948.
6. Hammar, P., P. Leroy, …, J. Elf. 2012. The *lac* repressor displays facilitated diffusion in living cells. *Science.* 336:1595–1598.
7. Elf, J., G.-W. Li, and X. S. Xie. 2007. Probing transcription factor dynamics at the single-molecule level in a living cell. *Science.* 316:1191–1194.
8. Slutsky, M., M. Kardar, and L. A. Mirny. 2004. Diffusion in correlated random potentials, with applications to DNA. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 69:061903.
9. Hu, T., A. Y. Grosberg, and B. I. Shklovskii. 2006. How proteins search for their specific sites on DNA: the role of DNA conformation. *Biophys. J.* 90:2731–2744.

10. Li, G.-W., O. G. Berg, and J. Elf. 2009. Effects of macromolecular crowding and DNA looping on gene regulation kinetics. *Nat. Phys.* 5:294–297.
11. Bauer, M., and R. Metzler. 2012. Generalized facilitated diffusion model for DNA-binding proteins with search and recognition states. *Biophys. J.* 102:2321–2330.
12. Cartailler, J., and J. Reingruber. 2015. Facilitated diffusion framework for transcription factor search with conformational changes. *Phys. Biol.* 12:046012.
13. Bénichou, O., Y. Kafri, …, R. Voituriez. 2009. Searching fast for a target on DNA without falling to traps. *Phys. Rev. Lett.* 103:138102.
14. Lomholt, M. A., B. van den Broek, …, R. Metzler. 2009. Facilitated diffusion with DNA coiling. *Proc. Natl. Acad. Sci. USA.* 106:8204–8208.
15. Koslover, E. F., M. A. Díaz de la Rosa, and A. J. Spakowitz. 2011. Theoretical and computational modeling of target-site search kinetics in vitro and in vivo. *Biophys. J.* 101:856–865.
16. Kolesov, G., Z. Wunderlich, …, L. A. Mirny. 2007. How gene order is influenced by the biophysics of transcription regulation. *Proc. Natl. Acad. Sci. USA.* 104:13948–13953.
17. Godec, A., and R. Metzler. 2016. Universal proximity effect in target search kinetics in the few-encounter limit. *Phys. Rev. X.* 6:041037.
18. Normanno, D., L. Boudarène, …, M. Dahan. 2015. Probing the target search of DNA-binding proteins in mammalian cells using TetR as model searcher. *Nat. Commun.* 6:7357.
19. Caccianini, L., D. Normanno, …, M. Dahan. 2015. Single molecule study of non-specific binding kinetics of LacI in mammalian cells. *Faraday Discuss.* 184:393–400.
20. Doi, M., and S. F. Edwards. 1986. The Theory of Polymer Dynamics. Oxford Clarendon Press, Oxford, United Kingdom.
21. Amitai, A., A. Seeber, …, D. Holcman. 2017. Visualization of chromatin decompaction and break site extrusion as predicted by statistical polymer modeling of single-locus trajectories. *Cell Rep.* 18:1200–1214.
22. Pulkkinen, O., and R. Metzler. 2013. Distance matters: the impact of gene proximity in bacterial gene regulation. *Phys. Rev. Lett.* 110:198101.
23. Cheviakov, A. F., and M. J. Ward. 2011. Optimizing the principal eigenvalue of the Laplacian in a sphere with interior traps. *Math. Comput. Model.* 53:1394–1409.
24. Robinett, C. C., A. Straight, …, A. S. Belmont. 1996. In vivo localization of DNA sequences and visualization of large-scale chromatin organization using lac operator/repressor recognition. *J. Cell Biol.* 135:1685–1700.
25. Izeddin, I., V. Récamier, …, X. Darzacq. 2014. Single-molecule tracking in live cells reveals distinct target-search strategies of transcription factors in the nucleus. *eLife.* 3:e02230.
26. Slutsky, M., and L. A. Mirny. 2004. Kinetics of protein-DNA interaction: facilitated target location in sequence-dependent potential. *Biophys. J.* 87:4021–4035.
27. Reingruber, J., and D. Holcman. 2011. Transcription factor search for a DNA promoter in a three-state model. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 84:020901.
28. Sheinman, M., and Y. Kafri. 2009. The effects of intersegmental transfers on target location by proteins. *Phys. Biol.* 6:016003.
29. Bénichou, O., C. Chevalier, …, R. Voituriez. 2011. Facilitated diffusion of proteins on chromatin. *Phys. Rev. Lett.* 106:038102.
30. Lomholt, M. A., T. Ambjörnsson, and R. Metzler. 2005. Optimal target search on a fast-folding polymer chain with volume exchange. *Phys. Rev. Lett.* 95:260603.
31. Shvets, A. A., and A. B. Kolomeisky. 2016. The role of DNA looping in the search for specific targets on DNA by multisite proteins. *J. Phys. Chem. Lett.* 7:5022–5027.
32. Mirny, L., M. Slutsky, …, A. Kosmrlj. 2009. How a protein searches for its site on DNA: the mechanism of facilitated diffusion. *J. Phys. A Math. Theor.* 42:434013.
33. Amitai, A., and D. Holcman. 2017. Polymer physics of nuclear organization and function. *Phys. Rep.* 678:1–83.
34. Lennard-Jones, J. E. 1924. On the determination of molecular fields: II: from the equation of state of a gas. *Proc. R. Soc. Lond. A.* 106:463–477.
35. de Gennes, P. G. 1979. Scaling Concepts in Polymer Physics. Cornell University Press, Ithaca, NY.
36. Karlin, S., and H. Taylor. 1981. A Second Course in Stochastic Processes. Academic Press, Cambridge, MA.
37. Woodcock, C. L., and R. P. Ghosh. 2010. Chromatin higher-order structure and dynamics. *Cold Spring Harb. Perspect. Biol.* 2:a000596.
38. Schuss, Z. 2009. Diffusion and Stochastic Processes. An Analytical Approach. Springer-Verlag, New York, NY.
39. Knight, S. C., L. Xie, …, R. Tjian. 2015. Dynamics of CRISPR-Cas9 genome interrogation in living cells. *Science.* 350:823–826.
40. Cisse, I. I., I. Izeddin, …, X. Darzacq. 2013. Real-time dynamics of RNA polymerase II clustering in live human cells. *Science.* 341:664–667.
41. Amitai, A., and D. Holcman. 2013. Diffusing polymers in confined microdomains and estimation of chromosomal territory sizes from chromosome capture data. *Phys. Rev. Lett.* 110:248105.

Biophysical Journal, Volume *114*

Supplemental Information

Chromatin Configuration Affects the Dynamics and Distribution of a Transiently Interacting Protein

Assaf Amitai

MODELING THE INTERACTION OF A PROTEIN WITH CHROMATIN

We study here the following dynamical process. A protein (point particle) starts from a distance a from monomer, part of a long polymer. The protein diffuses until encountering a monomer again. Upon encounter, it binds for a certain amount of time and finally disassociates and diffuses until encountering another monomer to which it can bind (main text Fig.1a,b).

Our goal is to estimate the overall protein dynamics and the time it takes for the protein to find a specific monomer. Chromatin is represented by flexible chain polymer [1], which describes a linear chain of beads (monomers) connected with the following potential:

$$\phi(\boldsymbol{R}) = \frac{\kappa}{2}\sum_{n=2}^{N}\left(\boldsymbol{R}_n - \boldsymbol{R}_{n-1}\right)^2, \tag{1}$$

where $\boldsymbol{R} = (\boldsymbol{R}_1, \boldsymbol{R}_2, ..., \boldsymbol{R}_N)$, the spring constant $\kappa = dk_BT/b^2$ is related to the standard-deviation b of the distance between adjacent monomers [1], with k_B is the Boltzmann coefficient, T the temperature and d the dimensionality (dim 3). Since a chromatin locus (monomer) moves 3 orders of magnitude slower than a protein, we shall assume that the particle is moving while the monomer positions are drawn from the polymer equilibrium distribution in a domain Ω, which we will take to be a sphere of radius A representing the nucleus.

1. The monomer site are traps of size ϵ

The protein starts at $t = 0$ at a distance a from monomer n, whose position is given by $\boldsymbol{R}_n$ (main text Fig.1b). The protein diffuses until being absorbed on another monomer. Absorption occurs at a distance ϵ from the monomer (main text Fig.1b). The pdf of the protein position ($p(\boldsymbol{x}, t)$) satisfies the Forward-Fokker-Planck equation with mixed boundary conditions

$$\begin{aligned}
\frac{1}{D}\frac{\partial p(\boldsymbol{x},t)}{\partial t} &= \Delta p(\boldsymbol{x},t),\ \boldsymbol{x} \in \tilde{\Omega}_{\epsilon,N}, \\
P(\boldsymbol{x},0) &= \delta(\boldsymbol{x} - (\boldsymbol{R}_n + \boldsymbol{a})), \\
\frac{\partial p}{\partial \boldsymbol{n}_i} &= 0 \text{ for } \boldsymbol{x} \in \partial\Omega, \\
p(\boldsymbol{x},t) &= 0 \text{ for } \boldsymbol{x} \in \partial\Omega_a = \cup_{i=1}^{N}\partial\Omega_{\epsilon_i},
\end{aligned} \tag{2}$$

where D is the diffusion coefficient of the particle, $\tilde{\Omega}_{\epsilon,N} = \Omega \setminus \Omega_a = \cup_{i=1}^{N} \Omega_{\epsilon_i}$, $\partial\Omega_{\epsilon_i}$ is the surface of a ball centered around monomer i

$$\Omega_{\epsilon_i} := B(\boldsymbol{R}_i(t), \epsilon). \tag{3}$$

Finding the MFPT $\tau(\boldsymbol{x})$ for the protein to find any of the traps (monomers), starting from position $\boldsymbol{x}$ is equivalent to solving the following boundary value problem (b.v.p) [2]

$$\begin{aligned} \Delta\tau(\boldsymbol{x}) &= -\frac{1}{D},\ \boldsymbol{x} \in \tilde{\Omega}_{\epsilon,N}, \\ \partial_n \tau &= 0 \text{ for } \boldsymbol{x} \in \partial\Omega, \\ \tau(\boldsymbol{x}) &= 0 \text{ for } \boldsymbol{x} \in \partial\Omega_a = \cup_{i=1}^{N} \partial\Omega_{\epsilon_i}. \end{aligned} \tag{4}$$

2. The transition probability depends on the Neumann Green's in the spherical domain with traps

We first find the probability that the protein finds monomer l before finding any other monomer, which we denote the *transition* probability. Finding this probability is equivalent to solving the following b.v.p [3]

$$\begin{aligned} \Delta u_l(\boldsymbol{x}) &= 0,\ \boldsymbol{x} \in \tilde{\Omega}_{\epsilon,N}, \\ \partial_n u_l &= 0 \text{ for } \boldsymbol{x} \in \partial\Omega, \\ u_l(\boldsymbol{x}) &= 0 \text{ for } \boldsymbol{x} \in \partial\Omega_a = \cup_{i=1, i\neq l}^{N} \partial\Omega_{\epsilon_i}, \\ u_l(\boldsymbol{x}) &= 1 \text{ for } \boldsymbol{x} \in \partial\Omega_{\epsilon_l}. \end{aligned} \tag{5}$$

Starting from position $\boldsymbol{x}$, the probability u_l to find monomer l is given by [3]

$$u_l(\boldsymbol{x}) = \frac{C_l}{N\bar{C}} + 4\pi\epsilon C_l \left[G(\boldsymbol{x}, \boldsymbol{R}_l) - \frac{1}{N\bar{C}} \left(\sum_{j=1}^{N} C_j G(\boldsymbol{x}, \boldsymbol{R}_j) \right) \right] + \epsilon\chi_l, \tag{6}$$

where C_l is the capacity of monomer l in units of ϵ and $\bar{C}$ is the average capacity. Indeed, the absorbing monomers can be thought of as capacitors with certain boundary condition on their surface. The higher their capacity is, the larger is the probability that the particle will be captured by them. If we were to take non-spherical traps, that would change the capacity by some numerical constant ($\mathcal{O}(1)$) and thus, the capture time and probability.

$G(\boldsymbol{x},\boldsymbol{y})$ is the Neumann Green's function of the Laplacian in a sphere of radius A, which is given by [3]

$$G(\boldsymbol{x},\boldsymbol{y}) = \frac{1}{4\pi|\boldsymbol{x}-\boldsymbol{y}|} + \frac{A}{4\pi|\boldsymbol{x}|\boldsymbol{r}'} + \frac{1}{4\pi A}\log\left[\frac{2A^2}{A^2 - |\boldsymbol{x}||\boldsymbol{y}|\cos\theta + \boldsymbol{x}\boldsymbol{r}'}\right] + \frac{1}{6|\Omega|}\left(\boldsymbol{x}^2+\boldsymbol{y}^2\right) - \frac{7}{10\pi A}, \tag{7}$$

where $\boldsymbol{r}' = |\boldsymbol{x}'-\boldsymbol{y}|$, $\boldsymbol{x}' = \frac{\boldsymbol{x}A^2}{|\boldsymbol{x}|^2}$, $|\Omega|$ is the volume of the domain and

$$\chi_l = -\frac{4\pi C_l}{N\bar{C}}\left[(\boldsymbol{G}\boldsymbol{C})_l - \frac{1}{N}\boldsymbol{C}^T\boldsymbol{G}\boldsymbol{C}\right], \tag{8}$$

where

$$\boldsymbol{G}_{i,j} = \begin{cases} Q_i & i=j, \\ G_{ij}, & i\neq j, \end{cases} \tag{9}$$

where $G_{i,j} = G(\boldsymbol{R}_i(t),\boldsymbol{R}_j(t))$ and $Q_i(\boldsymbol{R}_i(t))$ is the regular part of the Green's function (self interacting term). The monomers capacity vector is

$$\boldsymbol{C}^T = (C_1, C_2, ..., C_N) \tag{10}$$

We take the monomer traps to be spherical. For spherical traps (eq.3) we have $C_i = \epsilon$. Thus

$$\chi_l = -\frac{4\pi C_l}{N\bar{C}}\left[\sum_{i\neq l} C_i G_{li} + C_l Q_l - \frac{1}{N\bar{C}}\left(\sum_{ij,i\neq j} C_i C_j G_{ij} + \sum_i C_i^2 Q_i\right)\right]. \tag{11}$$

We shall take $C_i = C$ for $i = 1..N$, thus

$$\chi_l = -\frac{4\pi C}{N}\left[\sum_{i\neq l} G_{li} + Q_l - \frac{1}{N}\left(\sum_{ij,i\neq j} G_{ij} + \sum_i Q_i\right)\right]. \tag{12}$$

From now on will work with spherical traps for which $C = 1$.

To facilitate the calculation we estimate G_{ij} using the equilibrium distribution of the monomers positions.

$$P_{eq}(\boldsymbol{R}) = \left(\frac{2\pi}{\kappa}\right)^{3(N-1)/2}\exp\left[-\frac{\kappa}{2}\sum_{n=2}^{N}\left(\boldsymbol{R}_n - \boldsymbol{R}_{n-1}\right)^2\right], \tag{13}$$

where we used the harmonic bead potential (eq.1). This pdf is that of a flexible polymer in bulk (free domain). When the size of the domain is of the order of the polymer radius of gyration ($R_g = \sqrt{N}b/6$) [1], the polymer configurations are affected by the domain wall.

Thus, we consider the limit $A \gg R_g$ ($R_g = \sqrt{N}b/6$) where of P_{eq} is valid. Additionally, to simplify the calculations we shall anchor the first monomer $n = 1$ at the origin.

Next we compute the stationary Green's function

$$\tilde{G}_{ij} = \int_{\Omega_N} G(\boldsymbol{R}_i, \boldsymbol{R}_j) P_{eq}(\boldsymbol{R})\, d\boldsymbol{R}. \tag{14}$$

We now estimate the different terms for the average Green's function

$$I_1 = \int_{\Omega_N} \frac{1}{4\pi|\boldsymbol{R}_i - \boldsymbol{R}_j|} P_{eq}(\boldsymbol{R})\, d\boldsymbol{R} = \frac{1}{b^3}\left[\frac{3}{2\pi|i-j|}\right]^{3/2} \int \frac{1}{4\pi|\boldsymbol{r}|} e^{-\frac{3r^2}{2|i-j|b^2}}\, d\boldsymbol{r} \tag{15}$$

$$= \frac{1}{b^3}\left[\frac{3}{2\pi|i-j|}\right]^{3/2} \int_0^\infty r e^{-\frac{3r^2}{2|i-j|b^2}}\, dr, \tag{16}$$

where we use the marginal pdf

$$P(\boldsymbol{R}_i, \boldsymbol{R}_j) = \frac{1}{b^3}\left[\frac{3}{2\pi|i-j|}\right]^{3/2} e^{-\frac{3(\boldsymbol{R}_i - \boldsymbol{R}_j)^2}{2|i-j|b^2}}. \tag{17}$$

Solving the integral, we get

$$I_1 = \frac{3^{1/2}}{(2\pi)^{3/2}|i-j|^{1/2}b}. \tag{18}$$

We assumed here that ϵ is small with respect to the distance between monomers: $|R_i - R_j| = \sqrt{i-j}b \ll \epsilon$. When ϵ is larger, this assumption breaks and the expression for the distance between the initial release site and the surface of the capture ball on the target monomer is more complicated. We further assume that we are in the limit $A \gg R_g$ where

$$\left\langle \frac{A}{4\pi|\boldsymbol{x}|\boldsymbol{r'}} \right\rangle \approx \frac{1}{4\pi A}, \tag{19}$$

and

$$\frac{1}{4\pi A}\left\langle \log\left[\frac{2A^2}{A^2 - |\boldsymbol{x}||\boldsymbol{y}|\cos\theta + \boldsymbol{x}\boldsymbol{r'}}\right]\right\rangle \approx \frac{1}{4\pi A}\log(1) = 0. \tag{20}$$

Finally

$$\langle R_n^2 \rangle = nb^2. \tag{21}$$

Substituting all the terms we find the average Green's function

$$\tilde{G}_{ij} = \frac{3^{1/2}}{(2\pi)^{3/2}|i-j|^{1/2}b} + \frac{b^2}{6|\Omega|}(i+j) - \frac{3A^2}{5|\Omega|}. \tag{22}$$

The regular part of the Green's function is

$$\tilde{Q}_i = \frac{2b^2 i}{6|\Omega|} - \frac{3A^2}{5|\Omega|}. \tag{23}$$

3. Estimating χ_l

First, we will approximate the double series sum in eq.(12)

$$\sum_{i\neq j,i=1}^{N} \tilde{G}_{ij}. \tag{24}$$

We will calculate each term separately. First

$$S_j = \sum_{i\neq j,i=1}^{N} \frac{1}{|i-j|^{1/2}} = \sum_{i=1}^{j-1} \frac{1}{(j-i)^{1/2}} + \sum_{i=j+1}^{N} \frac{1}{(i-j)^{1/2}} = S_{j1} + S_{j2}. \tag{25}$$

We will approximate this series sum using the Euler-Maclaurin (EM) formula

$$\sum_{i=a}^{b} f(i,c) \approx \int_a^b f(i,c)di + \frac{f(a,c)+f(b,c)}{2} + \frac{1}{12}(f^{'}(b,c) - f^{'}(a,c)), \tag{26}$$

where

$$f(i,c) = \frac{1}{(c-i)^{1/2}}. \tag{27}$$

Taking $a = 1, b = c-1$, the boundary terms are

$$\frac{f(a,c)+f(b,c)}{2} = \frac{1}{2}\left(\frac{1}{|c-1|^{1/2}} + 1\right). \tag{28}$$

The first derivative of (27) is

$$f'(i,c) = \frac{1}{2}\frac{1}{(c-i)^{3/2}}. \tag{29}$$

Hence, the boundary derivative terms are

$$\frac{1}{12}(f^{'}(i,c) - f^{'}(i,c)) = \frac{1}{12}\left(\frac{1}{2} - \frac{1}{2|c-1|^{3/2}}\right) \tag{30}$$

and the integral

$$\int_1^{c-1} \frac{1}{(c-i)^{1/2}} di = -2(c-i)^{1/2}|_1^{c-1} = 2(c-1)^{1/2} - 2. \tag{31}$$

Combining all the terms:

$$\begin{aligned} S_{j1} = \sum_{i=1}^{j-1} \frac{1}{(j-i)^{1/2}} &= 2(j-1)^{1/2} - 2 + \frac{1}{2}\left(S_{j1} = \frac{1}{|j-1|^{1/2}} + 1\right) \\ &+ \frac{1}{12}\left(\frac{1}{2} - \frac{1}{2|j-1|^{3/2}}\right), \end{aligned} \tag{32}$$

while

$$S_{j2} = \sum_{i=j+1}^{N} \frac{1}{(i-j)^{1/2}} = 2(N-j)^{1/2} - 2 + \frac{1}{2}\left(\frac{1}{|N-j|^{1/2}} + 1\right) + \frac{1}{12}\left(\frac{1}{2} - \frac{1}{2|N-j|^{3/2}}\right). \tag{33}$$

Thus

$$S_j = \sum_{i\neq j, i=1}^{N} \frac{1}{|i-j|^{1/2}} = 2((N-j)^{1/2} + (j-1)^{1/2}) + \frac{1}{2}\left(\frac{1}{|N-j|^{1/2}} + \frac{1}{|j-1|^{1/2}}\right) - \frac{1}{12}\left(\frac{1}{2|N-j|^{3/2}} + \frac{1}{2|j-1|^{3/2}}\right) - \frac{35}{12}, \quad \text{for } j = 2..N-1, \tag{34}$$

while

$$S_N = \sum_{i=1}^{N-1} \frac{1}{|N-i|^{1/2}} = 2(N-1)^{1/2} + \frac{1}{2}\frac{1}{|N-1|^{1/2}} - \frac{1}{24}\frac{1}{|N-1|^{3/2}} - \frac{35}{24}. \tag{35}$$

and

$$S_1 = \sum_{i=2}^{N} \frac{1}{(i-j)^{1/2}} = 2(N-1)^{1/2} + \frac{1}{2}\frac{1}{|N-1|^{1/2}} - \frac{1}{24}\frac{1}{|N-1|^{3/2}} - \frac{35}{24}. \tag{36}$$

We now sum S_j over all the monomers:

$$S = S_1 + S_N + \sum_{j=2}^{N-1} S_j = 2\left(\sum_{j=1}^{N-1}(N-j)^{1/2} + \sum_{j=2}^{N}(j-1)^{1/2}\right) + \frac{1}{2}\left(\sum_{j=1}^{N-1}\frac{1}{|N-j|^{1/2}} + \sum_{j=2}^{N}\frac{1}{|j-1|^{1/2}}\right) - \frac{1}{12}\left(\sum_{j=1}^{N-1}\frac{1}{2|N-j|^{3/2}} + \sum_{j=2}^{N}\frac{1}{2|j-1|^{3/2}}\right) - \frac{35}{12} - \sum_{j=2}^{N-1}\left(\frac{35}{12}\right). \tag{37}$$

Again, we use EM formula to estimate the series terms in S (eq.37)

$$s_1 = \sum_{j=1}^{N-1}(N-j)^{1/2} \approx \frac{2}{3}(N-1)^{3/2} + \frac{1}{2}\left((N-1)^{1/2}+1\right) + \frac{1}{24}\left((N-1)^{-1/2}-1\right), \tag{38}$$

$$s_2 = \sum_{j=2}^{N}(j-1)^{1/2} \approx \frac{2}{3}(N-1)^{3/2} + \frac{1}{2}\left((N-1)^{1/2}+1\right) + \frac{1}{24}\left((N-1)^{-1/2}-1\right), \tag{39}$$

$$s_3 \approx \sum_{j=1}^{N-1}\frac{1}{(N-j)^{1/2}} \approx 2 - 2(N-1)^{1/2}, \tag{40}$$

$$s_4 \approx \sum_{j=2}^{N}\frac{1}{(j-1)^{1/2}} \approx 2(N-1)^{1/2} - 2, \tag{41}$$

$$s_5 \approx \sum_{j=1}^{N-1} \frac{1}{(N-j)^{3/2}} \approx -2(N-1)^{-1/2} + 2, \tag{42}$$

$$s_6 \approx \sum_{j=1}^{N-1} \frac{1}{(N-j)^{3/2}} \approx -2(N-1)^{-1/2} + 2, \tag{43}$$

Introducing back into eq.(37)

$$\begin{aligned} S &= 2(s_1+s_2) + \frac{1}{2}(s_3+s_4) - \frac{1}{24}(s_5+s_6) - \frac{35(N-1)}{12} \\ &= \frac{8}{3}(N-1)^{3/2} + 2(N-1)^{1/2} + \frac{1}{3}(N-1)^{-1/2} - \frac{55}{12} - \frac{35N}{12}. \end{aligned} \tag{44}$$

For large N, we take only terms of $\mathcal{O}(1)$ or higher in S:

$$S \approx \frac{8}{3}(N-1)^{3/2} - \frac{35N}{12} + 2(N-1)^{1/2} - \frac{55}{12}. \tag{45}$$

Now we use

$$\sum_{1}^{N} i = \frac{N(N+1)}{2}. \tag{46}$$

Thus, the N^{-2} term in χ_l (eq.12), when substituting eqs.23, 45 and 46

$$\begin{aligned} \sum_{ij;j\neq i} G_{ij} + \sum_i Q_i = \frac{3^{1/2}}{(2\pi)^{3/2}b}\left(\frac{8}{3}(N-1)^{3/2} - \frac{35N}{12} + 2(N-1)^{1/2} - \frac{55}{12}\right) \\ + \frac{b^2N^2(N+1)}{6|\Omega|} - \frac{3A^2N^2}{5|\Omega|}. \end{aligned} \tag{47}$$

To estimate the N^{-1} term in χ_l (eq.12) we introduce eq.(34), taking only terms of $\mathcal{O}(1)$ or higher. For $2 < l < N$, we find

$$\begin{aligned} \sum_{i\neq l} G_{li} + Q_l = \frac{3^{1/2}}{(2\pi)^{3/2}b}\left(2\left((N-l)^{1/2} + (l-1)^{1/2}\right) - \frac{35}{12}\right) \\ + \frac{N}{|\Omega|}\left[\frac{b^2(N+1+2l)}{12} - \frac{3A^2}{5}\right], \end{aligned} \tag{48}$$

while

$$\sum_{i\neq 1} G_{1i} + R_1 = \frac{3^{1/2}}{(2\pi)^{3/2}b}\left(2(N-1)^{1/2} - \frac{35}{24}\right) + \frac{N}{|\Omega|}\left[\frac{b^2(N+3)}{12} - \frac{3A^2}{5}\right], \tag{49}$$

and

$$\sum_{i\neq N} G_{Ni} + R_N = \frac{3^{1/2}}{(2\pi)^{3/2}b}\left(2(N-1)^{1/2} - \frac{35}{24}\right) + \frac{N}{|\Omega|}\left[\frac{b^2(3N+1)}{12} - \frac{3A^2}{5}\right]. \tag{50}$$

Introducing (48), (49), (50) and (47) into (12):

$$\chi_l = -\frac{4\pi}{N}\left[\frac{3^{1/2}}{(2\pi)^{3/2}b}\left(2(N-l)^{1/2}+2(l-1)^{1/2}-\frac{8}{3}\frac{(N-1)^{3/2}}{N}-\frac{2(N-1)^{1/2}}{N}+\frac{55}{12N}\right)\right.$$
$$\left.+\frac{Nb^2(2l-(N+1))}{12|\Omega|}\right] \approx -\frac{4\pi}{N}\left[\frac{3^{1/2}}{(2\pi)^{3/2}b}\left(2(N-l)^{1/2}+2(l-1)^{1/2}-\frac{8(N-1)^{3/2}}{3N}\right)\right.$$
$$\left.+\frac{Nb^2(2l-(N+1))}{12|\Omega|}\right]. \tag{51}$$

We we took terms of order $\mathcal{O}\left(N^{1/2}\right)$ or higher

$$\chi_1 = -\frac{4\pi}{N}\left[\frac{3^{1/2}}{(2\pi)^{3/2}b}\left(2(N-1)^{1/2}-\frac{8}{3}\frac{(N-1)^{3/2}}{N}-\frac{2(N-1)^{1/2}}{N}+\frac{35}{24}+\frac{55}{12N}\right)-\frac{Nb^2(N+1)}{12|\Omega|}\right]$$
$$\approx -\frac{4\pi}{N}\left[\frac{3^{1/2}}{(2\pi)^{3/2}b}\left(2(N-1)^{1/2}-\frac{8(N-1)^{3/2}}{3N}\right)-\frac{Nb^2(N+1)}{12|\Omega|}\right] \tag{52}$$

For the last monomer:

$$\chi_N = -\frac{4\pi}{N}\left[\frac{3^{1/2}}{(2\pi)^{3/2}b}\left(2(N-1)^{1/2}-\frac{8}{3}\frac{(N-1)^{3/2}}{N}-\frac{2(N-1)^{1/2}}{N}+\frac{35}{24}+\frac{55}{12N}\right)+\frac{Nb^2(N+1)}{12|\Omega|}\right]$$
$$\approx -\frac{4\pi}{N}\left[\frac{3^{1/2}}{(2\pi)^{3/2}b}\left(2(N-1)^{1/2}-\frac{8(N-1)^{3/2}}{3N}\right)+\frac{Nb^2(N+1)}{12|\Omega|}\right]. \tag{53}$$

4. Computing the transition probability $u_l(n)$

We continue to estimate the site-to-site transition probability $u_l(n)$ by averaging over eq.6. We assume that the particle started at a distance $a \ll b$ from monomer $\boldsymbol{R}_n$. We thus need to estimate

$$S_{\boldsymbol{x},l} = \sum_{j=1}^{N} G(\boldsymbol{x}, \boldsymbol{R}_j), \tag{54}$$

that we write as

$$S_{\boldsymbol{x},l} = G(\boldsymbol{x}, \boldsymbol{R}_n) + \sum_{j=1, j\neq n}^{N} G(\boldsymbol{x}, \boldsymbol{R}_j). \tag{55}$$

Since we take $|\boldsymbol{x} - \boldsymbol{R}_n(t)| = a$, we estimate the sum over the Green's function as

$$\sum_{j=1, j\neq n}^{N} G(\boldsymbol{x}, \boldsymbol{R}_j) \approx \sum_{j=1, j\neq n}^{N} G(\boldsymbol{R}_n, \boldsymbol{R}_j). \tag{56}$$

We replace again the Green's function $G(\boldsymbol{R}_n, \boldsymbol{R}_j)$ by $\tilde{G}_{nj}$:

$$S_{\boldsymbol{x},n} \approx \langle G(\boldsymbol{x}, \boldsymbol{R}_n)\rangle + \sum_{j=1, j\neq n}^{N} \tilde{G}_{nj}. \tag{57}$$

Introducing expression 48 to eq.(57)

$$\begin{aligned} S_{\boldsymbol{x},n} \approx & \frac{1}{4\pi a} + \frac{3^{1/2}}{(2\pi)^{3/2} b}\left(2\left((N-n)^{1/2} + (n-1)^{1/2}\right) - \frac{35}{12}\right) \\ & + \frac{N}{|\Omega|}\left[\frac{b^2(N+1+2n)}{12} - \frac{3A^2}{5}\right], \end{aligned} \tag{58}$$

where we used that

$$\left\langle \frac{1}{4\pi|\boldsymbol{r}|} \right\rangle = \int \frac{1}{4\pi|\boldsymbol{r}|} P(\boldsymbol{r}, a) r^2 \sin(\theta) dr d\theta d\phi = \frac{1}{4\pi a}, \tag{59}$$

with $\boldsymbol{r} = \boldsymbol{x} - \boldsymbol{R}_n$ and the probability distribution of the particle is

$$P(\boldsymbol{r}, a) = \frac{\delta(\boldsymbol{r} - a)}{4\pi \boldsymbol{r}^2}. \tag{60}$$

Substituting eqs. (22), (23), (58), (51) all the elements to eq.6, we find that for $2 < n < N$

$$\begin{aligned} u_l(n) &= \frac{1}{N} + 4\pi\epsilon\left[G(\boldsymbol{x}, \boldsymbol{R}_l) - \frac{1}{N}\left(\sum_{j=1}^{N} G(\boldsymbol{x}, \boldsymbol{R}_j)\right)\right] + \epsilon\chi_l \\ &= \frac{1}{N} + 4\pi\epsilon\left[\frac{3^{1/2}}{(2\pi)^{3/2} b}\left(|n-l|^{-1/2} - \frac{2}{N}\Big((N-l)^{1/2} + (l-1)^{1/2} + (N-n)^{1/2}\right.\right. \\ &\quad \left.\left. + (n-1)^{1/2} - \frac{4(N-1)^{3/2}}{3N} - \frac{35}{24}\Big)\right) - \frac{1}{4\pi a N}\right], \text{ for } 2 < l < N \text{ , } l \neq n \end{aligned} \tag{61}$$

$$\begin{aligned} u_1(n) = \frac{1}{N} + 4\pi\epsilon\Bigg[\frac{3^{1/2}}{(2\pi)^{3/2} b}\Bigg(|n-1|^{-1/2} - \frac{2}{N}\Big(2(N-1)^{1/2} + (N-n)^{1/2} + (n-1)^{1/2} \\ - \frac{4(N-1)^{3/2}}{3N} - \frac{35}{24}\Big)\Bigg) - \frac{1}{4\pi a N} + \frac{b^2}{6|\Omega|}\Bigg]. \end{aligned} \tag{62}$$

$$\begin{aligned} u_N(n) = \frac{1}{N} + 4\pi\epsilon\Bigg[\frac{3^{1/2}}{(2\pi)^{3/2} b}\Bigg(|N-n|^{-1/2} - \frac{2}{N}\Big(2(N-1)^{1/2} + (N-n)^{1/2} + (n-1)^{1/2} \\ - \frac{4(N-1)^{3/2}}{3N} - \frac{35}{24}\Big)\Bigg) - \frac{1}{4\pi a N} - \frac{b^2}{6|\Omega|}\Bigg]. \end{aligned} \tag{63}$$

$$u_n(n) = \frac{1}{N} + 4\pi\epsilon\Bigg[\frac{3^{1/2}}{(2\pi)^{3/2}b}\Bigg(- \frac{4}{N}\Big((N-n)^{1/2} + (n-1)^{1/2} - \frac{2(N-1)^{3/2}}{3N} - \frac{35}{48}\Big)\Bigg) + \frac{1}{4\pi a}\left(1 - \frac{1}{N}\right)\Bigg]. \tag{64}$$

For the end monomer $(n = 1)$

$$u_1(1) = \frac{1}{N} + 4\pi\epsilon\Bigg[\frac{3^{1/2}}{(2\pi)^{3/2}b}\Bigg(- \frac{4}{N}\Big((N-1)^{1/2} - \frac{2(N-1)^{3/2}}{3N} - \frac{35}{96}\Big)\Bigg) + \frac{1}{4\pi a}\left(1 - \frac{1}{N}\right) + \frac{b^2}{6|\Omega|}\Bigg]. \tag{65}$$

For $2 < l < N$

$$u_l(1) = \frac{1}{N} + 4\pi\epsilon\Bigg[\frac{3^{1/2}}{(2\pi)^{3/2}b}\Bigg((l-1)^{-1/2} - \frac{2}{N}\Big((N-l)^{1/2} + (l-1)^{1/2} + (N-1)^{1/2} - \frac{4(N-1)^{3/2}}{3N} - \frac{35}{48}\Big)\Bigg) - \frac{1}{4\pi aN}\Bigg], \tag{66}$$

and for $l = N$

$$u_N(1) = \frac{1}{N} + 4\pi\epsilon\Bigg[\frac{3^{1/2}}{(2\pi)^{3/2}b}\Bigg((N-1)^{-1/2} - \frac{4}{N}\Big((N-1)^{1/2} - \frac{2(N-1)^{3/2}}{3N} - \frac{35}{96}\Big)\Bigg) - \frac{1}{4\pi Na} - \frac{b^2}{6|\Omega|}\Bigg]. \tag{67}$$

For $n = N$

$$u_N(N) = \frac{1}{N} + 4\pi\epsilon\Bigg[\frac{3^{1/2}}{(2\pi)^{3/2}b}\Bigg(- \frac{4}{N}\Big((N-1)^{1/2} - \frac{2(N-1)^{3/2}}{3N} - \frac{35}{96}\Big)\Bigg) + \frac{1}{4\pi a}\left(1 - \frac{1}{N}\right) - \frac{b^2}{6|\Omega|}\Bigg]. \tag{68}$$

For $2 < l < N$

$$u_l(N) = \frac{1}{N} + 4\pi\epsilon\Bigg[\frac{3^{1/2}}{(2\pi)^{3/2}b}\Bigg((N-l)^{-1/2} - \frac{2}{N}\Big((N-l)^{1/2} + (l-1)^{1/2} + (N-1)^{1/2} - \frac{4(N-1)^{3/2}}{3N} - \frac{35}{48}\Big)\Bigg) - \frac{1}{4\pi aN}\Bigg], \tag{69}$$

and finally

$$u_1(N) = \frac{1}{N} + 4\pi\epsilon\Bigg[\frac{3^{1/2}}{(2\pi)^{3/2}b}\Bigg((N-1)^{-1/2} - \frac{4}{N}\Big((N-1)^{1/2} - \frac{2(N-1)^{3/2}}{3N} - \frac{35}{96}\Big)\Bigg) - \frac{1}{4\pi Na} + \frac{b^2}{6|\Omega|}\Bigg]. \tag{70}$$

5. Computing the conditional MFPT $\tau_l(n)$ to monomer l starting from monomer n

We next compute the conditional mean first passage time (MFPT) $\tau_l(\boldsymbol{x})$ of the particle to a specific monomer site l before encountering any of the other monomers, given it started at position $\boldsymbol{x}$. The conditional dynamics of the particle obeys the Langevin equation [4]

$$d\boldsymbol{x} = \boldsymbol{a}(\boldsymbol{x})dt + \sqrt{2D}d\boldsymbol{\omega}, \tag{71}$$

where

$$\boldsymbol{a}(\boldsymbol{x}) = 2D\frac{\nabla u_l(\boldsymbol{x})}{u_l(\boldsymbol{x})}, \tag{72}$$

with $u_l(\boldsymbol{x})$ the transition probability (eq.5). Moving according to this Langevin equation, the particle experiences a drift pushing it away from all monomers except for its final destination l. The MFPT to find l is the solution of

$$\begin{aligned}
\Delta\tau_l(\boldsymbol{x}) + 2\frac{\nabla u_l(\vec{x})}{u_l(\vec{x})}\cdot\nabla\tau_l(\boldsymbol{x}) &= -\frac{1}{D},\ \boldsymbol{x}\in\tilde{\Omega}_{\epsilon,N}, \\
P(\boldsymbol{x},0) &= \delta(\boldsymbol{x}-(\boldsymbol{R}_n+\boldsymbol{a})), \\
\partial_n\tau_l &= 0 \text{ for } \boldsymbol{x}\in\partial\Omega, \\
\tau_l(\boldsymbol{x}) &= 0 \text{ for } \boldsymbol{x}\in\partial\Omega_a = \Omega_{\epsilon_l}.
\end{aligned} \tag{73}$$

To estimate $\tau_l(\boldsymbol{x})$, we define the variable $v_l = \tau_l(\boldsymbol{x})u_l(\boldsymbol{x})$. Hence, $v_l(\boldsymbol{x})$ is the solution of the b.v.p

$$\begin{aligned}
D\Delta v_l(\boldsymbol{x}) &= -u_l(\boldsymbol{x}),\ \boldsymbol{x}\in\tilde{\Omega}_{\epsilon,N}, \\
\partial_n v_l &= 0 \text{ for } \boldsymbol{x}\in\partial\Omega, \\
v_l(\boldsymbol{x}) &= 0 \text{ for } \boldsymbol{x}\in\partial\Omega_a = \cup_{i=1}^{N}\partial\Omega_{\epsilon_i}.
\end{aligned} \tag{74}$$

The solution $v_l(\boldsymbol{x})$ can be written as the eigenfunctions expansion

$$v_l(\boldsymbol{x}) = \sum_{i=0}^{\infty} a_i w_{i,\epsilon}(\boldsymbol{x}), \tag{75}$$

where $w_{i,\epsilon}(\boldsymbol{x})$ are the eigenfunctions of the Laplacian in the perturbed domain $\tilde{\Omega}_{\epsilon,N}$ [5], obeying the equation

$$\Delta w_{i,\epsilon}(\boldsymbol{x}) + \lambda_{i,\epsilon} w_{i,\epsilon}(\boldsymbol{x}) = 0, \tag{76}$$

with $\lambda_{i,\epsilon}(\boldsymbol{x})$ the corresponding eigenvalues of the homogeneous.

Introducing the expansion to eq.(74), we get

$$D\sum_{i=0}^{\infty} a_i \lambda_{i,\epsilon} w_{i,\epsilon}(\boldsymbol{x}) = u_l(\boldsymbol{x}), \tag{77}$$

To find the coefficients a_i we multiply by $w_{j,\epsilon}(\boldsymbol{x})$ and integrate over the whole domain. Using the orthonormality of the eigenfunctions

$$Da_j\lambda_{j,\epsilon} = \int_{\tilde{\Omega}_{\epsilon,N}} w_{j,\epsilon}(\boldsymbol{x}) u_l(\boldsymbol{x}) dV, \tag{78}$$

giving

$$a_j = \frac{\langle w_{j,\epsilon}(\boldsymbol{x}) u_l(\boldsymbol{x})\rangle}{D\lambda_{j,\epsilon}}. \tag{79}$$

Introducing back to eq.(75)

$$v_l(\boldsymbol{x}) = \sum_{i=0}^{\infty} \frac{\langle w_{j,\epsilon}(\boldsymbol{x}) u_l(\boldsymbol{x})\rangle w_{i,\epsilon}(\boldsymbol{x})}{D\lambda_{j,\epsilon}}. \tag{80}$$

We shall approximate $v_l(\boldsymbol{x})$ with the first term

$$v_l(\boldsymbol{x}) \approx \frac{\langle w_{0,\epsilon}(\boldsymbol{x}) u_l(\boldsymbol{x})\rangle w_{0,\epsilon}(\boldsymbol{x})}{D\lambda_{0,\epsilon}}. \tag{81}$$

The perturbed zero order eigenfunction is given by

$$w_{0,\epsilon}(\boldsymbol{x}) = \frac{1}{|\tilde{\Omega}_{\epsilon,N}|^{1/2}} - \frac{4\pi\epsilon}{|\tilde{\Omega}_{\epsilon,N}|^{1/2}} \sum_{j=1}^{N} G(\boldsymbol{x}, \boldsymbol{R}_j) + \mathcal{O}(\epsilon^2). \tag{82}$$

The coefficient is

$$\begin{aligned}
\langle w_{0,\epsilon}(\boldsymbol{x}) u_l(\boldsymbol{x})\rangle \approx \int_{\tilde{\Omega}_{\epsilon,N}} \Bigg[& \frac{1}{|\tilde{\Omega}_{\epsilon,N}|^{1/2}} \times \frac{1}{N} + \frac{1}{|\tilde{\Omega}_{\epsilon,N}|^{1/2}} \times 4\pi\epsilon \Bigg[G(\boldsymbol{x}, \boldsymbol{R}_l) \\
& - \frac{1}{N}\left(\sum_{j=1}^{N} G(\boldsymbol{x}, \boldsymbol{R}_j)\right) \Bigg] - \frac{4\pi\epsilon}{|\tilde{\Omega}_{\epsilon,N}|^{1/2}} \sum_{i=1}^{N} G(\boldsymbol{x}, \boldsymbol{R}_j) \times \frac{1}{N} + \frac{1}{|\tilde{\Omega}_{\epsilon,N}|^{1/2}} \epsilon\chi_l \Bigg] \\
= & \frac{|\tilde{\Omega}_{\epsilon,N}|^{1/2}}{N} + |\tilde{\Omega}_{\epsilon,N}|^{1/2}\epsilon\chi_l + \mathcal{O}(\epsilon^2),
\end{aligned} \tag{83}$$

where we used the definition of $u_l(\boldsymbol{x})$ (eq.6) and took only order $\mathcal{O}(\epsilon)$ terms. We also used

$$\int_{\tilde{\Omega}_{\epsilon,N}} G(\boldsymbol{x}, \boldsymbol{R}_j) = 0. \tag{84}$$

The corresponding eigenvalue is

$$\lambda_{0,\epsilon} = \frac{4\pi\epsilon N}{|\Omega|} - \frac{(4\pi\epsilon)^2}{|\Omega|}\left(\sum_{i,j;i\neq j} G_{ij} + \sum_i Q_i\right) + \mathcal{O}(\epsilon^2). \tag{85}$$

We recall that

$$\sum_{ij;j\neq i} G_{ij} + \sum_i Q_i = \frac{3^{1/2}}{(2\pi)^{3/2}b}\left(\frac{8}{3}(N-1)^{3/2} - \frac{35N}{12} + 2(N-1)^{1/2} - \frac{55}{12}\right) + \frac{b^2N^2(N+1)}{6|\Omega|} - \frac{3A^2N^2}{5|\Omega|}. \tag{86}$$

Thus

$$\lambda_{0,\epsilon} = \frac{4\pi C\epsilon N}{|\Omega|} - \frac{(4\pi C\epsilon)^2}{|\Omega|}\left(\frac{3^{1/2}}{(2\pi)^{3/2}b}\left(\frac{8}{3}(N-1)^{3/2} - \frac{35N}{12} + 2(N-1)^{1/2} - \frac{55}{12}\right) + \frac{b^2N^2(N+1)}{6|\Omega|} - \frac{3A^2N^2}{5|\Omega|}\right) + \mathcal{O}(\epsilon^2). \tag{87}$$

We recall that $\tau_l = v_l/u_l$

$$\tau_l(\boldsymbol{x}) \approx \frac{\langle w_{0,\epsilon}(\boldsymbol{x})u_l(\boldsymbol{x})\rangle w_{0,\epsilon}(\boldsymbol{x})}{D\lambda_{0,\epsilon}u_l(\boldsymbol{x})} = \frac{1}{D\lambda_{0,\epsilon}}\left(\frac{1}{N} + \epsilon\chi_l\right)\left[\frac{1}{u_l(\boldsymbol{x})} - 4\pi C\epsilon\sum_{j=1}^{N}\frac{G(\boldsymbol{x},\boldsymbol{R}_j)}{u_l(\boldsymbol{x})}\right]. \tag{88}$$

The particle starts its motion from the a neighborhood of monomer n. We shall use again the pre-averaging approximation

$$\int\int_{\Omega_N}\frac{1}{u_l(\boldsymbol{x})}P_{eq}\left(\boldsymbol{R}\right)\delta(|\boldsymbol{x}-\boldsymbol{R}_n| = a)d\boldsymbol{x}d\boldsymbol{R} \approx \frac{1}{u_l(n)}, \tag{89}$$

where $\boldsymbol{x} \in B(\boldsymbol{R}_n(t), a)$. For $n = l$

$$\begin{aligned}&\int\int_{\Omega_N}\frac{G(\boldsymbol{x},\boldsymbol{R}_j)}{u_l(\boldsymbol{x})}P_{eq}\left(\boldsymbol{R}\right)\delta(|\boldsymbol{x}-\boldsymbol{R}_l| = a)d\boldsymbol{x}d\boldsymbol{R}\\ &\approx \frac{1}{u_l(l)}\int_{\Omega_N}G(\boldsymbol{x},\boldsymbol{R}_j)P_{eq}\left(\boldsymbol{R}\right)d\boldsymbol{R} \approx \frac{\tilde{G}_{nj}}{u_n(n)},\end{aligned} \tag{90}$$

and

$$\tau_n(n) \approx \frac{1}{D\lambda_{0,\epsilon}}\left(\frac{1}{N} + \epsilon\chi_n\right)\left[\frac{1}{u_n(n)} - \frac{4\pi C\epsilon}{u_n(n)}\sum_{j=1}^{N}\tilde{G}_{nj}\right]. \tag{91}$$

When $n \neq l$, we use the Taylor approximation to expand the encounter probability $u_l(\boldsymbol{x})$ for small ϵ

$$\begin{aligned}&\sum_j\int\int_{\Omega_N}\frac{G(\boldsymbol{x},\boldsymbol{R}_j)}{u_l(\boldsymbol{x})}P_{eq}\left(\boldsymbol{R}\right)\delta(|\boldsymbol{x}-\boldsymbol{R}_n| = a)d\boldsymbol{x}d\boldsymbol{R} \approx \sum_j\int\int_{\Omega_N}G(\boldsymbol{x},\boldsymbol{R}_j)\Bigg[N\\ &-4\pi\epsilon N^2\left(G(\boldsymbol{x},\boldsymbol{R}_l) - \frac{1}{N}\sum_{i=1}^{N}G(\boldsymbol{x},\boldsymbol{R}_i) + \chi_l\right)\Bigg]P_{eq}\left(\boldsymbol{R}\right)\delta(|\boldsymbol{x}-\boldsymbol{R}_n| = a)d\boldsymbol{x}d\boldsymbol{R}\\ &\approx \int\int_{\Omega_N}G(\boldsymbol{x},\boldsymbol{R}_n)\left[N - 4\pi\epsilon N^2\left(G(\boldsymbol{x},\boldsymbol{R}_l) - \frac{1}{N}\sum_{i=1}^{N}G(\boldsymbol{x},\boldsymbol{R}_i) + \chi_l\right)\right]\\ &P_{eq}\left(\boldsymbol{R}\right)\delta(|\boldsymbol{x}-\boldsymbol{R}_n| = a)d\boldsymbol{x}d\boldsymbol{R} = N\sum_j\tilde{G}_{nj} - 4\pi\epsilon N^2\tilde{G}_{nn}\left[\tilde{G}_{nl} + \chi_l - \frac{1}{N}\sum_j\tilde{G}_{nj}\right]\end{aligned} \tag{92}$$

where, assuming that $a \ll b$, we only take one element in the sum over all monomer, where the singular term scales as a^{-1}. This happens for $j = n$. Finally, substituting expression (92) to eq.(88), we find an approximation for the MFPT

$$\tau_l(n) \approx \frac{N^{-1} + \epsilon\chi_l}{D\lambda_{0,\epsilon}} \left[\frac{1}{u_l(n)} - 4\pi\epsilon N \left(\sum_j \tilde{G}_{nj} - 4\pi\epsilon N \tilde{G}_{nn} \left[\tilde{G}_{nl} + \chi_l - \frac{1}{N} \sum_j \tilde{G}_{nj} \right] \right) \right] \quad (93)$$

We estimated here the MFPT using the first eigenvalue of the Laplacian and the long time asymptotic of the heat kernel [6]. Since the particle starts close to its target, higher eigenvalues will contribute to the MFPT [7], and our expression for the MFPT can be extended to include them by estimate the contribution of higher term eigenvalues.

6. Site-to-site transition rates as a CTMC

The particle can be either bound to one of the N monomers in the domain or diffuse between dis-association and capture events. We denote the unbound state as *in-transit* from one monomer site to the another. We assume that at site n, the particle remains bound for a characteristic time T_n, and leaves with a Poissonian rate. We define the process as a continuous time Markov chain (CTMC) (Fig.3a), where the release rate from site n is

$$q_{n,Tr_{nl}} = T_n^{-1} u_l(n), \text{ for } n = 1...N. \quad (94)$$

When the destination site l is far, the transition time is well described by a single exponent (eq.90 and [5]). We assume that the in-transit time until reattachment is Poisson-distributed. Thereby, the transition rate from the in-transit state Tr_{nl} to monomer site l has probability 1 and a rate

$$q_{Tr_{nl},l} = \tau_l(n)^{-1}. \quad (95)$$

When the particle is captured rapidly, multiple exponent contribute to the the transition time, that can be taken into account by defining multiple in-transit states. The $N(N+1) \times N(N+1)$ rate matrix Q whose entries are the rates $q_{n,Tr_{nl}}$ and $q_{Tr_{nl},l}$ describes the transitions between all states and can be used to find the time-dependent probability distribution function $P(\boldsymbol{S}, t)$, with the state vector $\boldsymbol{S} = (s_{b1}, s_{b2}, ..., s_{bN}, s_{it11}, s_{it12}, ..., s_{itNN})$, where the first N states are bound state and the next N^2, are in-transit states. The probability distribution function of the particle is given by

$$p(\boldsymbol{S}, t) = e^{Qt} p(\boldsymbol{S}, 0), \quad (96)$$

where $p(\boldsymbol{S}, 0)$ is the initial distribution of the particle in the different states.

THE REATTACHMENT PROBABILITY DEPENDS ON THE LOCAL DENSITY

To clarify the effect of monomer density on the reattachment probability, we estimated the local density around different monomers along the chain. For a given monomer m, the probability to find the n-th monomer at 3d distance between r and $r + dr$ is

$$p(r, n, m) = \left(\frac{1}{\sqrt{2\pi |m-n| \sigma^2}} \right)^3 e^{-\frac{r^2}{2|m-n|\sigma^2}}, \tag{97}$$

where $\sigma^2 = \frac{b^2}{3}$. The distribution $p(r, n)$ is normalized such that $4\pi \int_0^\infty p(r, n) r^2 dr = 1$.

The probability to find monomer n within a ball of radius h around m is ($\sigma_{m,n} = \sqrt{|n-m|}\sigma$)

$$\begin{aligned} P(h, n, m) &= 4\pi \int_0^h p(r, n) r^2 dr = \frac{2}{\sqrt{2\pi}} \int_0^{\frac{h}{\sigma_{n,m}}} e^{-\frac{x^2}{2}} x^2 dx = \frac{4}{\sqrt{\pi}} \int_0^{\frac{h}{\sqrt{2}\sigma_{n,m}}} e^{-x^2} x^2 dx \\ &= \mathrm{Erf}(y) - \frac{2}{\sqrt{\pi}} y e^{-y^2}, \end{aligned} \tag{98}$$

with $y = \frac{h}{\sqrt{2}\sigma_{n,m}}$.

The total monomer density around monomer m given that it is part of a polymer of length N is

$$\rho_m = \sum_{i=1}^{m-1} P(h, n, m) + \sum_{i=m+1}^{N} P(h, n, m). \tag{99}$$

In SM Fig.1 we plot the reattachment probability $u_n(n)$ as a function of ρ_m computed for different monomers along the chain. Since the middle monomer is closest to the center of mass of the polymer, the local density there is highest. Thus, the released particle is easily lost to other monomer and $u_n(n)$ is minimal. When the ball surrounding the release monomer m is larger, the average number of monomer within it will be larger (Fig.1a vs. b)

SELF AVOIDING POLYMER

We simulated the particle release and capture process for a self-avoiding polymer. For the purpose of the simulations, self-avoiding interactions were introduced by adding the

Lennard-Jones potential [8]

$$\phi_{LJ}(\boldsymbol{R}_1, ..\boldsymbol{R}_N) = \sum_{i,j,i\neq j} \phi^{i,j}_{\rm LJ}(\boldsymbol{r}_{i,j}), \tag{100}$$

with $\boldsymbol{r}_{i,j} = \boldsymbol{R}_i - \boldsymbol{R}_j$ and

$$\phi^{i,j}_{\rm LJ}(\boldsymbol{r}_{i,j}) = \begin{cases} 4\left[\left(\frac{\sigma}{\boldsymbol{r}_{i,j}}\right)^{12} - 2\left(\frac{\sigma}{\boldsymbol{r}_{i,j}}\right)^{6} + \frac{1}{4}\right] & \text{for } |\boldsymbol{r}_{i,j}| \geq 2^{1/6}\sigma \\ 0 & \text{for } |\boldsymbol{r}_{i,j}| < 2^{1/6}\sigma. \end{cases} \tag{101}$$

[1] M. Doi and S. F. Edwards, *The Theory of Polymer Dynamics* (Oxford: Clarendon Press, 1986).

[2] Z. Schuss, *Diffusion and Stochastic Processes. An Analytical Approach* (Springer-Verlag, New York, NY, 2009).

[3] A. F. Cheviakov and M. J. Ward, Math. Comput. Modelling **53**, 1394 (2011).

[4] S. Karlin and H. Taylor, *A Second Course in Stochastic Processes* (Academic Press, 1981).

[5] A. Amitai, I. Kupka, and D. Holcman, Phys. Rev. Lett **109**, 108302 (2012).

[6] D. S. Grebenkov and B.-T. Nguyen, SIAM Review **55**, 601 (2013).

[7] S. A. Isaacson and J. Newby, Phys. Rev. E **88**, 012820 (2013).

[8] J. E. Lennard-Jones, Proc. R. Soc. Lond. A **106**, 463477 (1924).

SUPPLEMENTARY FIGURES

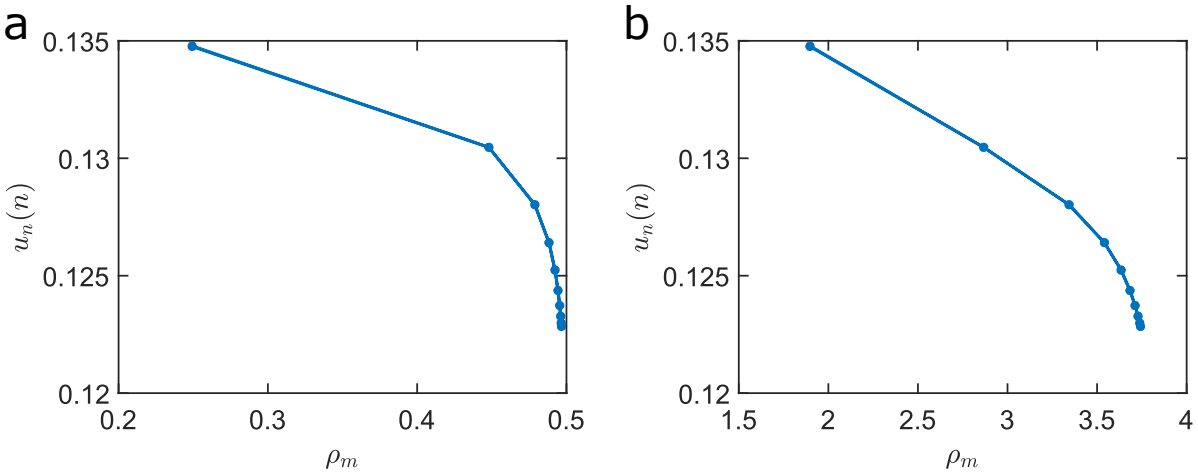

FIG. 1. **The reattachment probability depends on local monomer density.** The reattachment probability $u_n(n)$ as a function of the average local monomer density ρ_m estimated using different monomers along the chain (eq.99). The left most point corresponds to the chain extremities where the density is lowest, while the rightmost point corresponds to the middle monomer of the chain where the density is highest. The polymer has 20 monomer sites ($N = 20$), $a = 0.3b$, $\epsilon = 0.03b$, $A = 10b$. The surrounding ball (eq.98) is of radius $h = 1b$ (a) or $h = 3b$ (b).

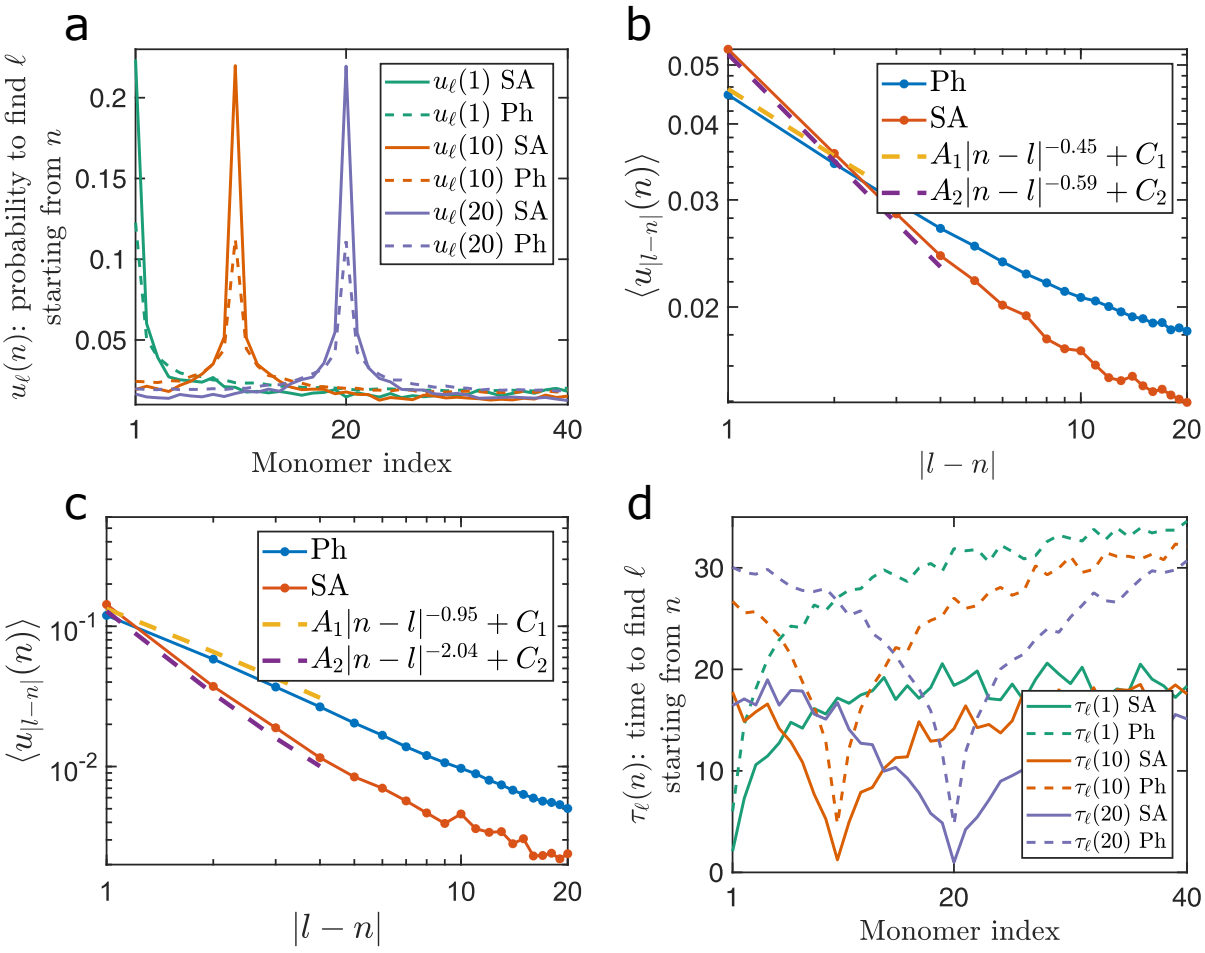

FIG. 2. **Transition between monomer sites.** (**a**) The transition probability $u_l(n)$ starting from site 1 (turquoise), 10 (orange) or 20 (purple). The polymer has 40 monomer sites ($N = 40$), $a = 0.3b$, $\epsilon = 0.07b$, $A = 10b$. The full lines are the result of Brownian simulations for a self avoiding polymer, where the LJ distance was taken to be $\sigma = 1b$ (eq.101). The dashed lines are the simulation results for a phantom polymer (without the LJ potential- see main text). (**b**) The transition probability averaged over all monomers $\langle u_{|l-n|}(n)\rangle$ for a phantom polymer (blue full curve) and a self avoiding polymer (red full curve). Also shown is the result of fitting a function of the form $A|l-n|^{-\alpha}+C$ to the two curves. (**c**) $u_l(n)$ was estimated from simulations with $\epsilon = 0.5b$ and $a = 0.7b$. The same analysis as in b was performed. (**d**) The mean first transition time $\tau_l(n)$ from site n to site l without interacting with any other site along the way. The full line lines are the result of BS for a self avoiding polymer and the dashed line is are the result for a phantom polymer. The results shown are for the case simulated in a,b.

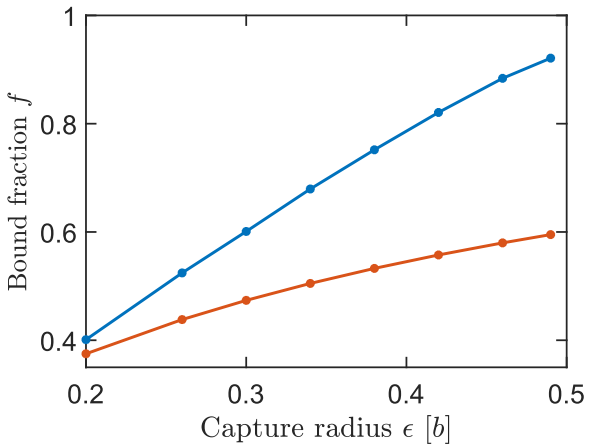

FIG. 3. **The bound fraction of particles to monomer site.** The fraction of bound particles f at monomer sites as a function of the capture radius ϵ. f was estimated from the long time behavior of the probability distribution function $p(\boldsymbol{S}, t)$, by solving the corresponding CTMC (described in the main text) (blue curve). In the Brownian simulations: $N = 100$, $a = 0.5b$, $A = 10b$. The time units are $10b^2/D$, with D the diffusion coefficient of the particle. Also shown is the bulk bound fraction estimate (red curve) given by $f_{bulk} = \frac{k_{on}}{k_{on}+k_{off}}$, with $k_{on} = \frac{4\pi\epsilon ND}{|\Omega|}$. This estimated assumes that the recapture probability and rate are uniform to all sites independent where the particle was released.