

Recognition in Confinement: The Dynamics of Homologous Gene Pairing

Ehud Haimov^{1,2}, Jonathan G. Hedley³, Natalia E. Simonowicz¹, Yaxuan Xiao¹, Andrew Stannard¹, Gleb Oshanin⁴, Angelo Rosa⁵, Yuval Elani⁶, and Alexei A. Kornyshev^{1,*}

¹*Department of Chemistry, Faculty of Natural Sciences, Imperial College London, Molecular Sciences Research Hub, White City Campus, Wood Lane, W12 0BZ, United Kingdom*

²*School of Mechanical Engineering, Tel Aviv University, Tel Aviv 69978, Israel*

³*Department of Engineering Science, University of Oxford, Oxford, OX1 3QD, United Kingdom*

⁴*Sorbonne Université, CNRS, Laboratoire de Physique Théorique de la Matière Condensée (UMR 7600), 4 Place Jussieu, 75252 Paris, Cedex 05, France*

⁵*Scuola Internazionale Superiore di Studi Avanzati (SISSA/ISAS), Via Bonomea 265, 34136 Trieste, Italy*

⁶*Department of Chemical Engineering, Faculty of Engineering, South Kensington Campus, Exhibition Road, London, SW7 2AZ, United Kingdom*



(Received 5 August 2025; accepted 14 July 2026; published 2 September 2026)

How can protein-independent, side-by-side alignment of homologous double-stranded DNA (dsDNA) arise as an early step in genome repair and genetic exchange, and what determines their kinetics under confinement? Although recombination proteins are known to facilitate this process, experiments indicate that an initial side-by-side arrangement of homologs can occur even in their absence. Helical Coherence Theory (HCT) proposes that sequence-dependent distortions of the double helix lead to commensurate charge patterns between interacting dsDNAs, favoring homologous over nonhomologous alignment. However, HCT has largely been developed for straight, rigid dsDNA rods, an approximation limited to roughly one persistence length ($\ell_B \approx 50$ nm), leaving the question open as to how recognition manifests in longer, fluctuating chains. Moreover, the dynamics and timescales of homolog searching under confinement, and the internal dynamics of the paired state, such as transient formation of local unpaired regions (“bubbles”) and end fraying, remain unexplored. Here we extend HCT to dsDNA chains of multiple persistence lengths under confinement using coarse-grained simulations with a custom HCT-based force field to promote our understanding of how homologous genes pair *in vivo*. We simulate two homologous chains in spherical cavities spanning 0.3 to 1 times the free-space gyration radius under two limiting ionic conditions, and use the ratio of confinement radius to free-space gyration radius as a control parameter. Pairing occurs on microsecond timescales, with intermediate confinement giving the most favorable balance between rapid encounter and avoidance of kinetically trapped states. This behavior reflects a crossover from diffusion-limited search under weak confinement to trapping-dominated dynamics under strong confinement. To resolve pairing dynamics, we develop and fit a kinetic theory to simulation trajectories, distinguishing rates of independent pairing modes underlying bubble formation and fraying. Together, these results establish a general physical principle: confinement acts as a control parameter that governs a crossover between diffusion-limited and kinetically trapped pairing regimes for homologous DNA chains, contributing to a minimal protein-free mechanism for homologous recognition on biologically relevant timescales.

DOI: [10.1103/n5cv-svz2](https://doi.org/10.1103/n5cv-svz2)

I. INTRODUCTION

The exchange of homologous genes in genetic recombination is a key process in meiosis and DNA repair. It speeds up evolution, enhances genetic diversity and underpins the transfer of genetic material between different bacteria and

virus species through horizontal gene transfer [1,2]. While the protein-mediated stages of recombination are well understood [3], the initial step, i.e., how homologous sequences identify one another to enable precise recombination at corresponding loci, remains unresolved, even in light of recent reports proposing a potential role for cohesin in the early stages of the recombination process during DNA repair [4,5]. Intriguingly, experimental evidence shows that homolog recognition between genomic regions with high sequence similarity can occur even in the absence of proteins *in vitro*, in simple physiological solutions [6–11].

While the origin of early homology recognition remains unresolved [12–14], several theoretical models support the possibility of protein-independent sequence recognition [15–24]. Recent experimental work [11] has substantiated

*Contact author: a.kornyshev@imperial.ac.uk

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

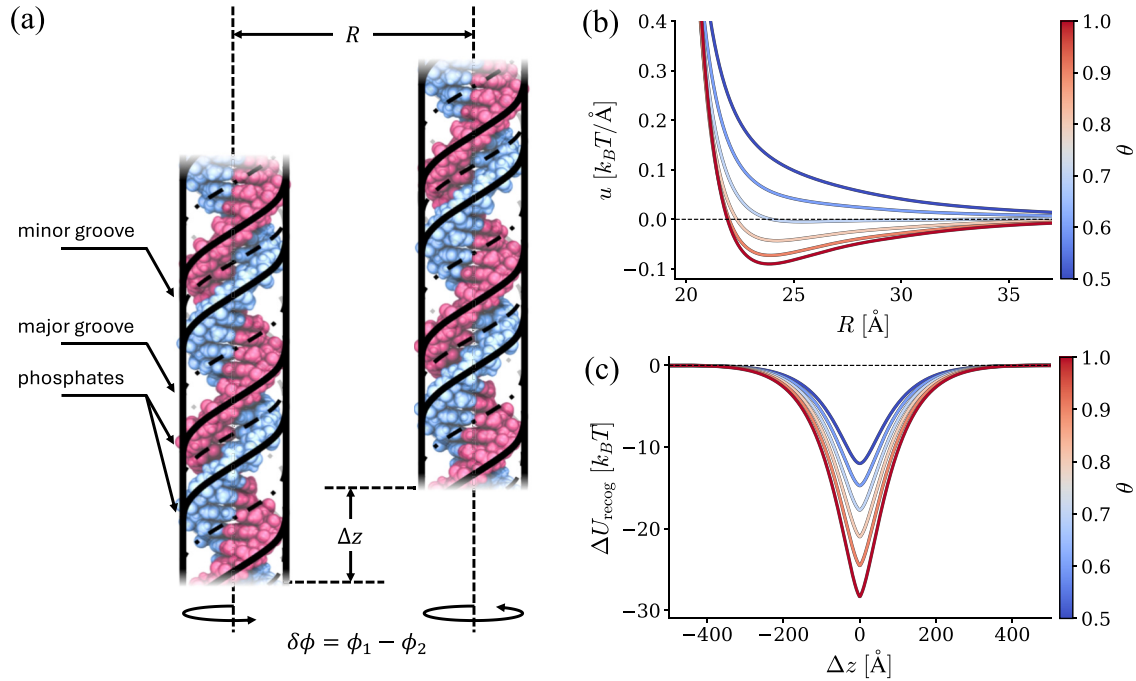


FIG. 1. Helical Coherence Theory. (a) Schematic of two parallel DNA duplexes at an interaxial separation R , relative azimuthal orientation $\delta\phi = \phi_1 - \phi_2$, and axial shift Δz ; the helical charge patterns in the theory arise from the negatively charged double-helical phosphate backbone, and the positively charged lines of condensed counterions in the minor and major grooves. (b) Electrostatic interaction energy per unit length, u [Eq. (1)], minimized w.r.t. $\delta\phi$. Negative energy indicates attraction; curves illustrate the progressive transition from repulsion to attraction with increasing charge compensation θ . (c) Recognition energy well, ΔU_{recog} [Eq. (2)] for a juxtaposition window of length $L = 1000$ Å, at a fixed interaxial separation of $R = 30$ Å, minimized w.r.t. $\delta\phi$. Here, we used the values: $f_1 = 0.2$, $f_2 = 0.8$, which are the fractions of counterions condensed in the minor and the major grooves respectively, $\lambda_c = 100$ Å is the helical coherence length (explained in the text), and the Debye length is $\kappa^{-1} = 7$ Å, the DNA radius $a = 10$ Å, the helical pitch of DNA, $H = 34$ Å, and the dielectric constants of water and the DNA core are $\epsilon = 80$ and $\epsilon_c = 2$ respectively.

one such mechanism based purely on the innate physical properties of DNA [18], supporting the view that homology recognition is an inherent, and potentially universal, feature of double-helical DNA [25]. This physical recognition mechanism, dubbed Helical Coherence Theory (HCT), rests on the well-established fact that DNA is not an ideal double-helix; rather, its local structure exhibits sequence-dependent distortions of structural characteristics such as helical twist and rise [26–31]. When two rigid double-stranded DNA molecules with identical or nearly identical (homologous) sequences are brought into a one-to-one parallel alignment, they can rotate about their long axes to adopt a mutual azimuthal orientation that aligns the negatively charged phosphate backbones of one molecule with the positively charged lines of adsorbed counterions in the grooves of the other [see Fig. 1(a)]. In particular, homologous sequences are considered to be perfectly structurally correlated with one another, and so can maintain this attractive register throughout their length of juxtaposition, thus forming an “electrostatic zipper” [32] between the DNAs. When the degree of phosphate-charge compensation by adsorbed counterions, θ , is sufficiently large (typically above 75%, i.e., $\theta \gtrsim 0.75$), this effect produces net attraction between the molecules at short interaxial distances; at lower charge compensations, it reduces repulsion. Formally, the electrostatic interaction energy per unit length between two such helices at separation R and relative azimuthal angle

$\delta\phi$ is

$$u \approx a_0(R) - a_1(R) \cos(\delta\phi) + a_2(R) \cos(2\delta\phi), \quad (1)$$

where expressions for the helical harmonics, $a_n(R)$, are given in Appendix A 3. Here a_0 describes a purely repulsive term with two contributions: (1) long-range monopole interaction and (2) interaction of helical charges with image charges due to difference in the dielectric constant inside and outside the dsDNA hydrophobic core. a_1 and a_2 are higher harmonics which account for the double helical charge patterns pertaining to the DNA structure. The relative azimuthal orientation that minimizes the energy (the “zipper” orientation) is then given by $\delta\phi_* = \pm \arccos(a_1/4a_2)$. Plotting the minimal interaction energy as a function of axial separation R in Fig. 1(b), it is clear that θ modulates the behavior of the two molecules between repulsion and attraction via the relative contributions of the helical harmonics, with a stable interaxial distance (global minimum energy) forming typically for $\theta \gtrsim 0.75$.

When two rigid and parallel homologous dsDNAs are shifted along their main axis, by some Δz relative to their favorable one-to-one configuration, they gradually lose correlation in sequence-dependent variations in base-pair twist and rise. The loss of registry with Δz , produces a “homology-recognition potential well” in the interaction energy, with half-width λ_c [25], the characteristic shift length over which

correlations are lost, called the *helical coherence length*.¹ This is a universal characteristic of DNA, which was first introduced theoretically [18] and later extracted from classical x-ray fiber diffraction experiments [33], found to be ~ 11 nm [34]. In the case of shifted homologs, Eq. (1) no longer describes the correct DNA-DNA interaction, and the correct formula is given by [25]

$$U(\Delta z) = L\{a_0(R) - a_1(R)v_1(L, \Delta z)\cos(\delta\phi) + a_2(R)v_2(L, \Delta z)\cos(2\delta\phi)\}, \quad (2)$$

where

$$v_n(L, \Delta z) = \begin{cases} \frac{2\lambda_c}{n^2L}(1 - e^{-\frac{n^2|\Delta z|}{\lambda_c}}) + (1 - \frac{2|\Delta z|}{L})e^{-\frac{n^2|\Delta z|}{\lambda_c}}, & |\Delta z| \leq L/2 \\ \frac{2\lambda_c}{n^2L}(1 - e^{-\frac{n^2L}{2\lambda_c}}). & |\Delta z| > L/2, \end{cases} \quad (3)$$

and L is the interaction length along which the two DNA molecules are juxtaposed in parallel and at proximity inter-axial distance R . The relative azimuthal angle $\delta\phi_*$ that minimizes the energy is now given by $\delta\phi = \pm \arccos(a_1v_1/4a_2v_2)$. The interaction energy between shifted homologs has a funnel-shape profile, characterized by half-width length λ_c and a depth set by the sequence-correlation strength and ionic conditions, thus quantifying the crossover from homologous registry at $\Delta z = 0$ to effective heterology for $\Delta z \gg \lambda_c$ [see Fig. 1(c) for the profile of energy difference between those two cases]. At a fixed R , increasing θ amplifies a_1 and a_2 , thus deepening this funneled-shape recognition well.

Several studies suggest that crowding agents can promote homologous pairing. In particular, Baldwin *et al.* [7] studied how homologous dsDNA oligomers pair under low charge neutralization, with gradually increasing osmotic pressure via added PEG. For mixtures of two distinct dsDNA sequence populations, domain segregation was observed despite net electrostatic repulsion between duplexes. Complementary evidence comes from single-molecule magnetic-tweezers experiments, which showed that homologous pairing can occur in simple electrolytes, but is accelerated in the presence of crowding agents [8]. From a theoretical perspective, it has also been shown that homologous sequences aligned in an antiparallel fashion interact no differently from heterologous sequences; in liquid crystalline assemblies, osmotic stress can drive a transition from antiparallel to parallel alignment, thereby restoring recognition [35]. Together, these observations indicate that crowding-induced compression, and the associated confinement of available configurations, can play a

central role in homologous pairing. In cellular environments, analogous compressive effects arise naturally from macromolecular crowding and spatial confinement.

Long homologous dsDNAs, spanning many persistence lengths, have been shown experimentally to form paired structures under physiological salt conditions [8,9]. These observations were further substantiated by a recent theoretical study based on HCT, which predicted the spontaneous formation of partially paired homologous structures, with localized “bubbles” of unpaired dsDNA between paired regions [40].

Despite the *in vitro* evidence of homologous gene pairing, the mechanism of this process *in vivo* is still debated. One potentially significant factor, that has not yet been explored, is confinement, which in cells can be mediated by membranes or biomolecular condensates, or both. It induces a compressive effect on long dsDNA, altering its motion and interactions.

Beyond the question of how protein-free recognition is possible in longer chains, a deeper conceptual gap exists: it remains unknown whether confinement, which is a ubiquitous feature of cellular environments, acts merely as a passive geometric constraint, or whether it qualitatively reshapes the kinetic pathways through which homologous sequences find one another. This distinction matters because it determines whether protein-free pairing can proceed on biologically relevant timescales inside cells.

To elucidate the mechanism of homologous pairing in living system, here, we systematically investigate two central aspects of homologous pairing of long dsDNA under confinement:

- (i) The equilibrium and quasi-equilibrium statistics of pairing, in which homologous molecules adopt paired configurations that are stable over long timescales
- (ii) The dynamics of the pairing process.

This work makes three conceptual contributions. First, it extends HCT from the rigid-rod limit to long, thermally fluctuating chains, connecting microscopic sequence-dependent electrostatics to mesoscopic polymer dynamics. Second, it identifies confinement as a control parameter which affects the competition between the enhancement of the encounter rate of homologous sections, and kinetic trapping. Third, it demonstrates that a cooperative kinetic model quantitatively captures the dominant pairing modes observed in simulation, revealing how each rate is modulated by confinement and ionic conditions. Together these establish a general physical framework for the dynamics of pairing between homologous DNA chains which will pave the way for future work on first passage time for homologous structures to form in the first place.

II. MODELING APPROACH

A. Coarse-grained description

To simulate two homologous DNA chains confined within a sphere, each double-stranded DNA was coarse-grain modeled as a chain of spherical beads of diameter $\ell_0 = 2$ nm, the value of which is chosen to reflect the realistic width of the molecule. Our coarse-grained model is described in Fig. 2. To capture the mechanical extensibility and large-scale conformational properties of the polymer, we implemented a set of standard mechanical interactions between

¹The concept of recognition energy, “the difference in the interaction energy between heterologous sequences and between homologous ones in favourable register,” was first introduced in Ref. [18] and subsequently elaborated on in recent follow-up studies [36,37]. As mentioned, the interaction itself is primarily electrostatic, akin to the electrostatic zipper effect [32], with further possible contributions of hydration forces and water structure [38,39], and it decays exponentially with increasing distance between the molecules, the attractive component of this interaction has a non-trivial decay length that depends on the helical pitch [38].

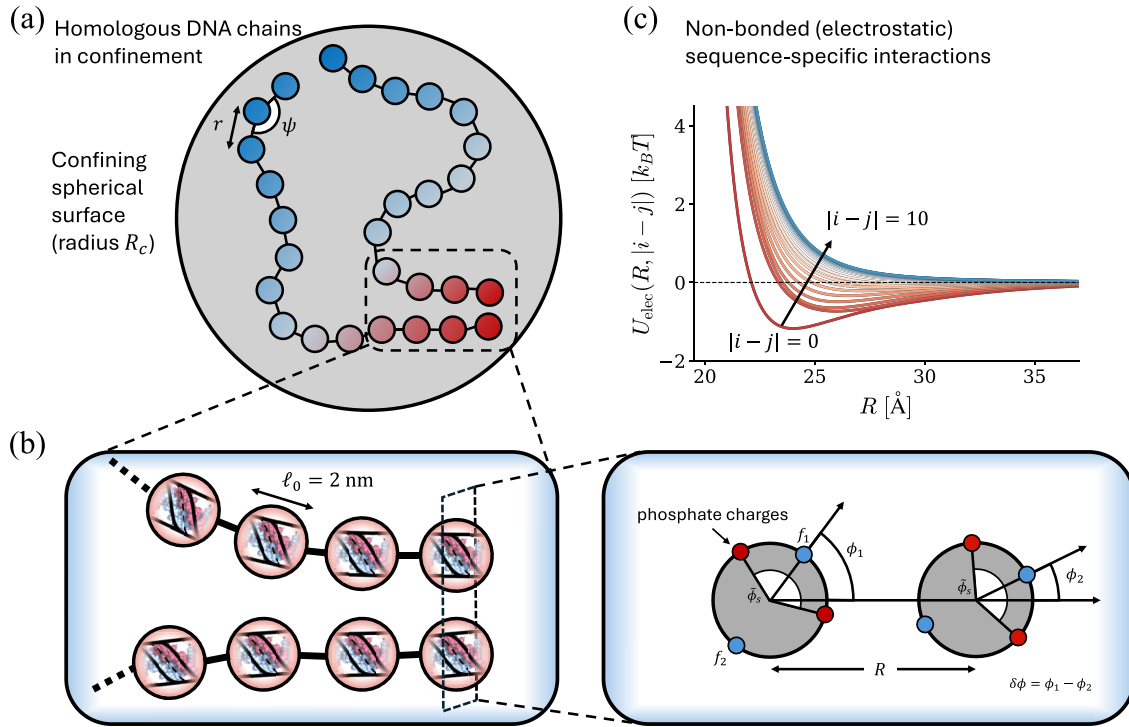


FIG. 2. Coarse-grained model of confined DNA. (a) Two homologous DNA molecules are enclosed inside a spherical confinement of radius R_c . Each molecule is represented as a series of beads with equilibrium bond length r and bond angle ψ ; bead colors grade from blue to red to denote sequence position, so matching colors identify homologous base-pair registers. (b) Enlarged view of the coarse-grained representation. Each bead corresponds to $\ell_0 = 2$ nm (~ 6 bp) of double-stranded DNA. Nonbonded interactions are computed from an analytical electrostatic model in which helical lines of phosphate charges (red) and partially neutralizing counterions (blue) populate the major (f_2) and minor (f_1) grooves. The pair potential depends on the center-center distance R and the azimuthal offset $\delta\phi$; the latter is minimized for every pair to obtain the effective interaction. (c) Electrostatic pair potential $U_{\text{elec}}(R, |i - j|)$ for selected sequence offsets $|i - j|$, computed from Eq. (4). Perfectly aligned homologous segments ($|i - j| = 0$, red) experience a short range attraction, whereas nonhomologous pairs ($|i - j| \gtrsim 5$, blue) are purely repulsive. Intermediate offsets yield progressively weaker attraction, illustrating how sequence specificity is encoded in the interaction.

neighboring beads, calibrated to reproduce the polymer's elastic response and gyration radius (see Appendix A2 for details).

To keep track of similar (homologous) sequences along the chain and to model their interactions accordingly, we indexed the beads of each DNA molecule. For two overall homologous chains, the electrostatic interaction between bead “ i ” in the first chain and bead “ j ” in the second chain was modeled to reproduce the results of HCT [25] [Eq. (2)], in the limiting case of large interaction length $L \gg \lambda_c$ (see Appendix A3 for justification):

$$U_{\text{elec}}(R, |i - j|) = \ell_0 \left\{ a_0(R) - a_1(R) \cos(\delta\phi) e^{-|i-j|\frac{\ell_0}{\lambda_c}} + a_2(R) \cos(2\delta\phi) e^{-4|i-j|\frac{\ell_0}{\lambda_c}} \right\}, \quad (4)$$

where R is the center-to-center separation between beads, $\delta\phi \equiv \phi_i - \phi_j$ where ϕ_i and ϕ_j are the angles stretched between the interaxial line of the sections represented by beads i and j and the middle of their minor grooves. The functions $a_{0,1,2}(R)$ are the first few dominant harmonics of the electrostatic interaction between two parallel homologous DNA in register, depending on the helical charge patterns formed by the phosphate strands and condensed counterions. These manifest in the set of parameters, $\{\bar{\sigma}, \theta, f_1, f_2\}$, where $\bar{\sigma} \approx 16.3 \mu\text{C cm}^{-2}$ is the average DNA charge density, θ is the frac-

tion of DNA charge compensated by condensed counterions, and f_1, f_2 are the fractions of these counterions condensed in the minor and the major grooves respectively (see Appendix A3 for explicit expressions). The first and second harmonic coefficients are also modulated by

(i) Cosine functions of a multiple of the difference in the local angular azimuthal orientations of the DNA segments represented by the beads and

(ii) Exponentially decaying functions, which account for the effects of an interaxial shift from a 1-to-1 register of the two homologous sections.

The modeling of homologous interactions as presented in Eq. (4) was enabled by the fact that the electrostatic helical coefficients $a_{0,1,2}$ decay at least as fast as Debye length (~ 7 Å in cytoplasm)—actually faster, with a decay length related also to the helical pitch [19,20,32], whereas the homology prefactors decay exponentially with a characteristic length scale which is at least not smaller than $\lambda_c/4 \approx 2.5$ nm. The two length scales are separated by almost an order of magnitude and can be thus considered independently. In Eq. (4), for each pair of beads, we determined $\delta\phi$ by adopting an adiabatic approximation in which torsional degrees of freedom about the DNA long-axis relax instantaneously to their minimum energy configuration. Therefore, $\delta\phi$ is also a function of distance R between the beads (see Appendix A3).

The DNA molecules were encapsulated within a confining spherical surface of radius R_c . The interaction between a bead and the confining wall at a surface-to-surface distance x was taken as steric or hydration type force of the form

$$F_w(x) = 2\pi\ell_0\gamma\Lambda e^{-x/\Lambda}, \quad (5)$$

where $\gamma \approx 10^8 \text{ Jm}^{-3}$ [41] is the characteristic repulsion energy per unit volume and $\Lambda \approx 0.3 \text{ nm}$ is the diameter of a water molecule. The relevant geometries for this interaction were accounted for in Eq. (5) by applying a Derjaguin approximation [41].

The displacement of each bead, over each time step, was governed by Langevin dynamics which includes a viscous drag force and a thermal-noise random force, as well as non-random forces (mechanical and electrostatic). The viscous drag force was evaluated using Stokes' law corresponding to the bead's size and the viscosity of the solvent. The self-correlation magnitude of the random force was evaluated in a standard way using the fluctuation-dissipation theorem to ensure realistic diffusive behavior (see Appendix A3 for technical details).

Coarse-grained simulations of entire bacterial genomes are computationally challenging, as typical genomes span $\sim 5 \text{ Mbp}$, confined within micron-scale volumes. Any feasible simulation must then scale down the system. As our primary goal is to quantify the effects of geometric confinement, we identify the relevant dimensionless constant as R_c/R_g , the ratio of confinement radius to the polymer's unconfined gyration radius. However, when considering kinetics, density of the system cannot be ignored, which is not an independent parameter. The polymer volume to confinement volume ratio, scales as $V_{\text{DNA}}/V_c \propto R_g^2/R_c^3$.

To probe the role of geometric confinement, we simulated 2.4 kbp chains, confined within sphere of radius $R_c = 0.3 - 1$ times R_g , where $R_g \approx 118.3 \text{ nm}$, which correspond to DNA volume fraction $V_{\text{DNA}}/V_c = 0.05\% - 1.8\%$. Considering that simple bacteria show $V_{\text{DNA}}/V_c \approx 1\%$, this choice of system size allows us to systematically vary the confinement radius, while avoiding biologically unrealistic high densities.

B. Order parameter for tracking sequence pairing

To quantify the cumulative behavior of pairing of homologous beads, we defined the following order parameter, at a given time step t :

$$S(t) = \frac{|x(t) - y(t)|}{x(t) + y(t)}. \quad (6)$$

Here $x(t)$ is the average distance between all "matching" beads, and $y(t)$ is the average distance between all nonmatching beads. For each timestep, we calculate the distance tensor $D_{ij}(t)$, the components of which indicate the center-to-center distance between bead i in one DNA molecule, and bead j in the other DNA molecule. We can write $x(t)$ and $y(t)$ explicitly in terms of $D_{ij}(t)$ as

$$x(t) = \frac{1}{N} \sum_{i=1}^N D_{ii}(t),$$

$$y(t) = \frac{1}{N^2 - N} \sum_{i=1}^N \sum_{j=1}^N (1 - \delta_{ij}) D_{ij}(t), \quad (7)$$

where δ_{ij} is the Kronecker delta. In the case of ideal, complete pairing of two overall homologous DNA, the distance between matching beads would be much less than the distance between nonmatching beads. For such an extreme situation, S would have continuously evolved and eventually reached ~ 1 . In contrast, in the case of two overall heterologous molecules, S would remain at ~ 0 . Recent theoretical studies suggest, however, that pairing between long homologous sequences can never be fully complete due to thermal fluctuations, which give rise to persistent "bubbles" [40], i.e., unpaired regions distributed along the contour length, in agreement with earlier experimental observations [9]. We expect to see this in simulation results.

C. Counterion condensation

To examine how counterion condensation affects homologous pairing, we vary the parameter θ in our simulations, the fraction of phosphate charge neutralized by adsorbed cations. For a single type of DNA-adsorbing cations, this parameter can be approximated with a Langmuir isotherm [11]:

$$\theta(c) = \frac{c}{K_d + c}, \quad (8)$$

where c is the bulk concentration of the specifically-adsorbing cation in the simulation box and K_d is the dissociation constant for binding to DNA. We consider two cases: (1) an idealized limit of full charge neutralization, $\theta = 1$, corresponding to $c \gg K_d$ and (2) one that more realistically mimics conditions in bacteria, in which only a fraction of the DNA charge is compensated. To model this case, we use $\theta = 0.85$. Assuming that magnesium (the most abundant cation that condenses on DNA in prokaryotes) is the only species that significantly adsorbs on DNA, using the Mg^{2+} dissociation constant $K_d = 16.3 \text{ mM}$ for approximately physiological monovalent ion concentration (100 mM Na^+) reported by Ref. [11], we can relate $\theta = 0.85$ to Mg^{2+} concentration of approximately 92 mM . This is consistent with the approximate total intracellular Mg^{2+} concentration in simple bacteria [42]. However, most Mg^{2+} is bound to phosphate-rich biomolecules, resulting in a free ion concentration on the order of mM . However, the dynamic nature of Mg^{2+} association, along with the presence of the polyamine spermidine, which binds strongly to DNA, makes $\theta = 0.85$ a reasonable approximation.

III. RESULTS

A. Steady-state properties of paired homologous genes

Simulations were initialized with the 25 central beads of the two polymers placed in one-to-one homologous register. The remaining beads were then situated sequentially by a self-avoiding random walk, rejecting trial positions that would place a bead outside the confining sphere or cause overlap with an existing bead. This configuration was chosen to mimic a state in which an initial homologous contact had already formed, so that the simulations could focus on the subsequent dynamics of pairing rather than on the first encounter itself. Each trajectory was then evolved for $\sim 500 \mu\text{s}$, which was typically sufficient for the system to reach a steady state.

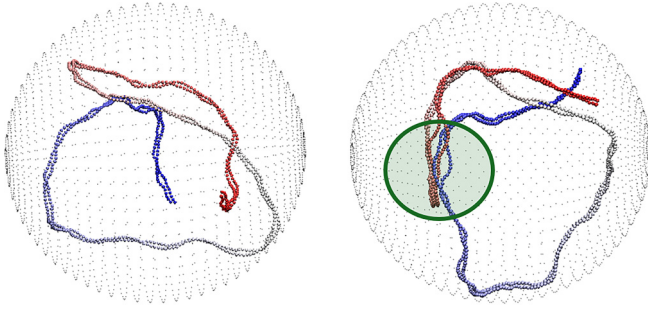


FIG. 3. Pairing of two homologous DNA molecules: Two types of steady-states. Panels display the snapshots of simulation outcomes for two overall homologous DNA molecules. Each bead number (from 3' to 5' ends) is represented with a specific color to distinguish different sections of homologous sequences from each other; each DNA duplex is painted by a color gradient from blue to red along its contour. Parameters used for both simulation runs were the same: $R_c/R_g = 0.7$, $\theta = 0.85$. The faint gray dots show the confining sphere. Outcomes: (a) Type 1: no kinetic traps, all beads are engaged in a periodic steady-state "breathing" from paired to unpaired local configurations and back. See movie for temporal evolution from initial state at [43]. (b) Type 2: system is kinetically stuck due to protruding duplex going between two paired regions. The highlighted green circle shows a problematic, stuck region where a duplex entered a looped section. See movie for temporal evolution from initial state at [44].

At the end of each simulation, one of two characteristic steady-state outcomes were observed. In the first, the polymers relaxed to an equilibrium configuration of homologous sections in one-to-one register, as shown in Fig. 3(a). This state was not static: paired regions continually formed and dissociated, producing a fluctuating pattern of paired segments and local unpaired bubbles along the duplexes. In the second, the system became kinetically trapped, as shown in Fig. 3(b). In these configurations, a protruding duplex or paired fragment became lodged between two already paired regions, obstructing relaxation to the fully ordered register.

Kinetic trapping was most common under strong confinement, $R_c/R_g \lesssim 0.4$. In this regime, confinement increases the frequency with which homologous beads encounter one another, but it also raises the probability that pairing is initiated in a nonoptimal geometry, which promotes the formation of protrusions. The detailed geometry of these trapped configurations varied across trajectories, but tighter confinements generally produced both more frequent and larger protrusions, with the precise behavior also depending on the ionic conditions. At intermediate confinements, $R_c/R_g \approx 0.5$ – 0.7 , nonoptimal encounter configurations were still sufficiently common for trapped states to arise transiently, but the duplexes retained enough configurational freedom to unpair, rearrange, and recover the correct one-to-one register. As a result, the ordered outcome could be reached either directly or after first passing through a trapped state.

To quantify these steady-state behaviors, we used the order parameter $S(t)$ and report its time-averaged value over the first steady-state plateau, $\langle S \rangle_{SS}^{(1)}$, for fully homologous chains across different confinement volumes and values of the intrinsic charge compensation θ . When multiple steady-state

plateaus were identified within a trajectory, separated by intervals of non-stationary behaviour, we defined $\langle S \rangle_{SS}^{(1)}$ using the earliest plateau. The method by which we determine plateaus and hence the steady-state order parameter is described in Appendix B.

As confinement becomes stronger, $\langle S \rangle_{SS}^{(1)}$ decreases, with a more pronounced reduction for $\theta = 0.85$ than for $\theta = 1$. This trend reflects a progressive shift from efficiently paired dynamics towards kinetically trapped configurations: tighter confinement increasingly favors nonoptimal encounters while limiting the configurational freedom needed to recover the fully paired state. This stronger reduction in $\langle S \rangle_{SS}^{(1)}$ observed at lower θ further indicates that when charge compensation, and hence pairing strength, is reduced, trapping emerges already under weaker confinement.

For nonstuck system configurations, the distribution of distances in steady-state between homologous beads was recorded; see Fig. 5 (left) for the histograms. This distribution was used to analyze meaningful statistics of the system's steady-state properties, within a statistical physics framework. The height of each bin, approximates the probability $p(R)$ of finding a two homologous beads at a distance R , and can be expressed as

$$p(R) = \Omega(R) \frac{e^{-U(R)/k_B T}}{Z}, \quad (9)$$

where Z is the partition function, $\Omega(R)$ is the multiplicity function (number of degenerate states) and $U(R)$ is the effective energy between each two homologous beads which accounts for not only the electrostatic nonbonding interactions, but also the bonding (force-field) contributions between the beads alongside any effects from the confinement. The multiplicity function $\Omega(R) = 4\pi R^2 \Delta R / \Lambda_T^3$, i.e., the ratio of the spherical shell volume associated with the specific bin position R and the cubed de Broglie thermal wavelength, $\Lambda_T^3 = h / \sqrt{2\pi m k_B T}$, where ΔR is the width of the bin, h is the Planck constant and m is the mass of a bead.

The effective energy $U(R)$ between two homologous beads is given by

$$\frac{U(R)}{k_B T} = -\ln \left(\frac{n(R)}{\Omega(R)} \right) + \text{const.} \quad (10)$$

To effectively remove the constant from Eq. (10) we simply set $U(R \rightarrow \infty) \rightarrow 0$, i.e., we vertically shift the entire curve such that the tail ends at zero. Results for $U(R)$ for different (nonstuck) cases are shown in Fig. 5 (right). Comparing the $U(R)$ in Fig. 5 with the result in Fig. 2, reveals how the collective behavior of the polymer beads impacts the interactions between two homologous beads. The well is wider and shallower for $\theta = 0.85$, reflecting how smaller charge compensation results in less tightly bound states.

B. Kinetics of homologous pairing: Evolution to steady state

The characteristic pairing time, τ , is defined as the time required to reach 70% of the first steady state [see Figs. 4(a) and 4(b)], reflecting the timescale over which a substantial fraction of homologous segments become aligned, allowing for homologous recombination (i.e., the exchange of genes) to locally start. Note that τ is defined relative to each system's

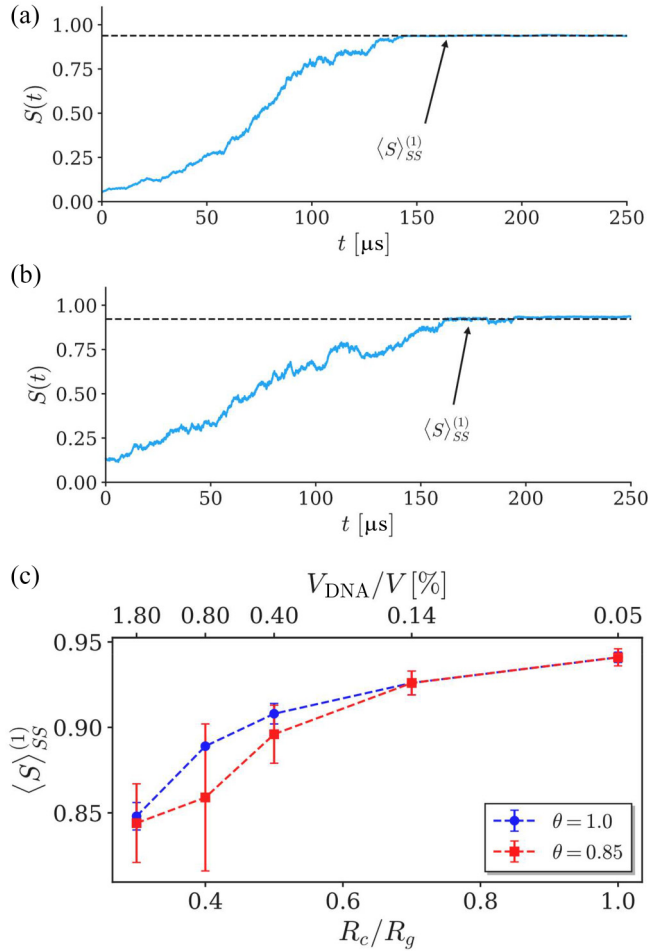


FIG. 4. Time evolution of the order parameter at the first steady state $\langle S \rangle_{SS}^{(1)}$; its dependence on confinement and fraction of charge compensated. (a) Time behavior of the order parameter tracks pairing throughout the simulation. For the case of $R_c/R_g = 1.0$ and $\theta = 1.0$, order parameter converges to an equilibrium value, as a perfectly paired state is formed. (b) For the case of $R_c/R_g = 0.7$ and $\theta = 0.85$, order parameter reaches a first steady value, corresponding to kinetically stuck state, and then goes up to reach another steady-state value, as a result of spontaneous unbinding and removing protrusions. (c) The dependence of the steady-state order parameter with confinement size for two charge-compensation levels, $\theta = 0.85$ (red) and $\theta = 1.0$ (blue).

own first steady-state plateau. Since $\langle S \rangle_{SS}^{(1)}$ itself varies with confinement, comparisons of τ across confinements should be interpreted as comparisons of the time required to reach the first steady-state pairing regime of each system, rather than the time required to reach a common fixed degree of global registration. At this point, the system is not always in equilibrium and sometimes exhibits kinetically trapped configurations with substantial amount of well aligned homologous sections with DNA protrusions in between. The transition from kinetically stuck configurations to a full equilibrium involves additional kinetic processes, such as local pairing and unpairing around the protrusion, which occur on different timescales and are beyond the scope of the present

analysis as here our focus is on the biologically motivated first steady-state configuration.

Combined with data in Fig. 4(c), this behavior represents a crossover between two physically distinct regimes: a diffusion-limited regime at weak confinement, where the dominant bottleneck is the time for homologous segments to encounter one another (with $R_c/R_g = 1$ marking the boundary of this regime, beyond which further increases in confinement leave polymer configurations largely unperturbed, see Appendix D), and a trapping-dominated regime at strong confinement, where protrusions due to multi-site attachment create long-lived metastable states. This crossover is an emergent consequence of the interplay between polymer flexibility, electrostatic recognition, and geometric confinement. We quantify the effect of self-crowding at tight confinements on the anomalous diffusion of the DNA molecules in Appendix E.

The temporal evolution of the distance between two specific homologous beads is shown in Fig. 6(b). Once in the paired state, we see that the beads exhibit steady-state “breathing mode” in which the distance between them fluctuates over time. Thus, while the order parameter gives a rough estimate of the pairing status, it lacks the ability to predict the number of paired homologous beads. Therefore, we define two homologous beads to be paired when they are at a surface-to-surface separation equal to, or lower than, the Debye length, $\lambda_D = 7 \text{ \AA}$ as a length scale for electrostatic interactions. Given this definition, we calculated the time evolution of the fraction of homologous paired beads, $n(t)$ as each simulation progresses, shown in Fig. 7(b) (gray points).

Simple Kinetic Model. To characterize the pairing and unpairing dynamics of beads, we derived a simplistic kinetic theory in which we consider two rates; k_{pair} , the rate of formation of a pair of beads and k_{unpair} , the rate of dissociation of a pair of beads. Constructing the corresponding rate equation for $n(t - t_0)$, fraction of beads paired at times larger than empirically chosen time t_0 , and solving it analytically (see Appendix C 1 for details), we obtain

$$n(t - t_0) = 1 + \frac{\beta}{2} - \frac{1 + \frac{\beta}{2} - n_0 + \gamma \tanh[\gamma k_{\text{pair}}(t - t_0)]}{1 + \frac{\beta + 2 - 2n_0}{2\gamma} \tanh[\gamma k_{\text{pair}}(t - t_0)]}, \quad (11)$$

where $\beta = k_{\text{unpair}}/k_{\text{pair}}$ and $\gamma = \frac{1}{2}\sqrt{\beta^2 + 4\beta}$. This simple model is expected to provide a reasonable description of the system near equilibrium, but less-accurate description in the stages where the system is far from equilibrium. Fitting Eq. (11) to the simulation data [dashed lines in Fig. 7(b)], this simple model gives qualitatively reasonable fits as pairing approaches equilibrium [dashed curves in Fig. 7(b)], but fails to capture early time pairing behavior as it doesn’t consider the correlations between the beads within each chain.

Detailed Master Equation. The simple kinetic model above only considered the evolution of the global paired fraction $n(t)$, and thus neglects correlations along the chain. In our simulations, the pairing and unpairing of beads is locally cooperative: beads within a paired domain are stabilized by their neighbors, whereas beads at domain boundaries and at the ends of the chain are more susceptible to unpairing, leading for example to fraying at the end of the paired dsDNA

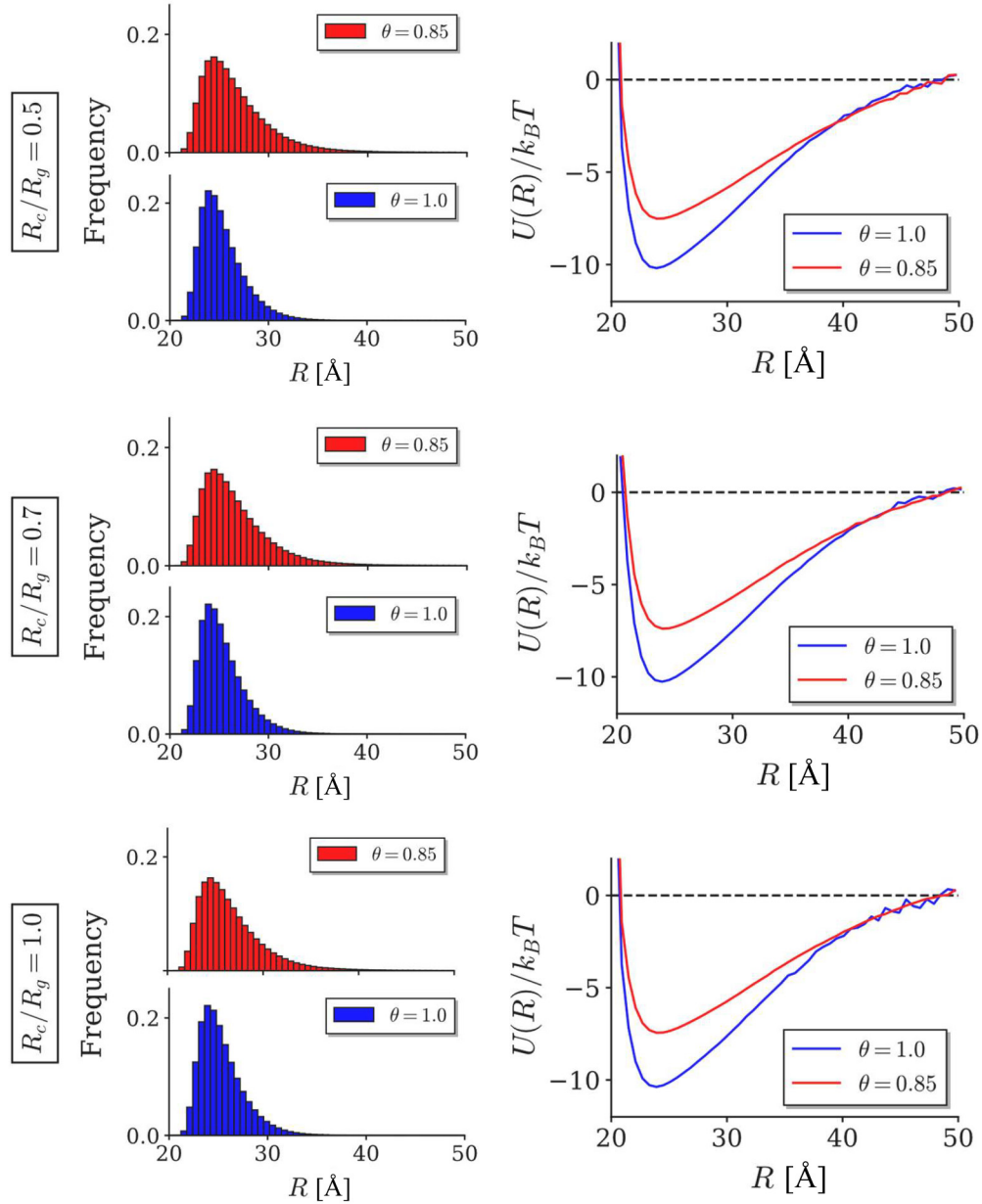


FIG. 5. Distributions of distances between homologous beads and the resulting effective free-energy profiles. Left column: histograms representing distribution of distances between homologous beads recorded in steady state are shown for three confinement radii, $R_c/R_g = 0.5$, 0.7 , and 1 , and indicated values of intrinsic charge compensation parameter θ . Statistics were gathered only for cases which achieved perfect pairing. Right column: Effective interaction potentials, $U(R)$ for two charge-compensation levels, $\theta = 1.0$ (blue) and $\theta = 0.85$ (red), computed from their corresponding histograms by Eq. (10).

chains. Collapsing these distinct behaviors into a single pair of effective rates, as considered in the simple kinetic model, effectively averages over local environments and therefore cannot reproduce the detailed shape of $n(t)$ at early stages of pairing.

To capture this, we introduce a nearest-neighbor master equation for the time-dependent probability that a bead i along the chain is paired, $P_i(t)$. The rate of change of P_i depends on whether its neighbors ($i-1$ and $i+1$) are paired, leading to three local environments: (1) both neighbors paired, (2) exactly one neighbor paired, and (3) both neighbors unpaired [see Fig. 7(a)]. Each environment $s \in \{1, 2, 3\}$ has

its own pair-formation and dissociation rates, $k_{\text{pair}}^{(s)}$ and $k_{\text{unpair}}^{(s)}$, giving

$$\begin{aligned} \frac{dP_i}{dt} = & k_{\text{pair}}^{(1)}(1 - P_i)P_{i-1}P_{i+1} \\ & + k_{\text{pair}}^{(2)}(1 - P_i)[P_{i-1}(1 - P_{i+1}) + P_{i+1}(1 - P_{i-1})] \\ & + k_{\text{pair}}^{(3)}(1 - P_{i-1})(1 - P_{i+1})(1 - P_i) \\ & - k_{\text{unpair}}^{(1)}P_{i-1}P_{i+1}P_i \\ & - k_{\text{unpair}}^{(2)}P_i[P_{i-1}(1 - P_{i+1}) + P_{i+1}(1 - P_{i-1})] \\ & - k_{\text{unpair}}^{(3)}P_i(1 - P_{i-1})(1 - P_{i+1}). \end{aligned} \quad (12)$$

TABLE I. Rate constants for pairing and unpairing in different modes. The rate constants are presented for various cases of confinement and charge compensation and were obtained as fitting parameters from fitting Eq. (12) to curves in Fig. 7(b).

	$\theta = 1.0$		$\theta = 0.85$	
	$R_c/R_g = 0.7$	$R_c/R_g = 1.0$	$R_c/R_g = 0.7$	$R_c/R_g = 1.0$
$k_{\text{pair}}^{(1)} [\text{ms}^{-1}]$	124.3	76.8	13.6	7.5
$k_{\text{unpair}}^{(1)} [\text{ms}^{-1}]$	13.8	8.5	2.1	1.2
$k_{\text{pair}}^{(2)} [\text{ms}^{-1}]$	26.7	23.4	13.3	10.8
$k_{\text{unpair}}^{(2)} [\text{ms}^{-1}]$	12.6	11.0	9.6	7.8
$k_{\text{pair}}^{(3)} [\text{ms}^{-1}]$	15.3	4.1	1.5	0.3
$k_{\text{unpair}}^{(3)} [\text{ms}^{-1}]$	36.8	9.9	7.2	1.4

The first three terms represent pair formation in the three neighbor configurations, while the last three represent dissociation in their corresponding configurations. We solve this system numerically and compare to simulations via the mean paired fraction $n(t) \approx (1/N) \sum_i P_i(t)$. To reduce the number of independent fitting parameters, we impose an approximate detailed-balance constraint within each local environment, relating $k_{\text{unpair}}^{(s)}$ to $k_{\text{pair}}^{(s)}$ via equilibrium pairing probabilities (see Appendix C 2). The resulting fits [solid curves in Fig. 7(b)] capture fine structure in $n(t)$ that is missed by the simple model, consistent with locally cooperative pairing and unpairing along the chain.

The fitted rate constants in Table I exhibit clear trends. Increasing confinement enhances all pairing rates, with the most pronounced effect on $k_{\text{pair}}^{(3)}$, which governs diffusive search and nucleation of new contacts. This increase reflects the reduced available volume at smaller R_c/R_g , which raises the probability of encounter. The rate $k_{\text{pair}}^{(1)}$, associated with the closure of isolated unpaired beads, also grows with tighter confinements. This may be associated with local paired-polymer properties that are themselves affected by global confinement. However, the ratio $k_{\text{pair}}^{(1)}/k_{\text{unpair}}^{(1)}$ remains roughly constant with confinement, suggesting that confinement may also be changing an effective local fluctuation or rearrangement timescale rather than simply altering the thermodynamic landscape of pairing. In contrast, $k_{\text{pair}}^{(2)}$, describing local zipping propagation, depends only weakly on confinement, consistent with its essentially local character. The combination of a nearly unchanged zipping rate and an enhanced nucleation rate explains the faster pairing observed at lower R_c/R_g , as well as the emergence of protrusions and kinetically stuck states.

All rates further increase with the fraction of charge neutralization θ , in agreement with pairing times shown in Fig. 6, indicating that stronger interactions accelerate pairing. Notably, $k_{\text{pair}}^{(3)}$ is the most sensitive to θ , highlighting the strong dependence of contact nucleation on interaction strength.

As the system approaches equilibrium, the pairing probabilities for individual beads, P_i (obtained numerically by solving Eq. (12) with parameters from Table I), converge to an approximately uniform value along the chain [see Fig. 7(c) for $\theta = 0.85$, $R_c/R_g = 0.7$]. In this regime, the Simple Kinetic Model and the Detailed Master Equation coincide [Fig. 7(b)], as the mean-field approximation underlying the simple model becomes a reasonable description. The terminal beads are an exception to this, as they exhibit lower pairing probabilities

due to their increased tendency to fray. However, since these beads account for only a small fraction of the chain, the Simple Kinetic Model still provides a reasonable description near equilibrium.

We emphasize that this kinetic model is phenomenological. The fitted rates are effective parameters obtained by matching the simulated paired-fraction curves, rather than local transition rates independently measured from trajectory statistics. The model is therefore used here as a descriptive tool to interpret the dominant local pairing modes, not as a parameter-free predictive framework. Direct extraction of local transition rates and validation against independent observables such as bubble-size distributions or fraying statistics are left for future work.

IV. DISCUSSION

The central result of this work is the identification of a general physical principle: geometric confinement acts as a control parameter which governs the kinetic pathways of homology pairing in DNA chains. It reveals a crossover between regimes with physically distinct rate-limiting steps. Our results show that confinement—a fundamental physical feature of cellular environments—acts on homologous DNA pairing in two distinct ways. By reducing the available configurational space, it initially accelerates encounters between homologous segments. But stronger confinement does not simply continue to improve pairing: as protrusions and nonoptimal contacts become more common, the increase in encounter rate is increasingly offset by the difficulty of rearranging the duplexes into the correct register. This stronger-confinement regime is also accompanied by increasingly subdiffusive monomer dynamics, consistent with self-crowding and suppressed bending fluctuations [45], as shown in auxiliary simulations presented in Appendix E. Confinement therefore affects not only how quickly homologous segments find one another, but also how efficiently those encounters can be converted into correctly paired states. More broadly, this highlights the importance of spatial organization in genome dynamics [46–48].

The role of charge compensation through θ shows that this confinement effect is not purely geometric. Stronger homologous interactions, corresponding to larger θ , make correctly paired states more robust and increase the likelihood that nonoptimal encounters can reorganize into the proper

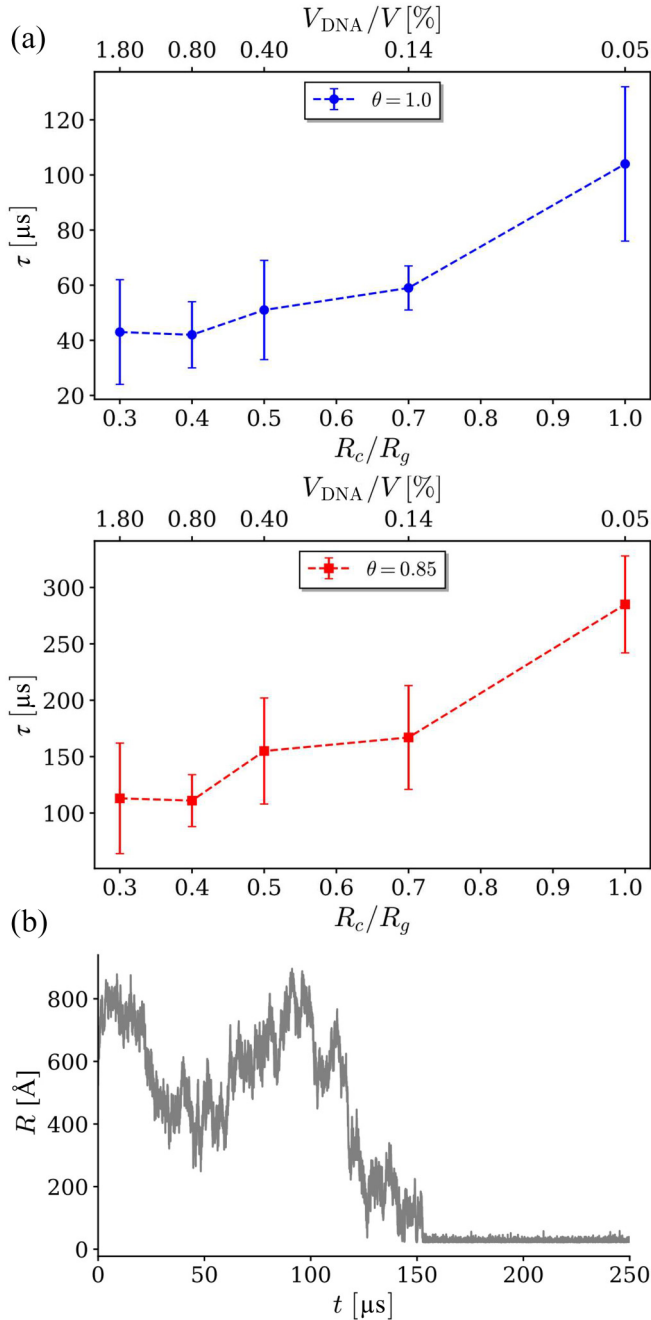


FIG. 6. Pairing kinetics. (a) The homologous pairing time, τ , as a function of confinement, for $\theta = 0.85$ (red) and $\theta = 1.0$ (blue). Points were obtained as a mean over four independent repeats and their error bars display standard deviation for the different repeats. (b) Time evolution of center-to-center separation between randomly selected couple of homologous beads, plotted for $\theta = 0.85$ and confinement $R_c/R_g = 0.7$.

register. When θ is lower, pairing is weaker, and the system becomes more susceptible to kinetically trapped states even under less severe confinement. The stronger reduction in the steady-state order parameter for $\theta = 0.85$, together with the earlier saturation of the confinement-induced speed-up, is consistent with this interpretation. Confinement therefore acts together with interaction strength: it brings homologous

segments into contact, but whether those contacts mature into correctly paired states depends on how strongly pairing is stabilized.

This picture reinforces the idea that homology recognition is an intrinsic physical property of DNA, arising from sequence-dependent structure and electrostatic interactions, and whose manifestation in long, flexible chains is strongly shaped by geometry and polymer dynamics. Our results further suggest that geometric confinement can act as an important control parameter for homologous pairing kinetics, with the dynamics crossing over between two distinct regimes. Under weak confinement, pairing is diffusion-limited, because excessive spatial freedom prolongs the homology search. Under strong confinement, the dynamics become trapping-dominated, because excessive restriction promotes kinetically trapped multisite attachments that impede a more complete pairing structure. The favorable confinement regime identified here reflects the balance between these two effects.

This is consistent with broader arguments that confined or compartmentalized environments can accelerate reactions by reducing the search space [49,50]. It also suggests that the efficiency of pairing is controlled jointly by geometry and electrostatics, a coupling that may be relevant when considering the diversity of cellular and subcellular dimensions observed in nature [51–54]. The emergence of a favorable confinement regime for pairing hints at the possibility that evolutionary pressures on biological size may, in part, reflect the need to support efficient genomic processes. We also note that $S(t)$ is a global, chain averaged order parameter. It quantifies the overall degree of homologous pairing, but does not by itself resolve local misregistration, protrusion length, or the detailed geometry of trapped regions. Developing local structural metrics to classify these trapped-state geometries is an important direction for future work.

Our kinetic modeling helps clarify how this behavior arises. The simple kinetic model captures the late-stage dynamics because, near steady state, the pairing probability becomes approximately uniform along most of the chain. The more detailed Master Equation shows, however, that the earlier stages are locally cooperative: confinement most strongly enhances the formation of new contacts, whereas local zipping is much less sensitive to confinement. This distinction helps explain why stronger confinement can accelerate the onset of pairing without necessarily improving the final paired state. The value of the more detailed model therefore lies not only in fitting the data more closely, but also in showing that the limiting step under confinement is not local zipping itself, but the balance between contact formation and subsequent reorganization.

These results should also be interpreted in light of the simulation protocol. We initialize all simulations with a centrally located paired stretch of 25 beads, chosen to mimic a successful initial homologous encounter that has already produced a stable nucleus. The timescales reported here should therefore be understood as those for growth and reorganization of an already nucleated paired state, rather than the full first-passage search time from completely separated chains. Placing the initial paired region at the center also minimizes asymmetry between the two flanking segments, and hence provides the most favorable geometry for subsequent pairing.

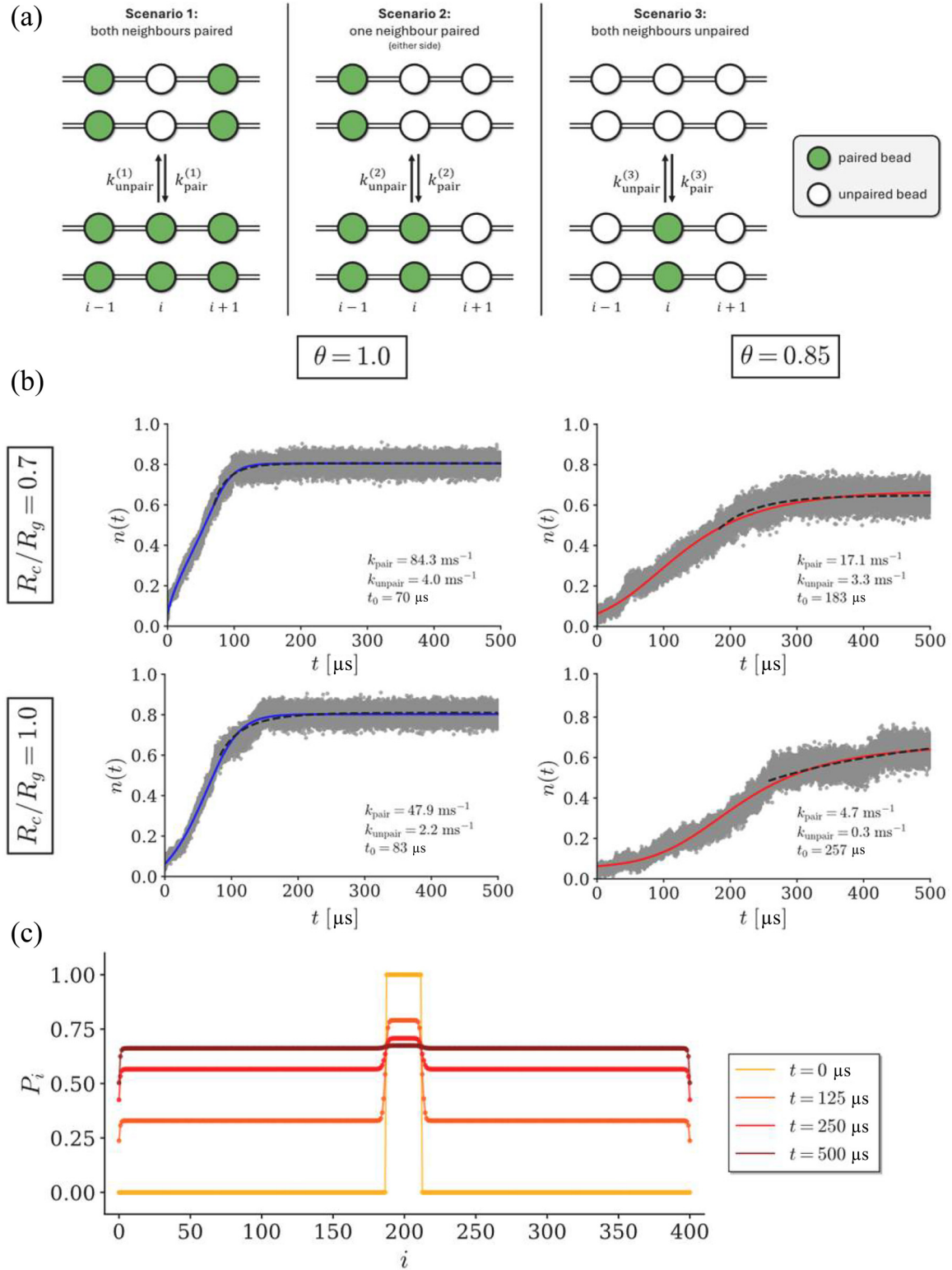


FIG. 7. The kinetics of homologous gene pairing. (a) Schematic of the model for homologous gene pairing used in the Detailed Master Equation, which takes into account correlations with the nearest neighbors. (b) The progression of pairing is shown for various cases of confinement and charge compensation. Raw simulation data are shown in gray. The fits to the Simple Kinetic Model, expressed by Eq. (11), are shown as dashed lines, with the fitting parameters, k_{pair} and k_{unpair} , shown in each panel. The fits to the Detailed Master Equation, expressed by Eq. (12), are shown as blue (for $\theta = 1.0$) and red (for $\theta = 0.85$) solid lines. The fitting parameters are displayed in Table I. (c) P_i as a function of bead index, i . Example plot for $\theta = 0.85$, $R_c/R_g = 0.7$ at four representative times.

An important simplification of our model is the neglect of torsional elasticity and chiral torques. The role of torsional elasticity in homolog pairing has been studied extensively in earlier work [55,56]. These studies show that torsional relaxation of mismatches can reduce the recognition energy, but not eliminate it, because torsional adaptation itself carries an energetic cost. Neglecting these degrees of freedom makes large-scale simulations of confined DNA pairing computationally feasible. In the present context, this approximation is reasonable because we focus on pairing between identical DNA molecules: once homologous segments encounter one another, torsional adjustment should be comparatively small, since mismatches in intrinsic helical structure are largely absent. By contrast, torsional adaptation is expected to matter more for nonhomologous interactions, which are unlikely to yield stable pairing after a homologous nucleus has formed. Likewise, neglecting chiral torques is appropriate for the early pairing regime considered here, where thermal fluctuations generate a substantial number of transient pairing bubbles [8,9,40], likely hindering the formation of stable supercoiled paired structures.

A limitation of the present model is that the two homologous chains are treated as sequence-identical. Real homologs contain sequence heterogeneity, which would make the local HCT recognition energy vary along the contour. However, homologous chromosomes are highly sequence-similar; a typical human genome is approximately 99.6% identical to a reference genome or to another individual genome [57]. Since the recognition energy associated with a single base pair is very small, approximately $-0.02 k_B T$ [11], isolated mismatches are expected to produce only weak local perturbations. In the extremely rare event of clustered mismatches or insertions and deletions in one of the homologous genes, there could be stronger effects, potentially slowing global alignment or promoting kinetic stuck states.

These insights hold significant promise for synthetic biology and artificial cell technologies [58,59]. In bottom-up approaches to synthetic cell construction, the incorporation of geometric confinement—through vesicles, coacervates, or droplet compartments—can be strategically harnessed to direct homologous DNA interactions and recombination events [60]. By embedding physical constraints into cell-free systems, it may be possible to recapitulate complex genome-scale functions without relying on the full complement of cellular protein machinery.

The biological implications of these results should also be stated with care. Cells naturally exploit compartmentalization to create distinct microenvironments that modulate reaction rates and coordinate genomic processes, as seen in DNA-rich structures such as the nucleoid and nucleolus in prokaryotes and eukaryotes [48,61]. *In vivo*, DNA operates in a protein-rich environment in which DNA-binding proteins, active processes, and biomolecular condensates can strongly influence accessibility, mobility, and intermolecular interactions, as well as directly facilitate homologous recombination. Our results therefore do not imply that confinement alone governs homologous pairing in cells. Rather, they show that geometric confinement, even in the absence of proteins, can already bias the pairing pathway and promote homologous association. In this sense, the present study builds on experimental

observations that homology recognition can occur even in protein-free systems [7–9,11], helps clarify the physical mechanisms underlying such protein-independent recognition, and provides a baseline against which the effects of additional biological components can later be assessed. An important direction for future work will be to investigate pairing in the presence of sequence-neutral condensates, modeled at different levels of coarse graining, which may introduce osmotic stress in addition to geometric confinement and thereby further promote or inhibit pairing.

V. CONCLUSIONS

We have shown that geometric confinement can substantially accelerate homologous gene pairing, but that this acceleration does not continue indefinitely as confinement is strengthened. Moderate confinement promotes encounters between homologous segments, whereas stronger confinement increasingly favors nonoptimal encounters and kinetically trapped configurations that hinder relaxation to the correct paired state. Homology search is therefore governed not only by how rapidly homologous segments meet, but also by how readily (misregistered) contacts can reorganize.

This behavior defines a favorable confinement regime for pairing, marking the transition between diffusion-limited pairing and kinetically trapped multinucleation behavior. Geometric confinement therefore emerges as an important physical control parameter for homologous pairing. More broadly, these results provide a baseline for understanding how homologous recognition in confined environments may be further shaped by proteins, condensates, and other cellular factors.

ACKNOWLEDGMENTS

The authors are thankful to Marco di Antonio, Geoff Baldwin, Fernando Bresme, Ming Chen, Claudia Danilowicz, Ralph Holden, Lorenzo di Michele, Mara Prentiss, and John Seddon for useful discussions throughout this work. This study was part of a project supported by the Leverhulme Trust, Grant No. RPG-2022-142. J.G.H. is supported by a Novo Nordisk Postdoctoral Fellowship in partnership with the University of Oxford, and the David Cockayne Junior Research Fellowship at Linacre College, Oxford.

A.A.K. and E.H. conceived the project. E.H., J.G.H., A.R., G.O., and A.A.K. developed methodology, N.E.S., J.G.H., and Y.X. performed simulations and curated the data, E.H., J.G.H., N.E.S., Y.X., A.S., G.O., and A.A.K. analyzed the data, N.E.S., J.G.H., and Y.X. performed their visualization. E.H. and A.A.K. led the writing of the initial draft, with drafting and revision contributions from J.G.H.; E.H., J.G.H., N.E.S., and A.A.K. developed its submitted version, with input from all authors. A.S. provided key conceptual suggestions during revision, informed by ongoing experimental work to test the paper's predictions. All authors participated in review and editing of the paper and approved the final manuscript. A.A.K. supervised the project and, with Y.E., acquired funding.

The authors declare no conflict of interests.

DATA AVAILABILITY

The data that support the findings of this article are not publicly available upon publication because it is not technically feasible and/or the cost of preparing, depositing, and hosting the data would be prohibitive within the terms of this research project. The data are available from the authors upon reasonable request.

APPENDIX A: SIMULATION METHODS

1. Simulation setup

The coarse-grained molecular dynamics simulations of two DNA molecules confined within a spherical cavity, were performed employing the Large-scale Atomic/Molecular Massively Parallel Simulator (LAMMPS) software package [62] using Imperial College London Research Computing service [63]. DNA molecules were modeled using bead-spring polymers, with each bead representing approximately 20 Å section (approximately six base pairs) of double-stranded DNA.

Initial polymer conformations were systematically generated using custom Python scripts. Twenty-five beads in the middle of each polymer were placed in a straight line at the center of the confining sphere at a separation of 23.9 Å, corresponding to the minimum of the potential well (see Fig. 5). The rest of the beads were placed sequentially in the confining sphere using a self-avoiding random walk, with collinear segments of 50 beads (matching the Kuhn length). Trial positions that exceeded the confinement radius or overlapped with any other bead were rejected and new random directions were sampled.

Confinement was imposed by a repulsive radial force acting near the boundary of the spherical volume, derived under the Derjaguin approximation to reflect the steric and electrostatic repulsion experienced by DNA near confining surfaces. This potential ensured that polymer beads remained within the spherical domain while allowing physically realistic proximity to the boundary (see main text for more details). Both bonded (polymer) and nonbonded (electrostatic) interactions were approximated through tailored potentials reflecting the physical electrostatic characteristics of DNA in solutions of the physiological electrolyte concentration, as described below.

Temperature was maintained at 37 °C using Langevin dynamics thermostating combined with NVE integration, employing a damping coefficient consistent with the DNA hydration environment. See below for estimation of this parameter. Equations of motion were integrated using a timestep of 500 fs.

Each simulation consisted of 500 million time steps, corresponding to a total simulation time of 250 μs. If the order parameter [defined in the main text, Eq. (6)] had not indicated convergence in pairing by the end of the simulation, the run was restarted using the final configuration. Computational runs were executed using parallel processing via Message Passing Interface (MPI), distributed across 64 CPU cores. Typical simulation runtimes for each 500 million timestep run were approximately 30 hours, enabling extensive parameter-space exploration within computationally tractable durations.

2. Interactions between beads in the same DNA chain

To maintain the mechanical properties of a DNA chain, we implemented a set of interactions between neighboring beads of the same polymer. For resistance to length deformations, a 12-6 Lennard-Jones (LJ) interaction was implemented between each two consecutive beads,

$$V_{\text{LJ}}(r) = 4\epsilon_{\text{LJ}} \left[\left(\frac{\sigma_{\text{LJ}}}{r} \right)^{12} - \left(\frac{\sigma_{\text{LJ}}}{r} \right)^6 \right], \quad (\text{A1})$$

where r is the center-to-center distance between the two consecutive beads and $\{\epsilon_{\text{LJ}}, \sigma_{\text{LJ}}\}$ is the set of LJ parameters estimated to reproduce the mechanical properties of DNA. Taking into consideration the radius of DNA, we set the minimum distance at $r_{\text{min}} = 2.1$ nm, from which we determined $\sigma_{\text{LJ}} = 2^{-1/6} r_{\text{min}} \approx 2.35$ nm. To set ϵ_{LJ} , we expanded Eq. (A1) to second order:

$$4\epsilon_{\text{LJ}} \left[\left(\frac{\sigma_{\text{LJ}}}{r} \right)^{12} - \left(\frac{\sigma_{\text{LJ}}}{r} \right)^6 \right] \approx \frac{18 \times 2^{2/3} \epsilon_{\text{LJ}}}{\sigma_{\text{LJ}}^2} (r - r_{\text{min}})^2 + O[(r - r_{\text{min}})^3], \quad (\text{A2})$$

and compared it with the elastic energy stored in a stretched rod characterized with the same Young modulus as that of a

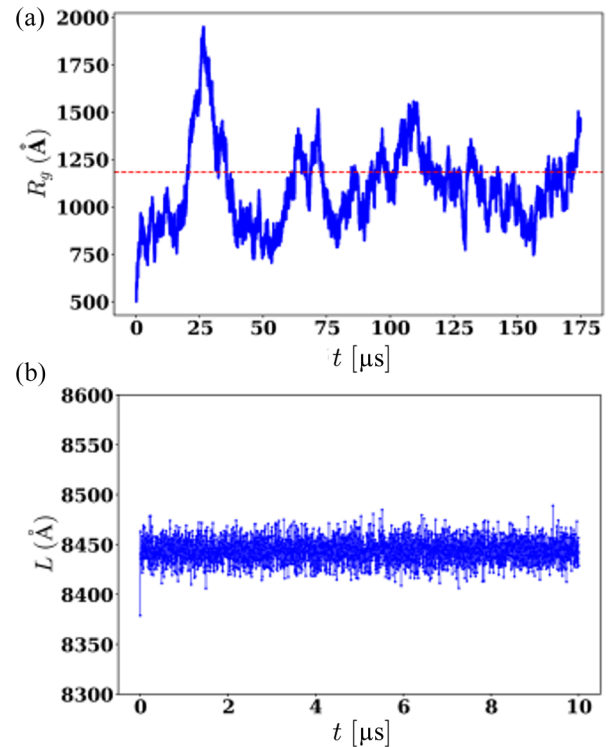


FIG. 8. A simulation of a single 2.4 kbp dsDNA chain in free space. (a) The time evolution of the gyration radius. Dashed line shows the theoretical prediction (1183.2 Å). (b) The time evolution of the total arc length of the chain. The value of the total arc length is shown to be practically a constant, albeit about 3% larger than the expected value due to the interspace separation of 1 Å between the surfaces of each two consecutive beads, and thermal expansion of the chain.

DNA molecule, $E \approx 1 \text{ nN nm}^{-2}$, to obtain $\epsilon_{\text{LJ}} \approx 21.4 k_B T$, where the temperature is set to 37°C .

To take into account resistance against bending, a cosine potential of the form

$$V_{\text{bend}}(\psi) = K(1 + \cos \psi) \quad (\text{A3})$$

was assigned for each two consecutive sections of beads, i.e., for each triplet of beads $\{i-1, i, i+1\}$. Here ψ is the angle between center-to-center line of beads $\{i-1, i\}$, to that of beads $\{i, i+1\}$. When the polymer is locally straight, we have $\psi = \pi$, which gives zero bending penalty. The K coefficient was set by expanding Eq. (A3) for ψ around the minimum energy to second order:

$$K(1 + \cos \psi) \approx \frac{1}{2}K(\psi - \pi)^2, \quad (\text{A4})$$

and comparing it with the elastic energy stored in a torsional spring having the same flexural rigidity as that of a DNA molecule, which came out as

$$K = \frac{\ell_p}{r_{\text{min}}} k_B T \approx 25 k_B T, \quad (\text{A5})$$

where $\ell_p \approx 50 \text{ nm}$ is the persistence length of DNA in normal physiological conditions.

To check if these interactions were implemented correctly, we ran the simulation for a single 2.4-kbp-long chain in free space, which is essentially a periodic simulation box with dimensions much larger than the theoretical gyration radius of

the polymer. We recorded the total arc length, L , and gyration radius, R_g , as a function of time. The gyration radius was calculated according to its definition,

$$R_g = \sqrt{\frac{1}{N} \sum_{i=1}^N |\mathbf{r}_i - \mathbf{r}_{\text{CoM}}|^2}, \quad \mathbf{r}_{\text{CoM}} = \frac{1}{N} \sum_{i=1}^N \mathbf{r}_i, \quad (\text{A6})$$

where $\mathbf{r}_i$ is the position of the center of the i -th bead, and $\mathbf{r}_{\text{CoM}}$ is the center of mass of the polymer.

The results for R_g and L are shown in Fig. 8, which indicate that the arc length of the chain is close to invariant under thermal fluctuations, as appropriate, and the mean steady-state gyration radius fluctuates about the value predicted by standard polymer theory, i.e., $R_g^{\text{(Theory)}} = \sqrt{\ell_c L/6}$, where $\ell_c = 2\ell_p$ is the Kuhn length of the polymer.

3. Homology recognition interaction coefficients

The interaction between homologous beads at center-to-center distance R is described by Eq. (4), inspired by the theory of electrostatic interaction of helical molecules [32]. Here the interaction is written in the form of an expansion in helical harmonics, $a_n(R)$, where we consider the first few dominant terms. For a local response of the electrolyte solution in which the DNA are solvated ($\epsilon = 80$), these are given by

$$a_0(R) = \frac{8\pi^2 \bar{\sigma}^2}{\epsilon} \left\{ (1 - \theta)^2 \frac{K_0(\kappa R)}{\kappa^2 K_1^2(\kappa \ell_0/2)} + \sum_{p,q=-\infty}^{\infty} \frac{f(p, \theta)^2}{\tilde{\kappa}_p^2 [K_p'(\tilde{\kappa}_p \ell_0/2)]^2} \frac{[K_{p-q}(\tilde{\kappa}_p R)]^2 I_q'(\tilde{\kappa}_2 \ell_0/2)}{K_q'(\tilde{\kappa}_p \ell_0/2)} \right\}, \quad (\text{A7})$$

$$a_{n \in \{1,2\}}(R) = \frac{16\pi^2 \bar{\sigma}^2}{\epsilon} \frac{f(n, \theta)^2}{\tilde{\kappa}_n^2} \frac{K_0(\tilde{\kappa}_n R)}{[K_n'(\tilde{\kappa}_n \ell_0/2)]^2}, \quad (\text{A8})$$

where I_n, K_n, I_n', K_n' are the modified Bessel functions and their derivatives, κ is the inverse Debye length evaluated at physiological conditions, $\bar{\sigma} \approx 16.3 \text{ } \mu\text{C cm}^{-2}$ is the average surface charge density of the uncompensated phosphate charge, $\tilde{\kappa}_n = \sqrt{\kappa^2 + 4\pi^2 n^2/H^2}$ is the n th harmonic effective inverse screening length, where H is the helical pitch of DNA. Lastly,

$$f(n, \theta) = f_1 \theta + f_2 (-1)^n \theta - \cos(n\tilde{\phi}_s/2), \quad (\text{A9})$$

where f_1 and $f_2 = 1 - f_1$ are the fractions of counterions adsorbed to the minor and major grooves, respectively, and $\tilde{\phi}_s \approx 0.8\pi$ is the azimuthal width of the minor groove. We assumed a type of salt that dominantly adsorbs to the major groove, such as Ca^{2+} or Mg^{2+} , for which we set $f_1 = 0.2$ and $f_2 = 0.8$ in all simulations.

Assuming that the rotation of the beads is instantaneous compared to diffusive properties of the polymer, the “relative azimuthal orientation” $\delta\phi$ between two beads can be taken as the path that minimizes the electrostatic interaction energy:

$$\delta\phi_{\text{min}}(R) = \begin{cases} \arccos \left[\frac{a_1(R)e^{-|i-j|\ell_0/\lambda_c}}{4a_2(R)e^{-4|i-j|\ell_0/\lambda_c}} \right], & \frac{a_1(R)e^{-|i-j|\ell_0/\lambda_c}}{4a_2(R)e^{-4|i-j|\ell_0/\lambda_c}} < 1; \\ 0 & \frac{a_1(R)e^{-|i-j|\ell_0/\lambda_c}}{4a_2(R)e^{-4|i-j|\ell_0/\lambda_c}} > 1. \end{cases} \quad (\text{A10})$$

The exact expression for the electrostatic energy of two beads originating from different DNA molecules depends on the interaction length, L , in a nonlinear fashion as shown in Eq. (3). It is computationally expensive to quantify the interaction length for interacting sections and produce, accordingly, a time-changing electrostatic profile. Therefore, we prescribed an approximated long interaction length ($L \gg \lambda_c$) forcefield for the beads. The electric force $f_{\text{elec}} = -\partial U_{\text{elec}}/\partial \Delta z$ that such a potential creates is a very good approximation for

small shifts $\Delta z \ll \lambda_c$, which is self-consistent with the final outcome of the simulation in which the genes that pair up do so with a very small shift $\Delta z \ll \lambda_c$. Furthermore, this approximation turns out to be quite accurate even for large Δz shifts, as depicted in Fig. 9.

4. Langevin dynamics

At each time step, the position of each bead, $\mathbf{r}$, was displaced according to the Langevin equation, which

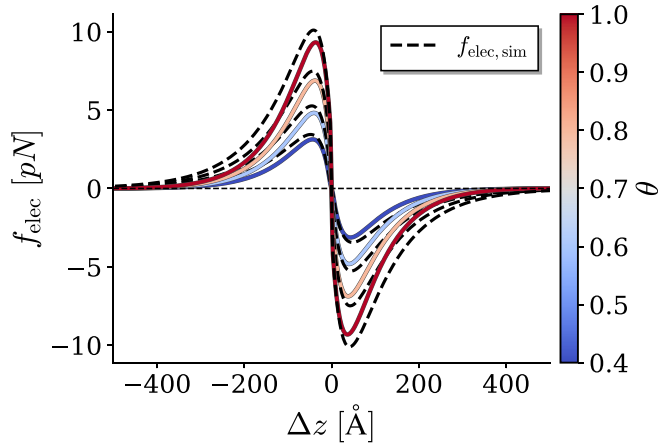


FIG. 9. Exact Theory vs Simulation Approximation. The electric force, $f_{\text{elec}} = -\partial U_{\text{elec}}/\partial \Delta z$, as a function of the axial shift, Δz , stemming from (1) the full electrostatic potential, Eq. (2) (solid curves), and (2) the approximated version, Eq. (4), used in the simulation (dashed), where $\Delta z = |i - j|\ell_0$ was taken to be continuous for visualization. Curves were plotted for a range of different values of intrinsic charge compensation, θ .

reads as

$$m \frac{d^2 \mathbf{r}}{dt^2} = -\gamma \frac{d\mathbf{r}}{dt} + \mathbf{F}_{\text{nonstochastic}}(t) + \mathbf{R}(t), \quad (\text{A11})$$

where m is the mass of the bead, γ is the hydrodynamic drag coefficient of the bead in creeping flow, $\mathbf{F}_{\text{nonstochastic}}$ is the sum of mechanical, electrostatic and hydration forces acting on the bead, and $\mathbf{R}$ is a random force taken from a Gaussian distribution which satisfies the following autocorrelation:

$$\langle \mathbf{R}_i(t) \mathbf{R}_j(t') \rangle = 2k_B T \delta_{ij} \delta(t - t'). \quad (\text{A12})$$

The hydrodynamic drag coefficient was approximated using Stokes' law:

$$\gamma = 6\pi\mu \left(\frac{\ell_0}{2} \right) \approx 1.3 \times 10^{-11} \text{ kg s}^{-1}, \quad (\text{A13})$$

where $\mu \approx 0.7 \times 10^{-3} \text{ Pa s}$ is the dynamic viscosity of water in 37°C and $\ell_0 = 2 \text{ nm}$ is the diameter of a single bead in the DNA chain. The molar mass of each bead is approximated to be 3960 g mol^{-1} , leading to a Langevin damping time of $\tau = m_{\text{bead}}/\gamma \approx 505 \text{ fs}$. By estimating this parameter for water our simulation replicates real time evolution.

APPENDIX B: DATA ANALYSIS

To identify plateaus in the order parameter, we defined the stationary condition:

$$\frac{|\Delta \bar{S}_i|}{\bar{S}_i} < 0.01, \quad (\text{B1})$$

where $\bar{S}_i$ is the order parameter averaged over a time window of $25 \mu\text{s}$ and $|\Delta \bar{S}_i| = \bar{S}_{i+1} - \bar{S}_i$ is the change of $\bar{S}_i$ over two consecutive windows. The $25 \mu\text{s}$ time window was chosen to be short enough to resolve the onset of a steady state, while remaining long enough to reduce sensitivity to small fluctuations and transient decreases in pairing rates. Thus, a plateau was defined as a region in which the stationary condition is

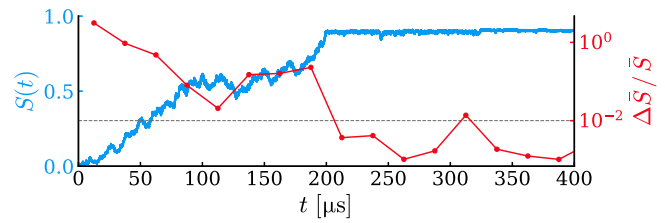


FIG. 10. Determination of plateau in pairing for the case of $R_c/R_g = 0.5$ and $\theta = 0.85$. Progression of order parameter over time, $S(t)$, is indicated in blue. The corresponding $|\Delta \bar{S}_i|/\bar{S}_i$ is in red, with the dashed line marking the stationary condition threshold $|\Delta \bar{S}_i|/\bar{S}_i < 0.01$.

satisfied. In Fig. 10 we show both $\Delta(\bar{S})/\bar{S}$ and S as a function of time for an example case.

APPENDIX C: KINETIC MODELS

1. Simple kinetic model

Here we describe the kinetics of homologous pairing, which treats homologous beads independent of their neighbors, in the case when the system does not get “trapped” on the way to reaching equilibrium. After initiating a segment of 25 beads in parallel juxtaposition, the time evolution of the fraction of paired homologous beads, $n(t) = N_{\text{paired}}(t)/N$, which is complementary to the fraction of *unpaired* homologous beads, can be approximated as

$$\frac{dn}{dt} = k_{\text{pair}}(1 - n)^2 - k_{\text{unpair}}n, \quad (\text{C1})$$

where k_{pair} and k_{unpair} are the kinetic coefficients for, respectively, creating or destroying the pairing between two homologous beads. Applying initial conditions, $n(t_0) = n_0$, where t_0 is empirically chosen time, when the kinetics start to follow the Simple Kinetic Model, Eq. (C1) is solved analytically by separation of variables:

$$\int_{t_0}^t dt = k_{\text{pair}}^{-1} \int_{n_0}^{n(t)} \frac{dn}{n^2 - (k_{\text{unpair}}/k_{\text{pair}} + 2)n + 1}, \quad (\text{C2})$$

which gives

$$t - t_0 = \frac{2}{\sqrt{k_{\text{unpair}}\sqrt{k_{\text{unpair}} + 4k_{\text{pair}}}}} \left\{ \text{arctanh} \left[\frac{k_{\text{unpair}} - 2(n - 1)k_{\text{pair}}}{\sqrt{k_{\text{unpair}}\sqrt{k_{\text{unpair}} + 4k_{\text{pair}}}}} \right] - \text{arctanh} \left[\frac{k_{\text{unpair}} - 2(n_0 - 1)k_{\text{pair}}}{\sqrt{k_{\text{unpair}}\sqrt{k_{\text{unpair}} + 4k_{\text{pair}}}}} \right] \right\}.$$

Rearranging the right-hand side of this equation to solve for $n(t)$, and defining $\beta = k_{\text{unpair}}/k_{\text{pair}}$, we obtain Eq. (11). At small times, $t \ll 2/k_{\text{pair}}\sqrt{\beta^2 + 4\beta}$, Eq. (11) gives a linear rise:

$$n \left(t - t_0 \ll \frac{2}{k_{\text{pair}}\sqrt{\beta^2 + 4\beta}} \right) = n_0 + k_{\text{pair}}(t - t_0), \quad (\text{C3})$$

and at long enough times, $t - t_0 \gg 2/k_{\text{pair}}\sqrt{\beta^2 + 4\beta}$, the number of paired sites approaches its steady-state value:

$$n_{\text{SS}} = 1 - \frac{1}{2}\beta \left(\sqrt{1 + \frac{4}{\beta}} - 1 \right), \quad (\text{C4})$$

which gives $n_{\text{SS}} \rightarrow 0$ for $\beta \rightarrow \infty$, and $n_{\text{SS}} \rightarrow 1$ for $\beta \rightarrow 0$, as required.

We systematically set t_0 such that $n(t_0) = 0.75n_{\text{eq}}$, i.e., the fraction of paired beads at that time equals 75% of its equilibrium value. This choice reflects the fact that the Simple Kinetic Model is not intended to describe the early-time dynamics, but instead provides an accurate description in the long-time regime approaching equilibrium.

2. Detailed master equation

a. Model formulation

To incorporate correlations along the chain, we track the probability that each bead is paired, $P_i(t)$. The key modeling assumption is nearest-neighbor dependence, i.e., the kinetics at bead i depends on whether beads $i - 1$ and $i + 1$ are paired. This introduces three local environments for bead i :

- (1) Both neighbors are paired
- (2) Exactly one neighbor is paired
- (3) Both neighbors are unpaired.

We assign each environment $s \in \{1, 2, 3\}$ its own pair-formation and dissociation rates $k_{\text{pair}}^{(s)}$ and $k_{\text{unpair}}^{(s)}$. By expressing the probability of each neighbor configuration in terms of P_{i-1} , P_i and P_{i+1} , the overall evolution of $P_i(t)$ can be written as gains from pair formation minus the losses from dissociation, yielding Eq. (12) in the main text. For compactness we let $P_i \equiv P_i(t)$.

b. Detailed-balance approximation

Directly fitting all six rates ($k_{\text{pair}}^{(s)}$, $k_{\text{unpair}}^{(s)}$ for $s = 1, 2, 3$) can be poorly constrained. We therefore reduce the number of independent parameters by imposing an approximate detailed-balance condition within each local environment s . At equilibrium, conditional on being in scenario s , we can therefore write

$$k_{\text{pair}}^{(s)} P^{\text{eq}}(\text{unpair}|s) = k_{\text{unpair}}^{(s)} P^{\text{eq}}(\text{pair}|s), \quad (\text{C5})$$

where $P^{\text{eq}}(\text{unpair}|s)$ and $P^{\text{eq}}(\text{pair}|s)$ are equilibrium probabilities that two homologous beads are unpaired and paired providing that scenario s occurs. Applying Eq. (C5) to Eq. (12) eliminates $k_{\text{unpair}}^{(s)}$ for each s , thereby reducing the number of free kinetic parameters used in the fitting, resulting in

$$\begin{aligned} \frac{dP_i}{dt} = & k_{\text{pair}}^{(1)} (1 - P_i) P_{i-1} P_{i+1} \\ & + k_{\text{pair}}^{(2)} (1 - P_i) [P_{i-1} (1 - P_{i+1}) + P_{i+1} (1 - P_{i-1})] \\ & + k_{\text{pair}}^{(3)} (1 - P_{i-1}) (1 - P_{i+1}) (1 - P_i) \\ & - k_{\text{pair}}^{(1)} \frac{P^{\text{eq}}(\text{unpair}|1)}{P^{\text{eq}}(\text{pair}|1)} P_{i-1} P_{i+1} P_i \end{aligned}$$

$$\begin{aligned} & - k_{\text{pair}}^{(2)} \frac{P^{\text{eq}}(\text{unpair}|2)}{P^{\text{eq}}(\text{pair}|2)} P_i [P_{i-1} (1 - P_{i+1}) \\ & + P_{i+1} (1 - P_{i-1})] \\ & - k_{\text{pair}}^{(3)} \frac{P^{\text{eq}}(\text{unpair}|3)}{P^{\text{eq}}(\text{pair}|3)} P_i (1 - P_{i-1}) (1 - P_{i+1}). \end{aligned} \quad (\text{C6})$$

c. Numerical solution and boundary conditions

We solve Eq. (C6) for $P_i(t)$ numerically using a fourth-order Runge-Kutta (RK4) integrator. Initial conditions match our protocol for MD simulations: $P_i(t = 0) = 0$ for $i \in [1, 187]$ and $i \in [213 - 400]$, and $P_i(t = 0) = 1$ for $i \in [188, 212]$. To implement open boundaries and allow end effects such as fraying, we introduce two “ghost” beads at $i = 0$ and $i = N + 1$ with fixed pairing probability $P_0(t) = P_{N+1}(t) = 0$. This prevents the “both-neighbors-paired” environment from occurring at terminal beads and naturally incorporates reduced stability at the chain ends.

d. Connection to observables and fitting

To compare the dynamics modeled in this master equation to simulation, we approximate the instantaneous paired fraction by the mean of bead-pairing probabilities:

$$n(t) \approx \frac{1}{N} \sum_{i=1}^N P_i(t). \quad (\text{C7})$$

For a given parameter set, we integrate the master equation to obtain $n(t)$ and fit to the simulation times series data. The resulting curves are shown as solid lines in Fig. 7(b) and fitted parameters in Table I.

APPENDIX D: GOING BEYOND $R_c = R_g$

To assess whether $R_c/R_g = 1$ represents the onset of the unconfined regime, we performed simulations at $R_c/R_g = 2$.

The observed effect is entropic in nature: the DNA volume fraction decreases by an order of magnitude, increasing the free energy gap between paired and unpaired states. Pairing therefore still proceeds through the same mechanism (zip-ping), but with at a reduced rate reflecting the larger entropic penalty [see Figs. 11(a) and 11(b)].

To illustrate the increase of entropic penalty of pairing, we compare effective free energy profiles for $R_c/R_g = 1$ and $R_c/R_g = 2$ in Fig. 11(c). The well is shallower for $R_c/R_g = 2$, indicating reduced stability of paired state, as predicted. We further compare the obtained pairing times; τ changes weakly with confinement for $R_c/R_g > 1$ compared to $R_c/R_g < 1$. Both results confirm our predictions: above $R_c/R_g = 1$ pairing happens through the same mechanism (zipping), which is slightly quicker for $R_c/R_g = 1$ reflecting smaller entropic cost of pairing. In contrast, going below $R_c/R_g = 1$, pairing times change more dramatically, as confinement promotes nucleation of new pairing contacts. Those results confirm

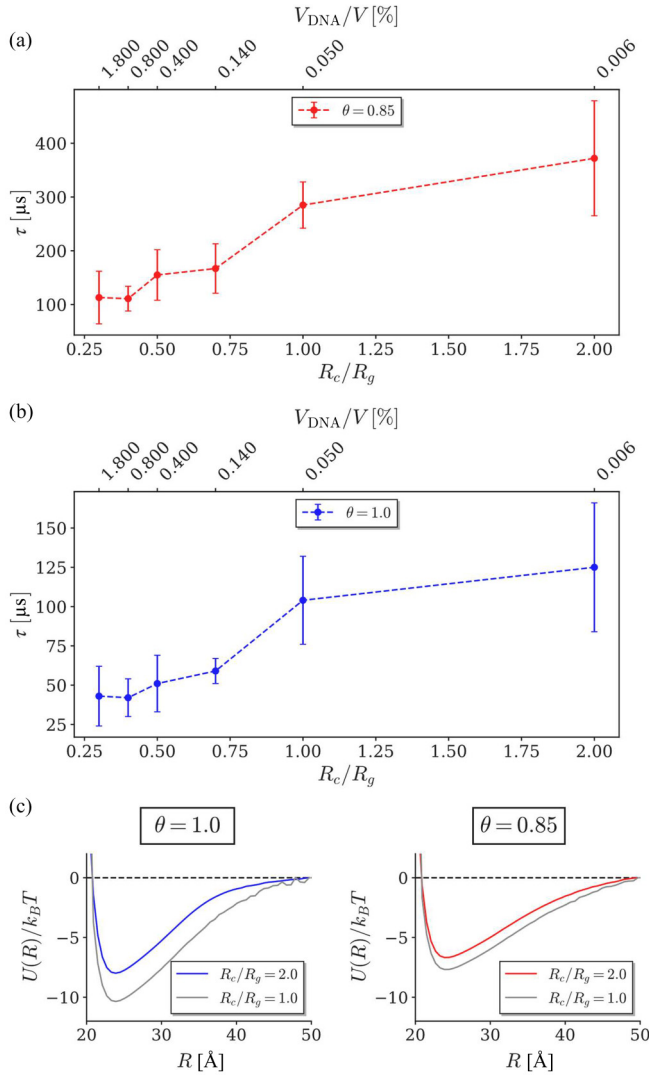


FIG. 11. Comparison of pairing at $R_c/R_g = 1.0$ and $R_c/R_g = 2$. Characteristic pairing time τ as a function of confinement for (a) $\theta = 0.85$ and (b) $\theta = 1.0$. Error bars present the standard deviation across different repeats. (c) Effective free energy profiles; comparison between $R_c/R_g = 1$ (in gray) and $R_c/R_g = 2$ (blue for $\theta = 1.0$, red for $\theta = 0.85$).

that $R_c/R_g = 1$ marks the transition between diffusion-limited regime and ‘reaction-limited’ regime.

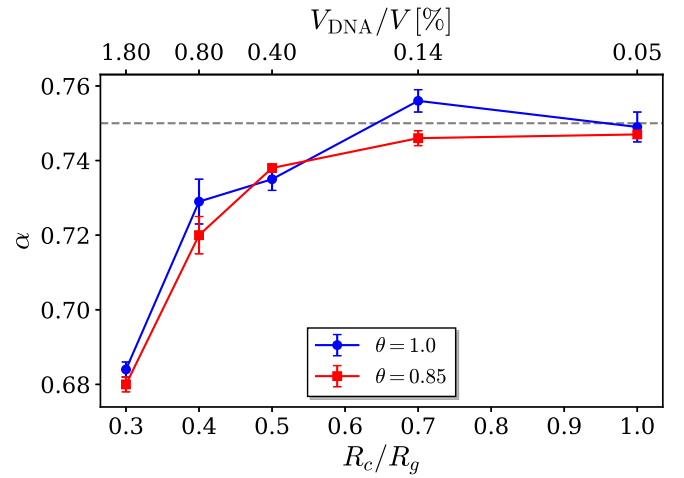


FIG. 12. Subdiffusion in confinement. Anomalous diffusion exponent α extracted from MSD fits of individual beads from the formula $\text{MSD} \sim t^\alpha$, as a function of confinement radius R_c/R_g . The horizontal line indicates $\alpha = 3/4$, the theoretical value for a monomer in a free semiflexible chain [45]. Bars represent standard deviation over independent simulations.

APPENDIX E: ANOMALOUS DIFFUSION UNDER CONFINEMENT

To quantify the effect of self-crowding on monomer dynamics, we performed additional simulations of two 2.4 kbp DNA molecules confined within spherical cavities of radii $R_c/R_g = 0.3$ – 1.0 , without electrostatic recognition interactions. The mean-squared displacement (MSD) of individual beads was computed and fitted to the general law of subdiffusion, $\text{MSD} \propto t^\alpha$, where t is time, and α is the anomalous exponent. The fitted values of α as a function of R_c/R_g are shown in Fig. 12. As confinement tightens, α decreases monotonically from its unconfined value of $3/4$, consistent with the known bending-mode subdiffusion of semiflexible chains [45], indicating progressive suppression of long-wavelength bending modes by self-crowding. Free-particle diffusion ($\alpha = 1$) is not expected for a subpersistence length segment within a semiflexible polymer, nor observed. This confirms that self-crowding at strong confinement increasingly hinders the diffusive encounters required for nucleation.

[1] J. Ringo, *Fundamental Genetics* (Cambridge University Press, Cambridge, 2004).
 [2] D. R. F. Leach, *Genetic Recombination* (Blackwell Science, Oxford, 1996).
 [3] A. Aquilera and R. Rothstein, *Molecular Genetic Recombination* (Springer, Berlin, 2007).
 [4] A. Marin-Gonzalez, A. T. Rybczynski, M. M. Nilavar, D. Nguyen, A. G. Li, V. Karwacki-Neisius, R. S. Zou, F. J. Avilés-Vázquez, M. T. Kanemaki, R. Scully, and T. Ha, Cohesin drives chromatin scanning during the rad51-mediated homology search, *Science* **390**, eadw1928 (2025).

[5] F. Teloni, Z. Takacs, M. Mitter, C. C. H. Langer, I. Prlesi, T. L. Steinacker, and V. P. Reuter, Cohesin guides homology search during DNA repair using loops and sister chromatid linkages, *Science* **390**, eadw0566 (2025).
 [6] S. Inoue, S. Sugiyama, A. A. Travers, and T. Ohyama, Self-assembly of double-stranded DNA molecules at nanomolar concentrations, *Biochemistry* **46**, 164 (2007).
 [7] G. Baldwin, N. J. Brooks, R. Robson, A. Wynveen, A. Goldar, S. Leikin, J. M. Seddon, and A. A. Kornyshev, DNA double helices recognize mutual sequence homology in a protein environment, *J. Phys. Chem. B* **112**, 1060 (2008).

- [8] C. Danilowicz, C. H. Lee, K. Kim, K. Hatch, V. W. Coljee, N. Kleckner, and M. Prentiss, Single molecule detection of direct, homologous, DNA/DNA pairing, *Proc. Natl. Acad. Sci. USA* **106**, 19824 (2009).
- [9] D. J. (O')Lee, C. Danilowicz, C. Rochester, A. A. Kornyshev, and M. Prentiss, Evidence of protein-free homology recognition in magnetic bead force-extension experiments, *Proc. R. Soc. A* **472**, 20160186 (2016).
- [10] J.-I. Nishikawa and T. Ohyama, Selective association between nucleosomes with identical DNA sequences, *Nucleic Acids Res.* **41**, 1544 (2013).
- [11] A. Stannard, E. Haimov, J. G. Hedley, Y. Xiao, M. Di Antonio, G. Oshanin, C. Danilowicz, M. Prentiss, L. Di Michele, and A. A. Kornyshev, Direct evidence and quantification of homologous recognition between DNA duplexes, *Proc. Natl. Acad. Sci. USA* **123**, e2530949123 (2026).
- [12] D. Zickler, From early homologue recognition to synaptonemal complex formation, *Chromosoma* **115**, 158 (2006).
- [13] A. Barzel and M. Kupiec, Finding a match: How do homologous sequences get together for recombination? *Nat. Rev. Genet.* **9**, 27 (2008).
- [14] S. Henikoff, Nuclear organization and gene expression: Homologous pairing and long-ranged interactions, *Curr. Opin. Cell Biol.* **9**, 388 (1997).
- [15] R. E. Wagner and M. Radman, A mechanism for initiation of genetic recombination, *Proc. Natl. Acad. Sci. USA* **72**, 3619 (1975).
- [16] G. G. Doyle, A general theory of chromosome pairing based on the palindromic DNA model of Sobell with modifications and amplification, *J. Theor. Biol.* **70**, 171 (1978).
- [17] D. R. Forsdyke, A stem-loop "kissing" model for the initiation of recombination and the origin of introns, *Mol. Biol. Evol.* **12**, 949 (1995).
- [18] A. A. Kornyshev and S. Leikin, Sequence recognition in the pairing of DNA duplexes, *Phys. Rev. Lett.* **86**, 3666 (2001).
- [19] A. A. Kornyshev, D. J. Lee, S. Leikin, and A. Wynveen, Structure and interactions of biological helices, *Rev. Mod. Phys.* **79**, 943 (2007).
- [20] A. A. Kornyshev, Physics of DNA: Unravelling hidden abilities encoded in the structure of 'the most important molecule', *Phys. Chem. Chem. Phys.* **12**, 12352 (2010).
- [21] E. Gladyshev and N. Kleckner, Direct recognition of homology between double helices of DNA in *Neurospora crassa*, *Nat. Commun.* **5**, 3509 (2014).
- [22] E. Gladyshev and N. Kleckner, Recombination-independent recognition of DNA homology for repeat-induced point mutation, *Curr. Genet.* **63**, 389 (2017).
- [23] A. K. Mazur, Homologous pairing between long DNA double helices, *Phys. Rev. Lett.* **116**, 158101 (2016).
- [24] A. K. Mazur, T.-S. Nguyen, and E. Gladyshev, Direct homologous dsDNA-dsDNA pairing: How, where and why? *J. Mol. Biol.* **432**, 737 (2020).
- [25] A. A. Kornyshev and A. Wynveen, The homology recognition well as an innate property of DNA structure, *Proc. Natl. Acad. Sci. USA* **106**, 4683 (2009).
- [26] R. Sinden, *DNA Structure and Function* (Academic Press, New York, 1994).
- [27] W. Kabsch, C. Sander, and E. N. Trifonov, The ten helical twist angles of B-DNA, *Nucleic Acids Res.* **10**, 1097 (1982).
- [28] W. K. Olson, A. A. Gorin, X. J. Lu, L. M. Hock, and V. B. Zhurkin, DNA sequence-dependent deformability deduced from protein-DNA crystal complexes, *Proc. Natl. Acad. Sci. USA* **95**, 11163 (1998).
- [29] A. A. Gorin, V. B. Zhurkin, and K. W. Olson, B-DNA twisting correlates with base-pair morphology, *J. Mol. Biol.* **247**, 34 (1995).
- [30] H. M. Berman, W. K. Olson, D. L. Beveridge, J. Westbrook, A. Gelbin, T. Demeny, S. H. Hsieh, A. R. Srinivasan, and B. Schneider, The nucleic acid database. A comprehensive relational database of three-dimensional structures of nucleic acids, *Biophys. J.* **63**, 751 (1992).
- [31] S. Neidle, *Principles of Nucleic Acid Structure* (Elsevier, London, 2008).
- [32] A. A. Kornyshev and S. Leikin, Electrostatic zipper motif for DNA aggregation, *Phys. Rev. Lett.* **82**, 4138 (1999).
- [33] A. A. Kornyshev, D. J. Lee, A. Wynveen, and S. Leikin, Signatures of DNA flexibility, interactions and sequence-related structural variations in classical X-ray diffraction patterns, *Nucleic Acids Res.* **39**, 7289 (2011).
- [34] A. Wynveen, D. J. Lee, A. A. Kornyshev, and S. Leikin, Helical coherence of DNA in crystals and solution, *Nucleic Acids Res.* **36**, 5540 (2008).
- [35] D. J. Lee, A. Wynveen, T. Albrecht, and A. A. Kornyshev, Which way up? Recognition of homologous DNA segments in parallel and antiparallel alignment, *J. Chem. Phys.* **142**, 045101 (2015).
- [36] J. G. Hedley, E. Haimov, and A. A. Kornyshev, Water structural effects on DNA-DNA interactions and homologous recognition, *Physica A* **647**, 129894 (2024).
- [37] E. Haimov, J. G. Hedley, and A. A. Kornyshev, Nonlocal structural effects of water on DNA homology recognition, *J. Phys.: Condens. Matter.* **36**, 40LT01 (2024).
- [38] A. A. Kornyshev and S. Leikin, Theory of interaction between helical molecules, *J. Chem. Phys.* **107**, 3656 (1997).
- [39] J. G. Hedley, K. Coshic, A. Aksimentiev, and A. A. Kornyshev, Electric field of DNA in solution: Who is in charge? *Phys. Rev. X* **14**, 031042 (2024)b.
- [40] E. Haimov and A. A. Kornyshev, Statistical mechanics of homologous pairing of long double-stranded DNA, *J. Chem. Phys.* **162**, 215102 (2025).
- [41] J. Israelachvili, *Surface Forces*, 3rd ed. (Academic Press, New York, 2011).
- [42] M. L. J. Moncany and E. Kellenberger, High magnesium content of *Escherichia coli* B, *Experientia* **37**, 846 (1981).
- [43] <https://youtu.be/xOvuRbSZvtQ>.
- [44] https://youtu.be/_D8-KT4FuJg.
- [45] R. Granek, From semi-flexible polymers to membranes: Anomalous diffusion and reptation, *J. de Physique II* **7**, 1761 (1997).
- [46] P. Szalaj and D. Plewczynski, Three-dimensional organization and dynamics of the genome, *Cell Biol. Toxicol.* **34**, 381 (2018).
- [47] B. Bonev and G. Cavalli, Organization and function of the 3D genome, *Nat. Rev. Genet.* **17**, 661 (2016).
- [48] A. Németh and G. Längst, Genome organization in and around the nucleolus, *Trends Genet.* **27**, 149 (2011).
- [49] J. Berry, C. P. Brangwynne, and M. Haataja, Physical principles of intracellular organization via active and passive phase transitions, *Rep. Prog. Phys.* **81**, 046601 (2018).

- [50] W. Stroberg and S. Schnell, Do cellular condensates accelerate biochemical reactions? Lessons from microdroplet chemistry, *Biophys. J.* **115**, 3 (2018).
- [51] K. D. Young, The selective value of bacterial shape, *Microbiol. Mol. Biol. Rev.* **70**, 660 (2006).
- [52] L. Willis and K. S. Huang, Sizing up the bacterial cell cycle, *Nat. Rev. Microbiol.* **606-620**, 15 (2017).
- [53] A. C. Chien, N. S. Hill, and P. A. Levin, Cell size control in bacteria, *Curr. Biol.* **22**, R340 (2012).
- [54] National Research Council (US) Steering group for the workshop on size limits of very small microorganisms, in *Size Limits of Very Small Microorganisms: Proceedings of a Workshop* (National Academies Press, Washington, DC, 1999), pp. 16–25.
- [55] A. A. Kornyshev and A. Wynveen, Nonlinear effects in torsional adjustment of interacting DNA, *Phys. Rev. E* **69**, 041905 (2004).
- [56] A. G. Cherstvy, A. A. Kornyshev, and S. Leikin, Torsional deformation of double helix in interaction and aggregation of DNA, *J. Phys. Chem. B* **108**, 6508 (2004).
- [57] National Human Genome Research Institute, Human genomic variation, <https://www.genome.gov/about-genomics/educational-resources/fact-sheets/human-genomic-variation> (2023).
- [58] K. P. Adamala, M. Dogterom, Yu Elani, P. Schwille, M. Takinoue, and T.-Y. Tang, Present and future of synthetic cell development, *Nat. Rev. Mol. Cell Biol.* **25**, 162 (2024).
- [59] Y. Elani and J. M. Seddon, What it means to be alive: A synthetic cell perspective, *Interface Focus* **13**, 20230036 (2023).
- [60] C. Guindani, L. Claire da Silva, S. Cao, T. Ivanov, and K. Landfester, Synthetic cells: From simple bio inspired modules to sophisticated integrated systems, *Angew. Chem.* **61**, e202110855 (2022).
- [61] C. A. Azaldegui, A. G. Vecchiarelli, and J. S. Biteen, The emergence of phase separation as an organizing principle in bacteria, *Biophys. J.* **120**, 1123 (2021).
- [62] A. P. Thompson, H. M. Aktulga, R. Berger, D. S. Bolintineanu, W. M. Brown, P. S. Crozier, P. J. in 't Veld, A. Kohlmeyer, S. G. Moore, T. D. Nguyen, R. Shan, M. J. Stevens, J. Tranchida, C. Trott, and S. J. Plimpton, LAMMPS—A flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales, *Comput. Phys. Commun.* **271**, 108171 (2022).
- [63] Imperial College London Research Computing Service, doi:10.14469/hpc/2232.