# Interphase Human Chromosome Exhibits Out of Equilibrium Glassy Dynamics

Guang Shi[1], Lei Liu[2], Changbong Hyeon[2] and D. Thirumalai[3*]

[1]*Biophysics Program,*

*Institute for Physical Science and Technology,*

*University of Maryland,*

*College Park, MD 20742*

[2]*Korea Institute for Advanced Study,*

*Seoul 02455, Republic of Korea*

[3]*Department of Chemistry,*

*University of Texas at Austin, Texas, 78712*

[*] Corresponding author: dave.thirumalai@gmail.com

**Abstract**

The structural organization of the condensed chromosomes, which hold the secrets to gene regulation, is being revealed using chromosome conformation capture experiments and super resolution imaging techniques. To determine the structure and dynamics of human interphase chromosomes, we created the Chromosome Copolymer Model (CCM) by representing the chromosomes as a self-avoiding polymer with two loci types corresponding to euchromatin and heterochromatin. Using advanced clustering algorithms we establish quantitatively that the simulated contact maps for chromosomes 5 and 10 and those inferred from Hi-C experiments are in agreement. Ward Linkage Matrix (WLM), constructed from spatial distance information, shows that the spatial organization of the Topologically Associating Domains (TADs) and compartments predicted from simulation are in agreement with inferred WLM computed using data from super resolution microscopy experiments. Glassy dynamics is manifested in the stretched exponential relaxation of the structure factor, a pronounced peak in the fourth order susceptibility accounting for the fluctuations in $F(k,t)$, and caging in the mean square displacement of individual loci, $\Delta_i(t) \sim t^{\alpha}$ with $0 < \alpha < 1$. Remarkably, the distribution of $\alpha$, is extremely broad suggestive of highly heterogeneous dynamics. Chromosome folding is hierarchical involving the formation of chromosome droplets (CDs) on short genomic scale followed by coalescence of the CDs, reminiscent of Ostwald ripening. We propose that glassy landscapes for the folded chromosomes might provide a balance between stability and biological functions.

## Introduction

The organization of chromosomes without topological entanglement or knot formation in the crowded tight space of the nucleus is truly remarkable. Understanding the structural organization and the dynamics of eukaryotic chromosomes and the mechanism of chromosome territories formation holds the key to enunciating genome functions. Glimpses into the structures of the chromosomes have emerged, thanks to spectacular advances in Chromosome Conformation Capture (3C, 4C, 5C, and Hi-C) experiments (1–3), from which the probability, $P(s)$, that two loci separated by a certain genomic distance ($s$) are in contact can be inferred. The set of $P(s)$ as function of $s$ constitutes the contact map. More recently, super resolution imaging methods (4, 5) have more directly determined the positions of the loci of single chromosomes, thus providing a much-needed link to the indirect measure of spatial organization revealed through contact maps. The experiments by Zhuang and coworkers (4, 5) are of great value because the number of constraints needed to unambiguously infer structures from contact maps alone is very large (6).

Contact maps, constructed from Hi-C experiments, revealed that chromosome is organized into compartments on genomic length scales exceeding mega base pairs (Mbps) (2, 3). The partitioning of the structure into compartments are highly correlated with histone markers of the chromatin loci (3), implying that contacts are enriched within a compartment and depleted between different compartments. The loci associated with active histone markers and those associated with repressive histone markers localize spatially in different compartments. Higher resolution Hi-C experiments (3) have also identified Topological Associated Domains (TADs) on scales on the order of hundreds of kilobases (1). The TADs are the square patterns along the diagonal of the contact maps in which the probability of two loci being in contact is more probable than between two loci belonging to distinct TADs. Super resolution imaging experiments (4) show most directly that TADs belonging to distinct compartments are spatially separated in the chromosomes.

The experimental studies have inspired a variety of polymer models (7–19), which have

provided insights into many aspects of chromosome organization. These studies are particularly important because finding a unique solution (if one exists) to the inverse problem of deriving spatial structures from contact maps is most challenging (6). Some of the features in the contact maps, such as the probability, $P(s)$ versus $s$ and compartments may be computed using a homopolymer model without accounting for epigenetic states whereas fine structures such as TADs require copolymer models.

Although the polymer models, which have a rich history in this area (7–19), have been most useful in predicting the probabilistic nature of chromosome organization inferred from Hi-C experiments. However, biological functions, such as search for genes by transcription factors or mechanism for DNA damage repair, not only depend on genome structure but also the associated dynamics. The latter has received much less attention (20, 21). Experiments have revealed that genome-wide chromatin dynamics (22–24) of chromatin fiber in mammalian cells exhibit heterogeneous sub-diffusive behavior. Thus, it is important to understand how the slow dynamics of the individual locus and long length scale coherent collective motions, observed in Human HeLa cells (23), emerge from the highly organized chromosomes.

Here, we develop a copolymer model to describe both the structure and dynamics of human interphase chromosomes. We show that in order to capture the structural features faithfully, at least two types of beads, representing active and repressive loci are needed. Simulations of the resulting Chromosome Copolymer Model (CCM) for human interphase chromosomes 5 and 10 show that the major structural characteristics, such as the scaling of $P(s)$ as a function of $s$, compartments, and TADs indicated in the Hi-C contact maps are faithfully reproduced. We use sophisticated clustering algorithms to quantitatively compare the simulated contact maps and those inferred from Hi-C experiments. The compartment feature noted in the Hi-C contact map is due to micro-phase separation between chromosome loci associated with different epigenetic states, implying that a copolymer model is needed for characterizing large scale genome organization. The TADs emerge by incorporating experimentally inferred positions of the loop anchors, whose formation is facilitated by CTCF motifs. Chromosome dynamics is highly hetero-

geneous and exhibits many of the characteristics of out of equilibrium glassy dynamics including stretched exponential decay of the structure factor ($F_s(k,t)$), a non-monotonicity behavior in the time dependence of the fourth order susceptivity associated with fluctuations in $F_s(k,t)$. Of particular note is the remarkable cell-to-cell and loci-to-loci variation in the time ($t$) dependence of the mean square displacement, $\Delta_i(t)$, of the individual loci. The distribution, $P(\alpha)$ of the exponent associated with the increase in $\Delta_i(t) \sim t^{\alpha}$ is broad with the simulated and measured $P(\alpha)$ being in excellent agreement. Our work shows that chromosomes structures are highly dynamic exhibiting great cell-to-cell variations. The rugged chromosome energy landscape with multiple minima separated by large barriers is perhaps needed to achieve a balance between genomic stability and dynamics for the execution of a variety of biological functions.

## Results

**Choosing the energy scale in the Chromosome Capture Model (CCM):** In the CCM (Fig.S1), the only unknown parameter is $\epsilon$, characterizing the strength of the interaction between the loci (Table I in the SI). We chose a $\epsilon$ value that produces the contact maps that is near quantitative agreement with the Hi-C data. As $\epsilon$ increases the structures of folded chromosome are arranged in such a way that segments with small $s$ are more likely to be in spatial proximity (Fig.4A) (see the discussion below). The snapshots of the folded structure, the good agreement between the calculated and Hi-C contact maps, and the accurate description of the spatial organization as assessed by the Ward Linkage Matrix (WLM) (see below and the SI) confirm that $\epsilon = 2.4k_{\mathrm{B}}T$ produces closest agreement with experiments. Furthermore, $P(s)$ as a function of $s$ obtained in simulations with $\epsilon = 2.4k_{\mathrm{B}}T$ is also consistent with experiments (see below). The $s$-dependent contact probability, $P(s)$ in Fig.1A, shows that there are two scaling regimes. As $\epsilon$ increases, the probability of short-range (small $s$) increases by several fold, while $P(s)$ for large $s$ decreases by approximately an order of magnitude (Fig.S12A). In particular, for $\epsilon = 1.0k_{\mathrm{B}}T$, $P(s)$ as a function of $s$, decreases much faster compared to experiment at small $s$. In contrast, we find that at $\epsilon = 2.4k_{\mathrm{B}}T$, $P(s) \sim s^{-0.75}$ for $s < 0.5$ Mbps and when $s$ exceeds

$\sim 0.5$ Mbps, $P(s) \sim s^{-1.25}$ (red curve in Fig.1A). Such a behavior, with $P(s)$ exhibiting two distinct scaling regimes, is in good agreement with experiments (black line in Fig.1A). It is worth pointing out that the two-scaling regimes in $P(s)$ is a robust feature of all 23 Human interphase chromosomes (Fig.S13).

**CCM accounts for compartments and Topologically associating Domains (TADs):** We fixed $N$, the size of the copolymer to $N = 10,000$, modeling a 12 Mbps (mega baseparis) chromatin fiber, corresponding to a selected region of the Human Cell line GM12878 Chromosome 5 (Chr 5) from 145.87 Mbps to 157.87 Mbps. Comparison of the contact maps between simulations and experiment illustrates compartment formation appearing as plaid or checkerboard patterns in Fig.1B, demonstrating good agreement with Hi-C data (2, 3). The dashed rectangles mark the border of one such compartment enriched predominantly with interactions between monomers of the same type. In order to make the comparison precise, we treated the contact maps as probabilistic matrices, and used a variety of mathematical methods to quantitatively compare large matrices. First, the checkerboard pattern in the contact map is more prominent when illustrated using the Spearman correlation map (see SI for details and Figs.S5, S6). Second, to quantitatively compare the simulated results with experiments, we use spectral co-clustering algorithm (25) to bi-cluster the computed Spearman correlation map (see SI for details). Finally, the similarity between the simulated and experimental data is assessed using the Adjusted Mutual Information Score (AMI) described in the SI. The CCM model, based only on epigenetic information, yields a AMI score that results in correctly reproducing $\approx 81\%$ of the compartments obtained from the experimental data. In contrast, a homopolymer model (SI), with an absolute AMI score that is 200 times smaller (Fig.S8), does not lead to the formation of compartments (correctly reproducing $\approx 51\%$ of the compartments, no better than the random assignments). Thus, the CCM is the minimal model needed to reproduce the essential features found in the contact map.

Our model also reproduces Topological Associating Domains (TADs), depicted as triangles of average length of 200 kbps along the diagonal of the contact map in which the interactions

between the loci are greatly enhanced (Fig.1C). It has been noted (3) that in a majority of cases, boundaries of the TADs are marked by a pair of CTCF motifs with high probability of interaction between them. They are visualized as peaks in the Hi-C map (Fig.1C). To quantitatively detect the boundary of TADs, we adopt the same procedure used in Ref. (1) to identify the position of each TAD. The boundaries of the TADs, shown in blue (Hi-C data) and green (simulations) are reproduced by the CCM (Figs.1C and 1D). The non-compact nature and irregular shapes of the TADs are illustrated schematically by typical snapshots for two TADs (Figs.1F and 1G).

**Active and repressive loci segregate:** The inset in Fig.2A, displaying a typical snapshot of the folded chromosome, reveals that active (A, green) and repressive (B, red) loci are clustered together, undergoing micro-phase separation (see Methods for definition of active and repressive loci). The tendency to segregate is vividly illustrated in the radial distribution functions $g_{AA}(r)$, $g_{BB}(r)$ and $g_{AB}(r)$, which shows (Fig.2A) that $g_{AA}(r)$ and $g_{BB}(r)$ have much higher values than $g_{AB}(r)$ implying that the probability of finding B near A is small. Such a micro-phase separation between the A-rich and B-rich regions directly gives rise to compartments in the contact map. Interestingly, the normalized radial density (Fig.2B) shows that active chromatin exhibits a peak at large radial distance, $r$ implying that the active loci localize on the periphery of the folded chromosome whereas repressive chromatin is more homogeneously distributed. Visual inspection of the simulation trajectories also suggests that active and repressive chromatins are often separated in a polarized fashion, in accord with a recent experimental study (4).

**Spatial organization of the folded chromosome:** In order to illustrate the spatial organization of the chromosome, we introduce the distance function,

$$R(s) = \left\langle \sum_{i<j}^{N} \frac{(\boldsymbol{r}_i - \boldsymbol{r}_j)^2 \delta(s - |i-j|)}{N-s} \right\rangle^{1/2} \tag{1}$$

where $\langle \cdot \rangle$ denotes both an ensemble and time average. We calculated $R(s)$, the mean end-to-end distance between the loci, by constraining the genomic distance $|i-j|$ to $s$. If the folded chromosome is maximally compact, we expect $R(s) \sim s^{1/3}$. However, the plot of $R(s)$ on a log-log scale shows that in the range $10^5 \lesssim s \lesssim 10^6$ bps, $R(s) \sim s^{0.2}$. The plateau at large $s$ arises due to $s$ reaching the boundary of the folded structure. The inset in Fig.3A, comparing

the simulation result and experimental data (4), both show the same scaling for $R(s)$.

By a systematic analysis of the Fluorescence in Situ Hybridization (FISH) data, Wang et al (4) established that the probability of contact formation, $P(s)$, is inversely proportional to a power of $R(s)$, with the latter providing a direct picture of spatial organization. Similarly, in this work we explored the relation between between $C_{ij}$ and $R_{ij}$ where $C_{ij}(= P_{ij}\sum_{i<j} C_{ij} \propto P_{ij})$ is the number of contacts between loci $i$ and $j$, and $R_{ij} = \langle(\boldsymbol{r}_i - \boldsymbol{r}_j)^2\rangle^{1/2}$ is the mean distance between them. The heat map of $(1/C_{ij}, R_{ij})$ in Fig.3B shows that the two matrices are proportional to each other. In accord with the FISH experiment (4), we find that $1/C_{ij} \propto R_{ij}^{\lambda}$ where $\lambda \approx 4$, suggesting that larger mean spatial distance between loci $i$ and $j$ implies smaller contact probability, which is the fundamental assumption when experimental Hi-C data is used to infer three dimensional chromosome organization. The decrease of $C_{ij}$ with increasing $R_{ij}$ with a large value of $\lambda$ is unexpected but is an important finding needed to link contact maps and spatial structures.

**Structural Variation between Cells:** The slope of the dashed line in Fig.3B obtained using the data in Ref. (4), is 4.1, which coincides with our simulation results. Mean field arguments (26) suggest that $P(s) \sim R(s)^{-3}$, which follows from the observation that the end of the chain is uniformly distributed over a volume $R^3(s)$. This is neither consistent with our simulations nor with experiments, implying that the distribution of the chain ends is greatly skewed. Although both the simulated and experimental results establish a strong correlation between $R(s)$ and $P(s)$, such a correlation is only valid in an averaged sense (over many loci and cell population) (see SI for additional discussions and also Ref. (27)). To quantitatively determine the cell-to-cell structural variations in the TADs, we calculated the distribution, $P(\langle R_g^2\rangle/\overline{\langle R_g^2\rangle})$, of the mean square radius of gyration $\langle R_g^2\rangle$ for the thirty-two Chr5 TADs in each trajectory normalized by the average $\overline{\langle R_g^2\rangle}$ of each individual TAD (Fig.3C). The bracket (bar) is the time (ensemble) average. The large dispersion in $P(\langle R_g^2\rangle/\overline{\langle R_g^2\rangle})$ (Fig. 3C) is consistent with recent FISH (28) and single-cell Hi-C experimental findings (29, 30) in which individual TAD compaction varies widely from an extended to a compact state depending on the cell. To decipher how the variation of the structure of the chromosome changes as a function of $s$, we

calculated the coefficient of variation, $\delta R(s) = (\langle R_s^2 \rangle - \langle R_s \rangle^2)^{1/2} / \langle R(s) \rangle$. Interestingly, $\delta R(s)$ first increases with $s$ up to $s \approx 10^5 \sim 10^6$ bps and then decreases as $s$ further increases (Fig.3D). Analysis of the experimental data from Ref. (4) shows a similar decreasing trend for $s > 10^5$bps (Fig.S11(C)). Higher resolution experiments are needed to resolve the variance for $s < 10^5$bps. The predicted non-monotonic dependence of $\delta R(s)$ on $s$ is amenable to experimental test.

**Chromosome Structures in terms of Ward Linkage Matrix (WLM):** To quantitatively analyze the spatial organization, we used the Ward Linkage Matrix (WLM) calculated from the distance matrix $\boldsymbol{R}$ in which the element, $R_{ij}$, is the distance between the loci $i$ and $j$ (see SI for details). We also constructed the experimental WLM by converting the Hi-C contact map to a distance map by exploiting the approximate relationship between $R_{ij}$ and $P_{ij}$ ($\propto R_{ij}^{-4.1}$) discussed previously (see Fig.3B). The advantages of using distance matrices instead of contact maps are two fold. First, matrix $\boldsymbol{R}$ is a direct depiction of the three dimensional organization of the chromosome. Because WLM, constructed from $\boldsymbol{R}$, is a cophenetic matrix, which can be used to reveal the hierarchical nature of the chromosome organization. Second, the contact map matrix elements do not obey triangle inequality. Therefore, it is not a good indicator of the actual 3D spatial arrangement of the loci. We show the WLM for the two $\epsilon$ values (Fig.4A) and the comparison between WLM computed based on experimental data and WLM for $\epsilon = 2.4k_{\mathrm{B}}T$ (Fig.4B). Visual inspection of the WLMs for $\epsilon = 2.4k_{\mathrm{B}}T$ shows distinct segregation in the spatial arrangement of the loci. It is clear from Fig.4B that the WLM, reconstructed from Hi-C data, and simulations result with $\epsilon = 2.4k_{\mathrm{B}}T$ are almost identical. Heatmaps of the WLMs show that for both $\epsilon = 1.0k_{\mathrm{B}}T$ and $\epsilon = 2.0k_{\mathrm{B}}T$ (Fig.S14), visual inspection suggests that loci with large genomic separation $s$ are in spatial proximity, which is inconsistent with the experimental WLM. The Pearson correlation coefficient between experimental result and CCM using $\epsilon = 2.4k_{\mathrm{B}}T$ is 0.96 (0.53 for $\epsilon = 1.0k_{\mathrm{B}}T$, 0.84 for $\epsilon = 2.0k_{\mathrm{B}}T$ and 0.75 for $\epsilon = 2.7k_{\mathrm{B}}T$). Thus, the poorer agreement between the simulated WLM (Fig.S14) as well as Spearman correlation matrix (Fig.S6) using $\epsilon = (1.0, 2.0, 2.7)k_{\mathrm{B}}T$ and experiments, compared to $\epsilon = 2.4k_{\mathrm{B}}T$, further justifies the latter as the optimum value in the CCM.

**Cell-to-cell variations in the WLM:** To assess the large structural variation between cells, we calculated the WLM for individual cells. We obtain the single cell WLM as a time average of trajectory generated for an individual cell. Fig.5 shows that there are dramatic variations between the WLM for individual cells with the ensemble average deviating greatly from the patterns in individual cells. Thus, it is likely that the chromosome structure is highly heterogenous. These findings are reflected in the broad width of the distribution of the Pearson Correlation between all pairs of cells. The mean of the distribution is $\bar{\rho} = 0.2$ small, implying little overlap in the WLMs between any two cells. Such large cell-to-cell variations in the structures, which are also revealed in recent experiments (30, 31), also have profound consequences on the chromosome dynamics (see below).

**Chromosome dynamics is glassy:** We probe the dynamics of the organized chromosome with $\epsilon = 2.4k_{\mathrm{B}}T$. We first computed the incoherent scattering function, $F_s(k,t) = (1/N)\left\langle \sum_{j=1}^{N} \mathrm{e}^{i\boldsymbol{k}(\boldsymbol{r}_j(t)-\boldsymbol{r}_j(0))} \right\rangle$ where $\boldsymbol{r}_j(t)$ is the position of $j^{th}$ loci at time $t$. The decay of $F_s(k,t)$ (orange line in Fig.6A) for $k \sim 1/r_s$ ($r_s$ is the position of the first peak in the radial distribution function ($g_{AA}(r)$ and $g_{BB}(r)$) (Fig.2A)) is best fit using the stretched exponential function, $F_s(k,t) \sim e^{-(t/\tau_\alpha)^\beta}$ with a small stretching coefficient, $\beta \approx 0.27$. The stretched exponential decay with small $\beta$ is one hallmark of glassy dynamics. For comparison, $F_s(k,t)$ decays exponentially for $\epsilon = 1.0k_{\mathrm{B}}T$, implying liquid-like dynamics (blue line in Fig.6A).

In the context of relaxation in supercooled liquids, it has been shown that the fourth order susceptibility (32), $\chi_4(k,t) = N[\langle F_s(k,t)^2\rangle - \langle F_s(k,t)\rangle^2]$ provides a unique way of distinguishing between fluctuations in the liquid and frozen states. As in structural glasses, the value of $\chi_4(k,t)$ increases with $t$ reaching peak at $t = t_M$ and decays at longer times. The peak in the $\chi_4(k,t)$ is an indication of dynamic heterogeneity, which in the chromosome is manifested as dramatic variations in the loci dynamics (see below). For $\epsilon = 2.4k_{\mathrm{B}}T$, $\chi_4(k,t)$ reaches a maximum at $t_M \approx 1s$ (Fig.6B) which surprisingly, is the same order of magnitude ($\sim 5s$) in which chromatin movement was found to be coherent on a length scale of $1\mu$m (23). The dynamics in $F_s(k,t)$ and $\chi_4(k,t)$ together show that Human Interphase chromosome dynamics is glassy (20), and highly

heterogeneous.

**Single loci Mean Square Displacements are heterogeneous:** In order to ascertain the consequences of glassy dynamics at the microscopic level, we plot the MSD, $\Delta(t) = \frac{1}{N}\langle\sum_{j=1}^{N}(\boldsymbol{r}_j(t) - \boldsymbol{r}_j(0))^2\rangle$ in Fig.7, from which we draw a few conclusions.

1. Because of the polymeric nature of the chromosome, the maximum excursion in $\Delta(t \to \infty) = 2R_g^2$, where $R_g \approx 0.7\mu$m is the radius of gyration of Chr 5. Consequently, for both $\epsilon = 1.0k_{\mathrm{B}}T$ (red) and $\epsilon = 2.4k_{\mathrm{B}}T$, $\Delta(t)$ in the long time limit is smaller than $2R_g^2$. For $\epsilon = 2.4k_{\mathrm{B}}T$ (green), $\Delta(t)$ shows a crossover at $t \approx 10^{-2}s$ from slow to a faster diffusion, another indicator of glassy dynamics (33). The slow diffusion is due to caging by neighboring loci, which is similar to what is typically observed in glasses. The plateau in $\Delta(t)$ (Fig.7A) is not pronounced, suggesting that the folded chromosome is likely on the edge of glassiness. The crossover is more prominent in the time-dependence of the mean squared displacement of single loci (see below).

2. The two dashed lines in Fig.7A show $\Delta(t) \sim t^{\alpha}$ with $\alpha = 0.45$. The value of $s$ is close to 0.5 for the condensed polymer, which can be understood using the following arguments. The total friction coefficient experienced by the whole chain is the sum of contributions from each of the $N$ monomers, $\xi_T = N\xi$. The time for the chain to move a distance $\approx R_g$ is $\tau_R = R_g^2/D_R \sim N^{2\nu+1}$. Let us assume that the diffusion of each monomer scales as $Dt^{\alpha}$. If each monomer moves a distance on the order of $R_g$ then the chain as a whole will diffuse by $R_g$. Thus, by equating $D\tau_R^{\alpha} \sim R_g^2$, we get $\alpha = 2\nu/(2\nu+1)$. For an ideal chain $\nu = 0.5$, which recovers the prediction by Rouse model, $\alpha = 0.5$. For a self-avoiding chain, $\nu \approx 0.6$, we get $\alpha \approx 0.54$. For a condensed chain, $\nu = 1/3$, we get $\alpha = 0.4$, thus rationalizing the findings in the simulations. Surprisingly, $\alpha = 0.45$ found in simulations is in good agreement with recent experimental findings (24). We also obtained the similar result using a very different chromosome model (not published yet).

3. We also calculated the diffusion of a single locus (sMSD), which for the $i^{th}$ locus is

$\Delta_i(t) = \langle(\boldsymbol{r}_i(t_0+t) - \boldsymbol{r}_i(t_0))^2\rangle_{t_0}$ where $\langle\cdot\rangle_{t_0}$ is the average over the initial time $t_0$. Distinct differences are found between the polymer exhibiting liquid-like and glassy-like dynamics. The variance in single loci MSD is large for $\epsilon = 2.4k_{\mathrm{B}}T$, illustrated in Fig.7B, which shows 10 typical trajectories for $\epsilon = 1.0k_{\mathrm{B}}T$ and $\epsilon = 2.4k_{\mathrm{B}}T$ each. For glassy dynamics, we found that the loci exhibiting high and low mobilities coexist in the chromosome, with orders of magnitude difference in the values of the effective diffusion coefficients, obtained by fitting $\Delta_i(t) = D_\alpha t^{\alpha_i}$. Caging effects are also evident on the timescale as long as seconds. Some loci are found to exhibit caging-hopping diffusion, which is a hallmark in glass forming systems (34, 35). Interestingly, such caging-hopping process has been observed in Human cell sometime ago (36).

4. The large variance in sMSD has been found in the motion of chromatin loci in both *E.coli* and Human cells (37–41). To further quantify heterogeneities in the loci mobilities, we calculated the Van Hove function $P(\Delta x)$, $P(\Delta x|\Delta t) = \langle(1/N)\sum_{i=1}^{N}\delta(\Delta x - [x_i(\Delta t) - x_i(0)])\rangle$. Fig.7C and 7D show the $P(\Delta x|\Delta t)$ and normalized $P(\Delta x/\sigma|\Delta t)$ for $\epsilon = 2.4k_{\mathrm{B}}T$ at different lag times $\Delta t$. For $\epsilon = 1.0k_{\mathrm{B}}T$, Van Hove function is well fit by a Gaussian at different lag times $\Delta t$ (Fig.S15). In contrast, for chromosome with glassy dynamics, all the $P(\Delta x|\Delta t)$ exhibit fat tail, which are well fit by an exponential function at large values of $\Delta x$ (Fig.7C, D) at all $\delta t$ values, suggestive of the existence of fast and slow loci (35).

5. The results in Fig.7 allow us to make direct comparisons with experimental data to establish signatures of dynamic heterogeneity. We calculated the distribution of effective diffusion exponent $\alpha_i$, $P(\alpha)$, where $\alpha_i$ is obtained by fitting the sMSD to $\sim t^{\alpha_i}$ within some lag time range. Fig.7E shows that $P(\alpha)$ calculated from simulations is in good agreement with experiments (42) in the same lag time range (0.42s $< \Delta t <$ 10s). The $P(\alpha)$ distribution for small lag time ($10^{-6}$s $< \Delta t <$ 0.42s) shows two prominent peaks, further validating the picture of coexisting fast and slow loci. The good agreement between the predictions of the CCM with data, showing large variance of mobilities among individual loci *in vivo*, further supports our conclusion that organized chromosome dynamics

is glassy.

**Active loci has higher mobility:** Fig.7F shows MSD for loci of active type and repressive type. For $\epsilon = 1.0k_{\mathrm{B}}T$, there is no difference between active and repressive loci in their mobilities. However, in the glassy state active loci diffuses faster than the repressive loci. The ratio between the effective diffusion coefficients (the slope of dashed line) of the active and repressive loci is $0.0116/0.008 \simeq 1.45$, in good agreement with experimental estimate $0.018/0.013 \simeq 1.38$ (24). Such a difference is surprising since the parameters charactering the A-A and B-B interactions are identical. Because active loci are predominantly localized on the periphery of the chromosome whereas repressive loci are confined to the interior of the folded chromosome the latter is more caged by the neighboring loci, thus restricting their mobility. Note that in the experimental result (24) comparison is made between the loci in the periphery and interior of the nucleus. It is well known that the nucleus periphery is enriched with heterochromatin (repressive loci) and interior is enriched with euchromatin (active loci).

## Discussion

**Transferability of the CCM:** In order to demonstrate the transferability of the CCM we simulated Chr 10 using exactly the same parameters as for Chr 5. Fig.S16 compares the WLM obtained from simulations for different $\epsilon$ values and the computed WLM using the Hi-C contact map. The contact map is translated to the distance $R_{ij}$ by assuming that $P_{ij} \propto R_{ij}^{-4.1}$ holds for Chr 10 as well. It is evident that the CCM nearly quantitatively reproduces the spatial organization of Chr 10 (Fig.S16). Thus, it appears that the CCM could be used for simulating the structures and dynamics of other chromosomes as well.

**Scale-dependent folding of chromosomes:** In order to reveal how chromosome folds and to link these processes to the experimentally measurable $P(s)$, we calculated the time dependent change in $P(s)$ as a function of $t$. At scales less (above) than $s^* \approx 5 \times 10^5$bps, $P(s)$ decreases (increases) as the chromosome folds. The $P(s) \sim s^{-0.75}$ scaling for $s < s^*$ (see also Fig.1A) is the result of folding on small genomic scale during the early stage of chromosome folding (Fig.8A). In

the initial stages folding starts by forming $\approx s^*$ sized chromosome droplets (CDs) as illustrated in Fig.8A. In the second scaling regime, $P(s) \sim s^{-1.25}$, global folding occurs by coalescence of the CDs (Fig.8A). Thus, our CCM model, which suggests a hierarchical chromosome organization on two distinct scales, also explains the two scaling in $P(s)$.

The hierarchical nature of structural organization is further illustrated using $A(s)$, the number of contacts that a $s$-sized subchain forms with the rest of the chromosome. For a compact structure, $A(s) \sim s^{2/3}$ and $A(s) \sim s$ for an ideal chain. Fig.S17 shows that $A(s)$ computed using the Hi-C data (black square line) varies as $s^{2/3}$, suggesting that chromosome is compact on all length scales. We also find that upon increasing $\epsilon$, the range of $A(s) \sim s^{2/3}$ expands. The pictorial view of folding (Fig.8A) and the $A(s)$ scaling show that chromosome folding occurs hierarchically with the formation of CDs and subsequent growth of the large CDs at the expense of smaller ones. We quantitively monitored the growth of CDs during the folding and found that the size of CD grows linearly with time at the intermediate folding stage (Fig.8B). Such a condensation process is reminiscent of the Lifshitz-Slazov mechanism (43) used to describe Ostwald ripening.

**Conclusions**

We developed the Chromosome Copolymer Model (CCM), a self-avoiding polymer with two epigenetic states and with fixed loop anchors whose locations are obtained from experiment to describe chromosome dynamics. The use of rigorous clustering techniques allows to demonstrate that the CCM nearly quantitively reproduces Hi-C contact maps and spatial organization gleaned from super-resolution imaging experiments. It should be borne in mind that contact maps are probabilistic matrices that are low dimensional representation of the three dimensional organization of genomes. Consequently, many distinct copolymer models are likely to reproduce the probability maps encoded in the Hi-C data. In other words, solving the inverse problem of going from contact maps to a energy function is not unique (see Ref. (6))

Chromosome dynamics is glassy implying that the free energy landscape has multiple equiv-

alent minima. Consequently, it is likely that in genomes only the probability of realizing these minima is meaningful, which is the case in structural glasses. The presence of multiple minima also leads to cell-cell heterogeneity with each cell exploring different local minimum in the free energy landscape. We speculate that the glass-like landscape might also be beneficial in chromosome functions because only a certain minimum needs to be accessed to carry out a specific function, which will minimize large scale structural fluctuations. In this sense, chromosome glassiness provides a balance between genomic stability and mobility.

## Methods

**Construction of the model:** Contact maps (2, 3) of interphase chromosomes show that they are partitioned into genome-wide compartments, displaying plaid (checkerboard) patterns. If two loci belong to the same compartment they have higher probability to be in contact than if they are in different compartments. Although finer classifications are possible, compartments (2) can be categorized broadly into two (open (A) and closed (B)) classes associated with distinct histone markers. Open compartment is enriched with transcriptional activity-related histone markers, such as H3K36me3, whereas the closed compartment is enriched with repressive histone markers, such as H3K27me3. Chromatin segments with repressive histone markers have effective attractive interactions, which models HP1 protein-regulated interactions between heterochromatin regions (44, 45). We assume that chromatin fiber, with active histone markers, also has such a similar attraction. From these considerations, it follows that the minimal model for Human chromosome should be a copolymer where the two types of monomers represent open and closed chromatin states. To account for the two states, we introduce the Chromosome Copolymer Model (CCM) as a self-avoiding polymer with two kinds of beads. Similar genre of models have been proposed in several recent studies (15–18, 21) to successfully decipher the organization of genomes.

The energy function in the CCM is, $E = U_{\mathrm{C}} + U_{\mathrm{LJ}}$, where $U_{\mathrm{C}}$ contains bond potential ($U^{\mathrm{S}}$) and loop interaction ($U^{\mathrm{L}}$), and $U_{\mathrm{LJ}}$ is the Lennard-Jones pairwise interaction between the monomers.

If two monomers belong to type A (B), the interaction strength is $\epsilon_{AA}(\epsilon_{BB})$. The interaction strength between A and B is $\epsilon_{AB}$. Each monomer represents 1,200 base pairs (bps), with six nucleosomes connected by six linker DNA segments. The size of each monomer, $\sigma$, is estimated by considering two limiting cases. If we assume that nucleosomes are compact then the value of $\sigma$ may be obtained by equating the volume of the monomer to $6v$ where $v$ is the volume of a single nucleosome. This leads to $\sigma \approx 6^{1/3} R_{\mathrm{N}} \approx 20$ nm where $R_{\mathrm{N}} \approx 10$ nm is the size of each nucleosome (46). Another limiting case may be considered by treating the six nucleosome array as a worm-like chain. The persistence length of the chromatin fiber is estimated to be $\sim$ 1,000 bps (47), which is about the size of one monomer. The mean end-to-end distance of a worm like chain whose persistence length is comparable to the contour length $L$ is $R \approx L\sqrt{2/e}$. The value of $L$ for a six nucleosome array is $6(16.5 + R_{\mathrm{N}})$nm where the length of a single linker DNA is 16.5nm. This gives us the upperbound of $\sigma$ to be 130nm. Thus, the two limiting values of $\sigma$ are 20 nm and 130 nm. We assume that the value of $\sigma$ is an approximate mean, yielding $\sigma = 70$ nm.

The type of monomer is determined using the Broad ChromHMM track (48). There are totally 15 chromatin states, out of which the first eleven are related to gene activity. Thus, we consider state 1 to state 11 as a single active state (A), and states 12-15 as a single repressive state (B). Details of the assignment are given in the Supplementary Information (SI).

**Simulations:** As detailed in the SI, we performed simulations using both Langevin Dynamics (low friction) and Brownian Dynamics (high friction) using a custom modified version of the molecular dynamics package LAMMPS. The use of Langevin Dynamics accelerates the sampling of the conformational space (49), needed for reliable computation of static properties. Realistic value of the friction coefficient is used in Brownian Dynamics simulations to investigate chromosome dynamics, thus allowing us to make direct comparisons with experiments.

We varied the values of $\epsilon_{AA}$, $\epsilon_{BB}$ and $\epsilon_{AB}$ to investigate the effect of interaction strength on the simulation results. For simplicity, we set $\epsilon_{AA} = \epsilon_{BB} \equiv \epsilon$. By fixing the ratio $\epsilon/\epsilon_{AB}$ to 11/9, $\epsilon$ is the only relevant energy scale in the CCM. The results presented in the main text are

obtained with $\epsilon = 2.4k_{\mathrm{B}}T$ unless stated otherwise.

**Acknowledgements:** We want to thank Abdul N Malmi-Kakkada, Xin Li for discussions and comments on the manuscript. We are grateful to the National Science Foundation (CHE 16-36424) and the Collie-Welch Chair (F-0019) for supporting this work.

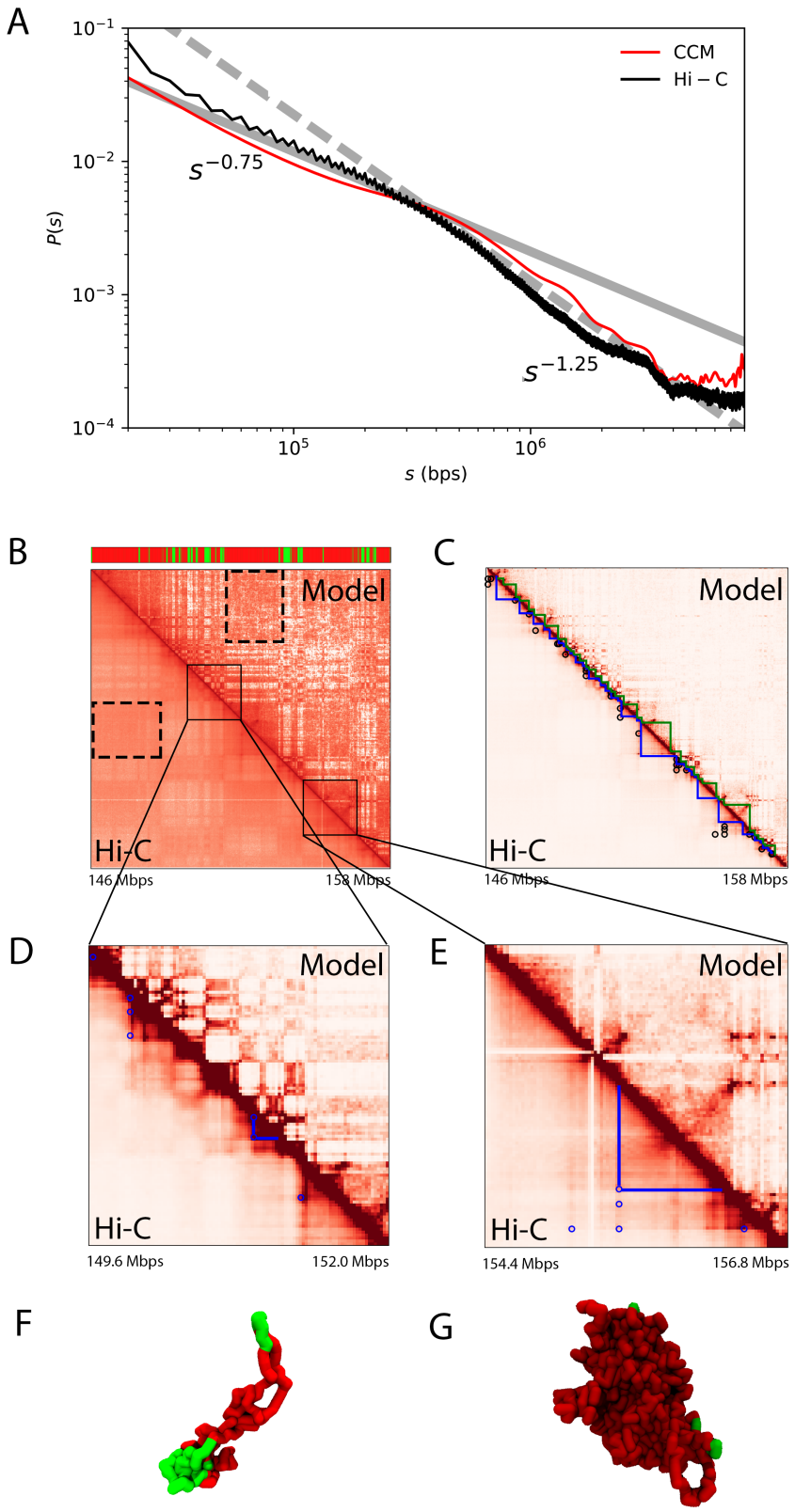
A
CCM
Hi – C
$s^{-0.75}$
$s^{-1.25}$
P(s)
s (bps)
B
Model
Hi-C
146 Mbps
158 Mbps
C
Model
Hi-C
146 Mbps
158 Mbps
D
Model
Hi-C
149.6 Mbps
152.0 Mbps
E
Model
Hi-C
154.4 Mbps
156.8 Mbps
F
G

FIG. 1: **(A)** Direct comparison between experimental data (3) (black) and simulated $P(s)$. Dashed and solid lines are plots of $s^{-1.25}$ and $s^{-0.75}$, respectively. The crossover point between the two scaling regimes at $s \sim 3 \cdot 10^5$bps is noticeable both in the experimental and simulated results. **(B)** Comparison of the contact maps inferred from Hi-C experiment (3) (lower triangle) and obtained from simulations (upper triangle) results. For easier visualization, the values of the contact probability are converted to a $\log_2$ scale. The bar above the map marks the epigenetic states with green (red) representing active (repressive) loci. The dashed black box is an example of a compartment. Such compartment-like structures emerge due to contacts between loci separated by large genomic distances giving rise to spatial order in the folded chromosome. **(C)** Illustration of Topologically Associated Domains (TADs). The blue and green triangles are from experiments and simulations, respectively. The black circles mark the positions of loops detected from experiment data, which are formed by two CTCF motifs. **(D)** The zoom in of the diagonal region for the chromosome segment between 149.6 Mbps to 152.0 Mbps. The blue circle marks the positions of CTCF loops found in experiment (3). **(E)** Same as **(D)** except for 154.4 Mbps to 156.8 Mbps. **(F)** and **(G)**. Snapshots of two TADs, marked by blue triangle in **(D)** and **(E)**, respectively. Green (Red) represent euchromatin (heterochromatin).

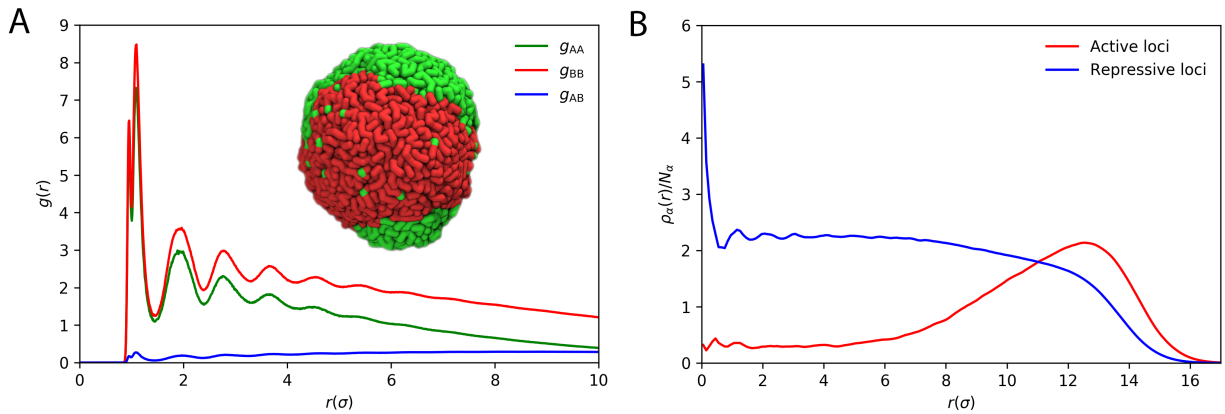


FIG. 2: **(A)** Radial distribution functions, $g(r)$, as a function of $r$ (in unit of $\sigma$) between active and repressive loci. Inset shows the typical conformation of the folded chromosome. Green and red segments correspond to active and repressive loci, respectively. The folded structure reveals vividly micro phase separation between active and repressive loci. **(B)** The normalized radial density, $\rho_\alpha(r)/N_\alpha = \langle N_\alpha(r)\rangle/(4\pi r^2 \Delta r N_\alpha)$, where $N_\alpha(r)$ is the number of loci of given type $\alpha$ found in the spherical shell between $r$ and $r + \Delta r$, $N_\alpha$ is the total number of loci of that type. The bracket $\langle\cdot\rangle$ is the ensemble average; $\rho_\alpha(r)/N_\alpha$ shows that the active loci are predominantly localized on the periphery of the folded chromosome. The repressive loci are more uniformly distributed.

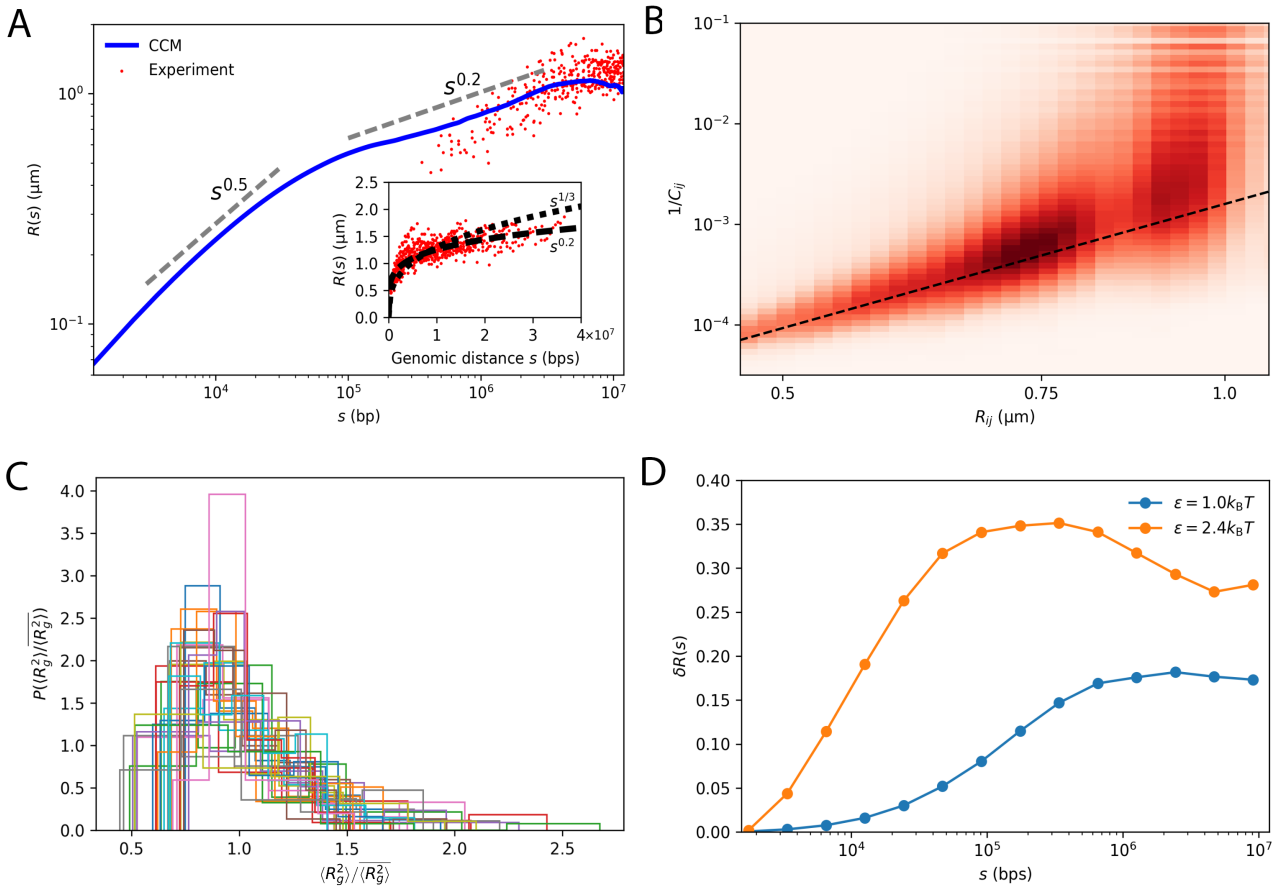
A
CCM
Experiment
$s^{0.2}$
$s^{0.5}$
$R(s)$ (μm)
$s$ (bp)
$s^{1/3}$
$s^{0.2}$
Genomic distance $s$ (bps)
B
$1/C_{ij}$
$R_{ij}$ (μm)
C
$P(\langle R_g^2\rangle/\overline{\langle R_g^2\rangle})$
$\langle R_g^2\rangle/\overline{\langle R_g^2\rangle}$
D
$\varepsilon = 1.0k_BT$
$\varepsilon = 2.4k_BT$
$\delta R(s)$
$s$ (bps)

FIG. 3: Organization and fluctuations of the chromosome structures. **(A)** The dependence of the spatial distance $R(s)$ (Eq.1) on the genomic distance, $s$. Grey dashed lines, indicating the slopes, are guides to the eye. The red dots are experimental data taken from Ref. (4) for $s < 1.2\times10^7$bps. The inset shows the complete set of experiment data. Short dashed and long dashed lines are $s^{1/3}$ and $s^{0.2}$, respectively. At small $s$ ($s < 10^5$bps), $R(s) \sim s^{0.5}$ implying that chromatin behaves as almost an ideal chain. **(B)** The heatmap of the 2D histogram of $(R_{ij}, 1/C_{ij})$. Dashed black line is the curve with scaling exponent 4.1, which coincides with the value obtained by fitting the experimental data (4). **(C)** Distribution $P(\langle R_g^2\rangle/\overline{\langle R_g^2\rangle})$, where $\langle R_g^2\rangle$ is the time average value of squared radius of gyration of a single trajectory and $\overline{\langle R_g^2\rangle}$ is the mean value averaged over all independent trajectories. Different colors represent $P(\langle R_g^2\rangle/\overline{\langle R_g^2\rangle})$ for the thirty two individual TADs. The distribution is surprisingly wide which suggests that TAD structures vary from cell-to-cell. **(D)** Coefficient of variation $\delta R(s) = (\langle R^2(s)\rangle - \langle R(s)\rangle^2)^{1/2}/\langle R(s)\rangle$, computed from simulations, shows a non-monotonic dependence on $s$ for $\epsilon = 2.4k_{\mathrm{B}}T$, increasing till $s \sim 10^5$bps and decreases at larger $s$.

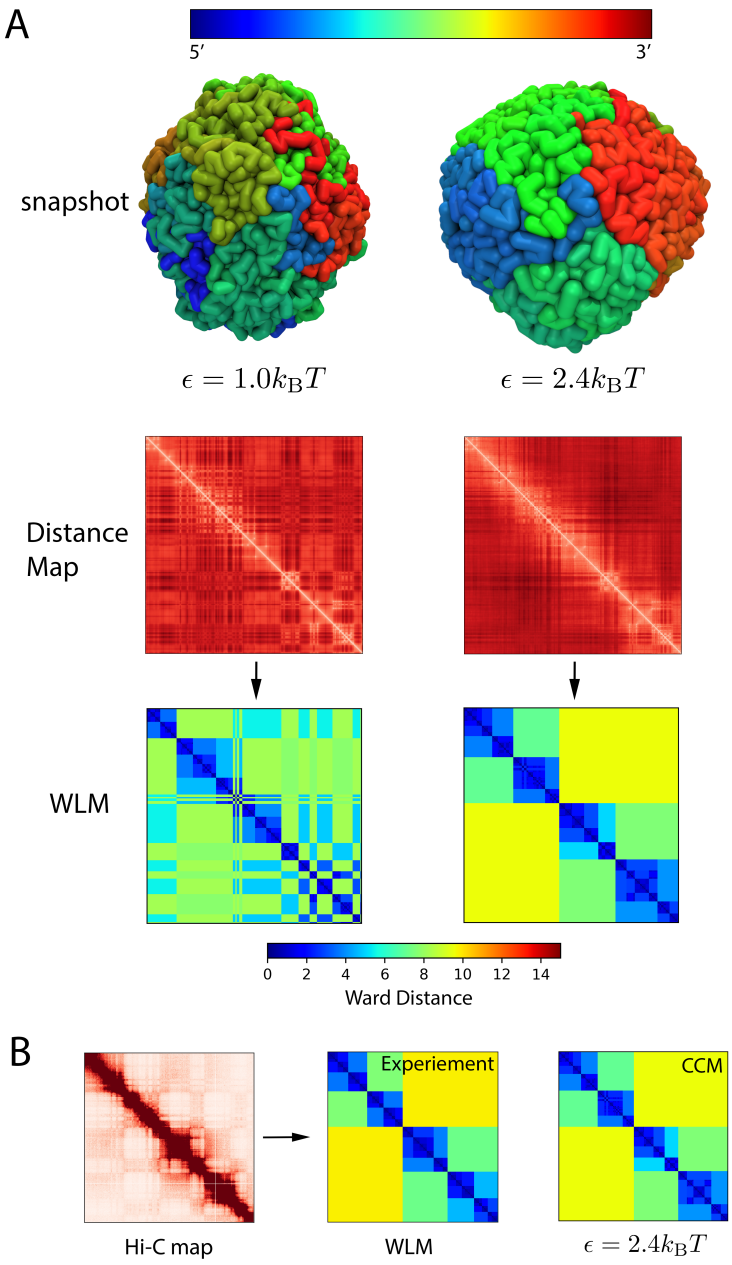
A
5′
3′
snapshot
$\epsilon = 1.0k_{\mathrm{B}}T$
$\epsilon = 2.4k_{\mathrm{B}}T$
Distance Map
WLM
0 2 4 6 8 10 12 14
Ward Distance
B
Hi-C map
Experiement
WLM
CCM
$\epsilon = 2.4k_{\mathrm{B}}T$

FIG. 4: **(A)** Typical conformations of the folded chromosome for $\epsilon = 1.0k_{\mathrm{B}}T$ (left) and $2.4k_{\mathrm{B}}T$ (right). The panel below shows the distance maps (red) and the Ward Linkage Matrix. The colorbar at bottom indicates the value of Ward distance (see SI) **(B)** Comparison of the ensemble average Ward Linkage Matrix (WLM) calculated using the experimental and simulated ($\epsilon = 2.4k_{\mathrm{B}}T$) contact maps.

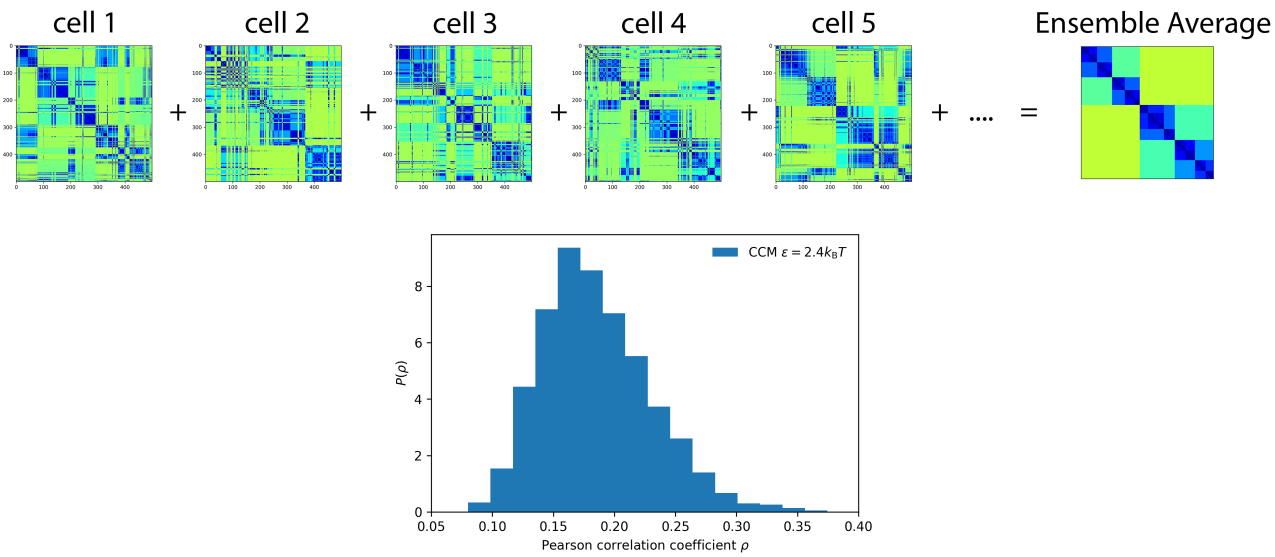


FIG. 5: Structural heterogeneity of the folded chromosome. The top figures show the Ward Linkage Matrices of different individual cells. The single cell WLM is the time average result over a single trajectory. The ensemble average WLM (right most) and the experimental WLM are in clear quantitative agreement (Fig.4B). However, the spatial organization varies from cell to cell. Each cell have very different WLMs, implying their structures are distinct. To quantify such heterogeneity, we calculated the Pearson correlation coefficient, $\rho$, between WLMs of any two cells. The bottom figure shows the distribution of $\rho$, $P(\rho)$, with a mean $\bar{\rho} = 0.2$. The distribution with low $\rho$ values is a further demonstration of structural heterogeneity in individual cells.

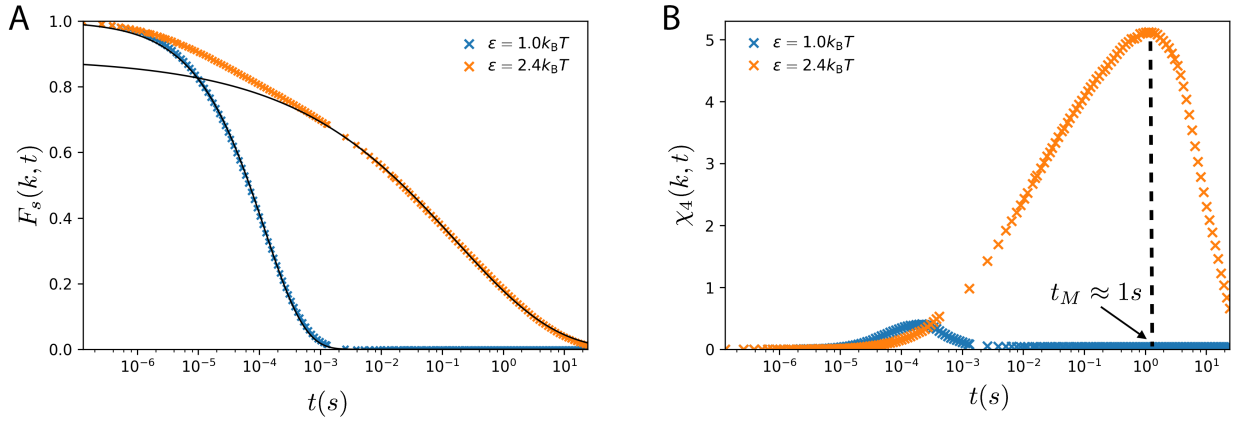


FIG. 6: **(A)** Intermediate scattering function obtained for $\epsilon = 1.0k_{\mathrm{B}}T$ (blue) and $\epsilon = 2.4k_{\mathrm{B}}T$ (orange). The line shows an exponential function fit, $F_s(k,t)$, for $\epsilon = 1.0k_{\mathrm{B}}T$. For $\epsilon = 2.4k_{\mathrm{B}}T$, $F_s(k,t) \sim e^{-(t/t_\alpha)^\beta}$ with $\beta = 0.27$, for $t$ exceeding a few milliseconds (black curve). **(B)** The fourth order susceptibility, $\chi_4(t)$, used as a function to demonstrate dynamic heterogeneity. The peak in $\chi_4(t)$ for $\epsilon = 2.4k_{\mathrm{B}}T$ around $t_M \approx 1$s is a signature of heterogeneity.

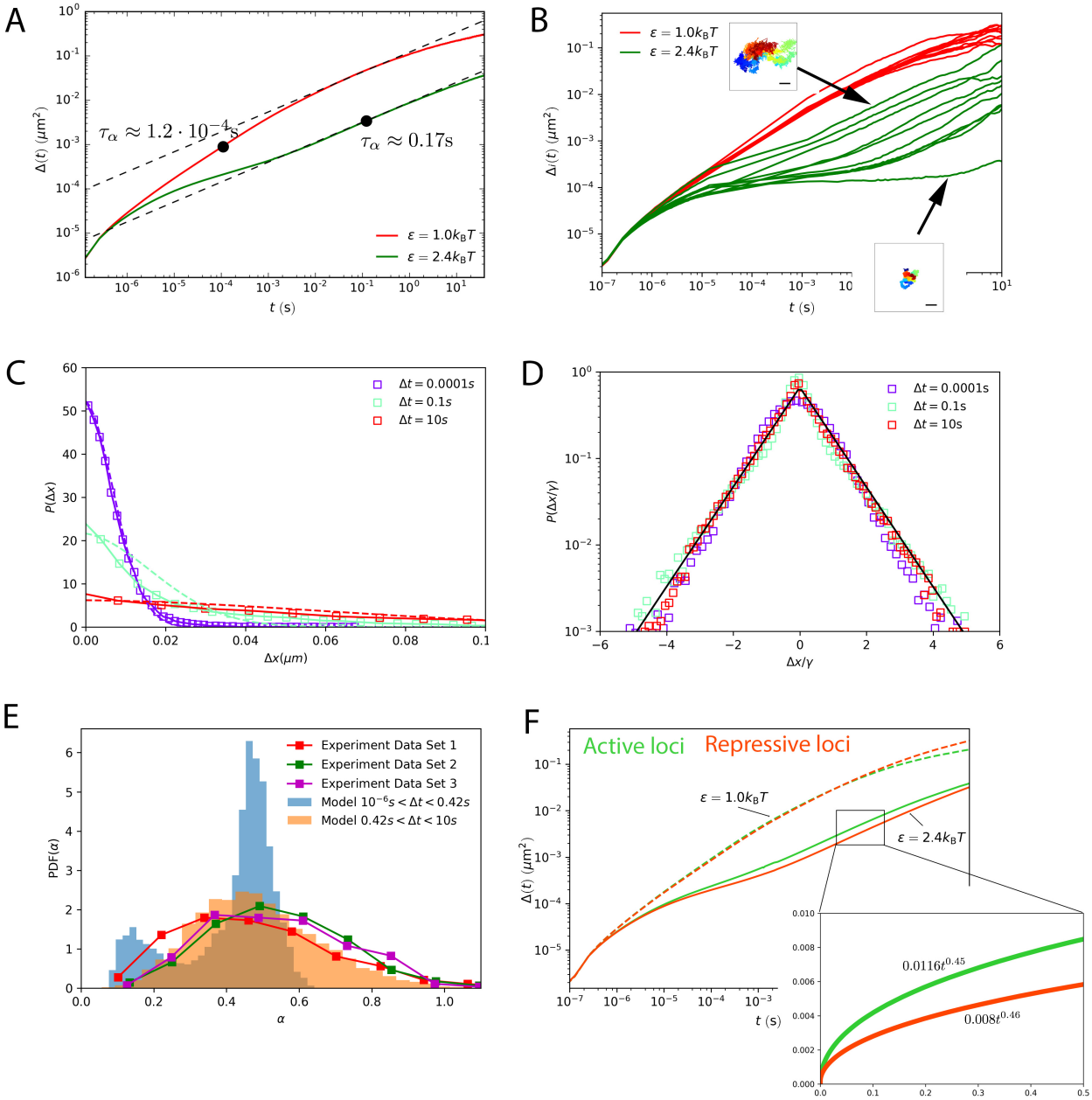
A
$\tau_\alpha \approx 1.2 \cdot 10^{-4}$s
$\tau_\alpha \approx 0.17$s
$\epsilon = 1.0k_BT$
$\epsilon = 2.4k_BT$
$\Delta(t)$ ($\mu m^2$)
$t$ (s)
B
$\epsilon = 1.0k_BT$
$\epsilon = 2.4k_BT$
$\Delta(t)$ ($\mu m^2$)
$t$ (s)
C
$\Delta t = 0.0001s$
$\Delta t = 0.1s$
$\Delta t = 10s$
$P(\Delta x)$
$\Delta x(\mu m)$
D
$\Delta t = 0.0001s$
$\Delta t = 0.1s$
$\Delta t = 10s$
$P(\Delta x/\gamma)$
$\Delta x/\gamma$
E
Experiment Data Set 1
Experiment Data Set 2
Experiment Data Set 3
Model $10^{-6}s < \Delta t < 0.42s$
Model $0.42s < \Delta t < 10s$
PDF($\alpha$)
$\alpha$
F
Active loci
Repressive loci
$\epsilon = 1.0k_BT$
$\epsilon = 2.4k_BT$
$\Delta(t)$ ($\mu m^2$)
$t$ (s)
$0.0116t^{0.45}$
$0.008t^{0.46}$

FIG. 7: (top) **(A)** Mean Square Displacement, $\Delta(t)$, as a function of time, $t$. The effective diffusion coefficients, D, computed from the fitted dashed lines are $0.122\mu\mathrm{m}^2/\mathrm{t}^{0.45}$and $0.009\mu\mathrm{m}^2/\mathrm{t}^{0.46}$ for $\epsilon = 1.0k_\mathrm{B}T$ and $\epsilon = 2.4k_\mathrm{B}T$, respectively. In the state exhibiting liquid-like behavior, the diffusion coefficient is more than one order magnitude larger than in the glassy state. Black dot marks the relaxation time $\tau_\alpha$ defined as $F_s(k, \tau_\alpha) = 1/e$ where $e \approx 2.718$. **(B)** Time dependence of 10 single loci MSD (sMSD, $\Delta_i(t)$) corresponding to $1^{st}, 100^{th}, ..., 10,000^{th}$ loci for $\epsilon = 1.0k_\mathrm{B}T$ and $\epsilon = 2.4k_\mathrm{B}T$. The insets show $\Delta_i(t)$ for two trajectories for fast (top) and slow (bottom) loci. Blue (red) indicates short (long) lag times. The scale bar corresponds to $1\mu m$. Caging effect can be clearly observed as plateau in $\Delta_i(t)$. (**C**) The Van Hove function $P(\Delta x)$ for $\epsilon = 2.4k_\mathrm{B}T$ at lag times $\Delta t = 0.0001$s, $\Delta t = 0.1$s and $\Delta t = 10$s. $P(\Delta x)$ has heavy tail at large $\Delta x$ and cannot be fit by a Gaussian (color dashed lines) except at $\Delta t = 0.0001$s at small $\Delta x$. **(D)** Same as **(C)** except displacement $\Delta x$ is normalized by the standard deviation $\gamma$. $P(\Delta x/\gamma)$ for different lag times collapse onto a master curve. Black line is an exponential fit, $\sim e^{-\eta(\Delta x/\gamma)}$ with $\eta \approx 1.3$. **(E)** Plot of Mean Square Displacement for active loci and repressive loci. The equation shown in the inset is the fit using formula $Dt^\alpha$, where $D$ is the diffusion coefficient and $\alpha$ is the diffusion exponent. (**F**) Normalized distribution, $P(\alpha)$, of the effective diffusion exponent $\alpha$. Comparison to experimental data (42) are also shown. The values of $\alpha$ are extracted from single loci trajectories by fitting sMSD, $\Delta_i(t) \sim t^\alpha$. The lag time range $0.42\mathrm{s} < \Delta t < 10\mathrm{s}$ is in the approximate same range used in the experiment. Experimental data set 1, 2, 3 are from Fig.2b, 2c and Fig.S5 of Ref. (42), respectively. The results from our simulation (orange) agree well with experimental data, shown as orange. The blue bar plot is $P(\alpha)$ for small lag times $10^{-6}\mathrm{s} < \Delta t < 0.42\mathrm{s}$. It shows two peaks, indicating the coexistence of two populations of loci with distinct mobilities.

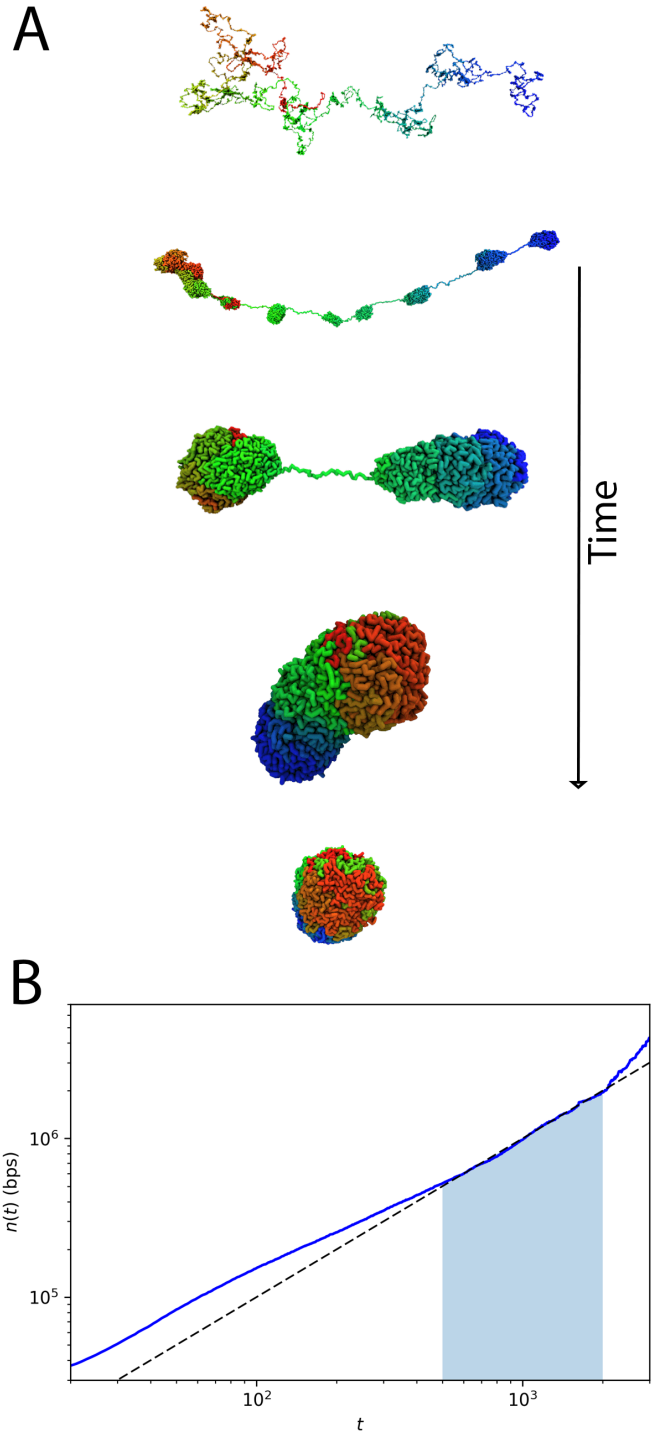
A
Time
B
$10^6$
$10^5$
$n(t)$ (bps)
$10^2$
$10^3$
$t$

FIG. 8: **(A)** Typical conformations during the chromosome folding process. Folding begins by formation of the chromosome droplets (CDs) connected by "tension strings". Typical CD size shown in the second subplot is about $s \approx 3 \cdot 10^5 \sim 6 \cdot 10^5$bps, being consistent with the approximate value of $s^*$. CDs merge to form larger clusters. In the final stage the two largest clusters condense to form the folded chromosome. **(B)** The growth of Chromosome droplets (CD), $n(t)$, as a function of $t$ during chromosome folding. The dashed line is fit in the region $5 \cdot 10^2$ to $2 \cdot 10^3$ timesteps with $n(t) \sim t^1$. $n(t)$ is the average number of basepairs in a CD at time $t$. The roughly linear increase of $n(t)$ with $t$ is consistent with Lifshitz-Slazov growth mechanism (43).

[1] Dixon JR, et al. (2012) Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* 485(7398):376–380.

[2] Lieberman-Aiden E, et al. (2009) Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* 326(5950):289–93.

[3] Rao SS, et al. (2014) A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* 159(7):1665–1680.

[4] Wang S, et al. (2016) Spatial organization of chromatin domains and compartments in single chromosomes. *Science* 353(6299):598–602.

[5] Boettiger AN, et al. (2016) Super-resolution imaging reveals distinct chromatin folding for different epigenetic states. *Nature* 529(7586):418–422.

[6] Bryngelson J, Thirumalai D (1996) Internal constraints induce localization in an isolated polymer molecule. *Phys. Rev. Lett.* 76(3):542.

[7] van den Engh G, Sachs R, Trask BJ (1992) Estimating genomic distance from DNA sequence location in cell nuclei by a random walk model. *Science* 257(5075):1410–1412.

[8] Sachs RK, van den Engh G, Trask B, Yokota H, Hearst JE (1995) A random-walk/giant-loop model for interphase chromosomes. *Proc. Natl. Acad. Sci. U. S. A.* 92(March):2710–2714.

[9] Marko JF, Siggia ED (1997) Polymer models of meiotic and mitotic chromosomes. *Mol. Biol. Cell* 8(11):2217–2231.

[10] Münkel C, Langowski J (1998) Chromosome structure predicted by a polymer model. *Phy. Rev. E* 57(5):5888.

[11] Rosa A, Everaers R (2008) Structure and dynamics of interphase chromosomes. *PLoS Comput. Biol.* 4(8).

[12] Mateos-Langerak J, et al. (2009) Spatially confined folding of chromatin in the interphase nucleus. *Proc. Natl. Acad. Sci. U. S. A.* 106:3812–3817.

[13] Rosa A, Becker NB, Everaers R (2010) Looping probabilities in model interphase chromosomes. *Biophys. J.* 98(11):2410–2419.

[14] Barbieri M, et al. (2012) Complexity of chromatin folding is captured by the strings and binders switch model. *Proc. Natl. Acad. Sci. U. S. A.* 109(40):16173–16178.

[15] Jost D, Carrivain P, Cavalli G, Vaillant C (2014) Modeling epigenome folding: Formation and dynamics of topologically associated chromatin domains. *Nucleic Acids Res.* 42(15):9553–9561.

[16] Zhang B, Wolynes PG (2015) Topology, structures, and energy landscapes of human chromosomes. *Proc. Natl. Acad. Sci. U. S. A.* 112(19):6062–6067.

[17] Brackley CA, Johnson J, Kelly S, Cook PR, Marenduzzo D (2016) Simulated binding of transcription factors to active and inactive regions folds human chromosomes into loops, rosettes and topological domains. *Nucleic Acids Res.* 44(8):3503–3512.

[18] Pierro MD, Zhang B, Aiden EL, Wolynes PG, Onuchic JN (2016) Transferable model for chromosome architecture. *Proc. Natl. Acad. Sci. U. S. A.* 113(43):12168–12173.

[19] Zhang B, Wolynes PG (2016) Shape Transitions and Chiral Symmetry Breaking in the Energy Landscape of the Mitotic Chromosome. *Phys. Rev. Lett.* 116(24).

[20] Kang H, Yoon YG, Thirumalai D, Hyeon C (2015) Confinement-induced glassy dynamics in a model for chromosome organization. *Phys. Rev. Lett.* 115(19):198102.

[21] Michieletto D, Orlandini E, Marenduzzo D (2016) Polymer model with Epigenetic Recoloring Reveals a Pathway for the *de novo* Establishment and 3D Organization of Chromatin Domains. *Phys. Rev. X* 6(4).

[22] Hihara S, et al. (2012) Local Nucleosome Dynamics Facilitate Chromatin Accessibility in Living Mammalian Cells. *Cell Rep.* 2(6):1645–1656.

[23] Zidovska A, Weitz Da, Mitchison TJ (2013) Micron-scale coherence in interphase chromatin dynamics. *Proc. Natl. Acad. Sci. U. S. A.* 110(39):15555–60.

[24] Shinkai S, Nozaki T, Maeshima K, Togashi Y (2016) Dynamic Nucleosome Movement Provides Structural Information of Topological Chromatin Domains in Living Human Cells. *PLoS Comput. Biol.* 12(10):e1005136.

[25] Dhillon IS (2001) Co-clustering documents and words using bipartite spectral graph partitioning in *Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining.* (ACM), pp. 269–274.

[26] Halverson JD, Smrek J, Kremer K, Grosberg AY (2014) From a melt of rings to chromosome territories: the role of topological constraints in genome folding. *Rep. Prog. Phys.* 77(2):022601.

[27] Fudenberg G, Imakaev M (2017) FISH-ing for captured contacts: towards reconciling FISH and 3c. *Nat. Methods* 14(7):673–678.

[28] Giorgetti L, et al. (2014) Predictive Polymer Modeling Reveals Coupled Fluctuations in Chromosome Conformation and Transcription. *Cell* 157(4):950–963.

[29] Nagano T, et al. (2013) Single-cell Hi-C reveals cell-to-cell variability in chromosome structure. *Nature* 502(7469):59–64.

[30] Stevens TJ, et al. (2017) 3D structures of individual mammalian genomes studied by single-cell Hi-C. *Nature* 544(7648):59–64.

[31] Finn E, et al. (2017) Heterogeneity and Intrinsic Variation in Spatial Genome Organization. *bioRxiv.*

[32] Kirkpatrick T, Thirumalai D (1988) Comparison between dynamical theories and metastable states in regular and glassy mean-field spin models with underlying first-order-like phase transitions. *Phys. Rev. A* 37(11):4439.

[33] Götze W (2008) *Complex dynamics of glass-forming liquids: A mode-coupling theory.* (OUP Oxford) Vol. 143.

[34] Thirumalai D, Mountain RD (1993) Activated dynamics, loss of ergodicity, and transport in supercooled liquids. *Phys. Rev. E* 47(1):479.

[35] Chaudhuri P, Berthier L, Kob W (2007) Universal nature of particle displacements close to glass and jamming transitions. *Phys. Rev. Lett.* 99(6):060604.

[36] Levi V, Ruan Q, Plutz M, Belmont AS, Gratton E (2005) Chromatin dynamics in interphase cells revealed by tracking in a two-photon excitation microscope. *Biophys. J.* 89(6):4275–4285.

[37] Bronstein I, et al. (2009) Transient anomalous diffusion of telomeres in the nucleus of mammalian cells. *Phys. Rev. Lett.* 103(1):018102.

[38] Weber SC, Spakowitz AJ, Theriot JA (2010) Bacterial chromosomal loci move subdiffusively through a viscoelastic cytoplasm. *Phys. Rev. Lett.* 104(23):238102.

[39] Javer A, et al. (2013) Short-time movement of E. coli chromosomal loci depends on coordinate and subcellular localization. *Nat. Commun.* 4.

[40] Chen B, et al. (2013) Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. *Cell* 155(7):1479–1491.

[41] Lucas JS, Zhang Y, Dudko OK, Murre C (2014) 3D Trajectories Adopted by Coding and Regulatory DNA Elements: First-Passage Times for Genomic Interactions. *Cell* 158(2):339–352.

[42] Bronshtein I, et al. (2015) Loss of lamin A function increases chromatin dynamics in the nuclear interior. *Nat. Commun* 6:8044.

[43] Lifshitz IM, Slyozov VV (1961) The kinetics of precipitation from supersaturated solid solutions. *J. Phys. Chem. Solids* 19(1-2):35–50.

[44] Canzio D, et al. (2013) A conformational switch in HP1 releases auto-inhibition to drive heterochromatin assembly. *Nature* 496(7445):377–381.

[45] Mulligan PJ, Koslover EF, Spakowitz AJ (2015) Thermodynamic model of heterochromatin formation through epigenetic regulation. *J. Phys.: Condens. Matter* 27(6):064109.

[46] Van Holde KE (2012) *Chromatin.* (Springer Science & Business Media).

[47] Sanborn AL, et al. (2015) Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. *Proc. Natl. Acad. Sci. U. S. A.* 112(47):E6456–E6465.

[48] Ernst J, et al. (2011) Mapping and analysis of chromatin state dynamics in nine human cell types. *Nature* 473(7345):43–49.

[49] Honeycutt J, Thirumalai D (1992) The nature of folded states of globular proteins. *Biopolymers* 32(6):695–709.