Topological Constraints with Optimal Length Promote the Formation of Chromosomal Territories at Weakened Degree of Phase Separation

Jiachen Wei^{12*}, Hao Tian³⁴⁵, Rui Zhou³⁴⁵, Yingfeng Shao¹⁶, Fan Song¹⁶ and Yi Qin Gao^{2345*}

- 1. State Key Laboratory of Nonlinear Mechanics and Beijing Key Laboratory of Engineered Construction and Mechanobiology, Institute of Mechanics, Chinese Academy of Sciences, Beijing 100190, China.
- 2. Shenzhen Bay Laboratory, 5F, No.9 Duxue Road, Nanshan District, 518055 Shenzhen, Guangdong, China.
- 3. Beijing National Laboratory for Molecular Sciences, College of Chemistry and Molecular Engineering, Peking University, Beijing 100871, China.
- 4. Biomedical Pioneering Innovation Center (BIOPIC), Peking University, Beijing 100871, China.
- 5. Beijing Advanced Innovation Center for Genomics (ICG), Peking University, Beijing 100871, China.
- 6. School of Engineering Science, University of Chinese Academy of Sciences, Beijing 100049, China.

E-mail: weijiachen@lnm.imech.ac.cn; gaoyq@pku.edu.cn

ABSTRACT

It is generally agreed that the nuclei of eukaryotic cells at interphase are partitioned into disjointed territories, with distinct regions occupied by certain chromosomes. However, the underlying mechanism for such territorialization is still under debate. Here we model chromosomes as coarse-grained block copolymers and to investigate the effect of loop domains (LDs) on the formation of compartments and territories based on dissipative particle dynamics. A critical length of LDs, which depends sensitively on the length of polymeric blocks, is obtained to minimize the degree of phase separation. This also applies to the two-polymer system: The critical length not only maximizes the degree of territorialization but also minimizes the degree of phase separation. Interestingly, by comparing with experimental data, we find the critical length for LDs and the corresponding length of blocks to be respectively very close to the mean length of topologically associating domains (TADs) and chromosomal segments with different densities of CpG islands for human chromosomes. The results indicate that topological constraints with optimal length can contribute to the formation of territories by weakening the degree of phase separation, which likely promotes the chromosomal flexibility in response to genetic regulations.

Modeling for human chromosomes

The human chromosome 1 (Chr1) and human chromosome 2 (Chr2) are modeled as chain of DPD beads of total length $L_{\alpha} = 228$ Mb and $L_{\beta} = 240$ Mb. (in unit of bead number) respectively, with each bead representing 40 kb segment. The genome can be divided into two sequentially, epigenetically, and transcriptionally distinct regions, namely forest (*f*) and prairie (*p*), based on its CGI (CpG island) density ^{1–3}. The bead is then in type of *f* (*p*), if more than half of the 40 kb segment of the bead belong to forests (prairies). Please note that in our model the centromeric regions for Chr1 and Chr2 are omitted, due to a lack of sequence information.

Similar to the modeling of loop domains (LDs) for -50-50- polymers, LDs for Chr1 and Chr2 are modelled as loops of fixed size and position. The binding sites for LDs are derived from the position of TAD boundaries obtained from Ref.[4]. If there exists TAD boundary in one segment, the corresponding bead is deemed as a cross-linking bead for LDs. Harmonic bonds between adjacent cross-linking beads are introduced to mimic the TAD constraints, with spring constant and equilibrium bond length same to that for -50-50- polymers.

We note that similar distributions of the length of the forest or prairie segments (F/P) $L_{f,p}$ and TADs L_{tad} are obtained for all human chromosomes. In Fig.S1 below, we compare the distributions of L_{tad} and $L_{f,p}$ for human chromosomes 1-5 (Chr1-Chr5). The distributions for F/P lengths can be well fitted by Gamma distributions (Γ). The prediction of the peak position and the sharp decrease at very small $L_{f,p}$ for the distribution of F/P lengths by Γ even outperforms the kernel density estimation (KDE). The distribution of TAD lengths L_{tad} also follows Γ . The full list of shape parameter α and the scale parameter β of Γ for Chr1-Chr5 are presented in Table.S1 below. The mean length of consecutive forest and prairie beads $\overline{L_{f,p}}$ fluctuates from 1.60 Mb to 2.72 Mb for Chr1-Chr5, thus in the main text we choose a polymer with $L_f = L_p = 1.96$ Mb to remove the effect of F/P lengths on phase behavior. The mean length of TADs $\overline{L_{tad}}$ fluctuates within a small range (450 kb-550 kb), which is very close to the critical LD length $L_{td}^*=440$ kb and $L_{td}^*=520$ kb measured in our main text.



Figure S1: Comparisons of the distributions of F/P lengths (a-e) $L_{f,p}$ and TAD lengths (f-j) L_{tad} for human chromosomes 1-5 (Chr1-Chr5) derived from embryonic stem cells (H1). Solid lines are fitting curves of kernel density estimation (KDE) and Gamma distributions (Γ). Dashed lines indicate the corresponding mean lengths of F/P segments $\overline{L_{f,p}}$ and TADs $\overline{L_{tad}}$.

Table S1: The full list of shape parameter α and the scale parameter β for Gamma

# Chr	$\alpha_{f,p}$	$\beta_{f,p}$	α_{ld}	β_{ld}
Chr1	1.33	0.83	2.94	6.43
Chr2	1.54	0.76	3.21	6.30
Chr3	1.26	0.47	3.43	6.72
Chr4	1.39	0.57	4.03	7.29
Chr5	1.40	0.72	2.97	5.54

distributions for Chr1-Chr5.

Note that the distributions of L_{tad} may vary across chromosomes taken from different cell lines or tissues. For instance, in Fig.S2 we compare the distributions of L_{tad} for Chr1 taken from H1 (embryonic stem cells), LV (liver primary tissue) and HepG2 (liver cancer cells). We estimate the mean length of TADs $\overline{L_{tad}} = 0.50$ Mb, $\overline{L_{tad}} = 0.60$ Mb and $\overline{L_{tad}} = 0.43$ Mb respectively for Chr1 taken from H1, LV and HepG2⁴. The largest variance of L_{tad} is obtained for Chr1 taken from LV.

TADs are of hierarchic structure with small loops forming inside bigger ones. These loops are dynamic, being generated and opened repeatedly in real time. In our coarse-grained simulations, both nested loops within TADs and the unfolding of TADs are neglected. This is because these loops are not forming consecutively, which shall exert limited influence on the equilibrium configurations of the entire chain (and thus the degree of phase separation). Additionally, the size of the nested loops can be smaller than the size of two beads (80 kb). The model provides no explanation for the formation of TADs in chromosomes, but rather serves as a tool to investigate the effect of topological constraints at length scales comparable with TADs (in form of consecutive loops) on phase separation and territory formation, simply based on polymer physics. In addition, our simulation results suggest that the degree of phase separation becomes stronger when loop domains are partially unfolded. This process could be compensated, to some extent, by the formation of smaller loops. The overall effect is the formation of the "average" TAD boundaries that are highly conserved across different cell lines and tissues. Nevertheless, we acknowledge that this is an idealized scenario, and the contact map of which could be quite different from experimental data.



Figure S2: the distributions of TAD lengths L_{tad} for human chromosome 1 (chr1) taken from H1, LV and HepG2.

Conversion of units

In this section we present how reduced units can be converted into real units (with insufficient accuracy). By comparing the root mean squared end-to-end distance R(s) measured from our simulations and that from recent experiments⁶, the unit length (i.e. cutoff for DPD interactions) \hat{r}_c can be estimated as 200*nm* in real unit.

Given the mass density of water $\hat{\rho}_s = 1 \text{ g/cm}^3 \rho_s = 3$ in simulations, the mass of the solvent beads corresponds to $\hat{m}_s = (\hat{\rho}_s / \rho_s) \hat{r}_c^{\ 3} = 2.67 \times 10^{-18} \text{ kg}$. The unit of time can then be estimated as $\hat{\tau} = (\hat{m}_s \hat{r}_c^{\ 2} / k_B \hat{T})^{1/2} = 5.1 \,\mu s$, with $\hat{T} = 298.15K$ the system temperature and k_B the Boltzmann constant. The time step used in the simulations then corresponds to $\delta \hat{t} = 0.1 \mu s$. The simulated diffusion coefficient for free solvent beads $D_t = 0.08$ can be converted into $\hat{D}_t = D_t \hat{r}_c^2 / \hat{\tau} = 0.63 \times 10^{-9} m^2 / s$, which is slightly lower than the diffusion coefficient of bulk water (2.30 × 10⁻⁹ m²/s). We note that such conversion shall only be taken as a rough guide.⁷⁻⁸

Analysis on polymers at swollen state



Figure S3: (a) Sketch of the gyration radius R_g of a LD. (b) The dependence of average gyration radius, $\overline{R_g}$, on $L_{\rm ld}$ for LDs that are formed exclusively by p beads at $A_{\rm sp} = 150$ (black) and $A_{\rm sp} = 160$ (red). (c) The dependence of radial density of polymer beads $\rho_{\rm c}$ and (d) solvent beads $\rho_{\rm s}$ on radial distance from the backbone δr at $A_{\rm sp} = 150$.

We measure the average gyration radius $\overline{R_g}$ of each loop formed by LD boundaries, as illustrated by Fig.S3(a), for polymer without hydrophobicity (i.e. polymer consists of exclusively *f* beads). As shown in the inset of Fig.S3(b), at $A_{sf} = 150$ the average gyration radius $\overline{R_g}$ of loops increases in power law of L_{ld} with exponent $\alpha_R = 0.65$, which is slightly larger than the theoretical prediction $\alpha_R^{\text{theory}} = 0.59$ for a 3D ring of self-avoiding chain.⁹ As indicated by the dashed line for $\overline{R_g} = 1$ in Fig. S3(b), there exists a critical length $L_{ld}^c = 560$ kb above which solvent beads or polymer beads on neighboring LDs can go inside the loop without energy compensation. This length is very close to the critical LD length $L_{ld}^* = 440$ kb for -50-50polymer measured in our main text. In poor solvent with $A_{sf} = 160$, $\overline{R_g}$ is barely changed for $L_{ld} < 520$ kb due to tight compaction, but is of smaller value for $L_{ld} \ge 520$ kb (in power law of L_{ld} with $\alpha_R = 0.56$). The result indicates LDs become loose enough to entrap solvent beads (or polymer beads from neighboring LDs) for $L_{ld} \ge 520$ kb in our model.

The infiltration of solvent beads can also be determined from the density distribution of polymer beads ρ_c and solvent beads ρ_s on radial distance from the backbone δr at different L_{ld} . To measure the cross-sectional density distributions, we define a backbone that indicates the alignment of the polymer, as presented in Fig.1(d) of the main text. The backbone is of coarsegrained degree $N_k = 3(L_{ld}/40 \text{ kb})$: the first point of the backbone chain is the center of mass of beads $1, 2, ..., N_k$, and the second point of the backbone is the center of mass of beads $2, 3, ..., N_k + 1$, and so forth. The radial density perpendicular to the backbone (i.e. $\delta r = 0$) decreases significantly with the increase of L_{ld} . The solvent density in bulk $\rho_s^{bulk} = 3.0$. Here $\rho_c \leq 3.0$ is reached at around $L_{ld} = 520 \text{ kb}$, which indicates the polymer becomes loose enough to entrap solvent beads for $L_{ld} \geq 520 \text{ kb}$.

Measuring compartmentalization based on graph theory

To measure compartmentalization of chromosome based on graph theory¹⁰, first let each prairie bead *p* on the polymer represent a node. Edges are created between nodes in close contact $(r_{ij} < r_c^G)$, with cutoff $r_c^G = 4.0$. The graph is then denoted as $G(E_g, V_g)$, with E_g and V_g

respectively representing the set of edges and nodes. Louvain algorithm¹¹ is implemented to determine the communities (i.e. sub-clusters) of $G(E_g, V_g)$, with the average size of communities denoted as S_c . The modularity Q_m measures the relative density of edges inside communities with respect to edges outside communities, which indicates the robustness of the division of the communities. To check whether S_c or Q_m better describes the degree of phase separation, in Fig. S4 we compare S_c and Q_m for three -50-50- polymers with different values of A_{sp} and L_{ld} . Both S_c and Q_m depend sensitively on A_{sp} . When TADs with $L_{ld} = 520$ kb are added at fixed $A_{sp} = 170$, the configuration of the polymer changes from a fully-segregated globule of p beads to a less-segregated "rod". This leads to about 20% increase of Q_m . However, S_c is almost unchanged by adding TADs. We thus choose Q_m as an order parameter for phase separation.



Figure S4: The (a) average size of communities S_c and (b) modularity Q_m for three -50-50-polymers at different A_{sp} and L_{ld} .

Fig. S5 presents the dependence of modularity Q_m of the same -50-50- polymer on A_{sp} at different L_{ld} . Different from $\Delta = 0.0$ obtained at $A_{sp} = 150$ for any given L_{ld} , here Q_m slightly decreases with L_{ld} at $A_{sp} = 150$ when $L_{ld} > 80$ kb. For $A_{sp} \ge 160$, Q_m depends also non-monotonically on L_{ld} . As shown in Fig. S5(b), the maximum value of Q_m at $A_{sp} = 170$ shall

locate at around 320 kb< L_{ld} . < 520 kb, which is in consistency with the critical length $L_{ld}^* =$ 440 kb that minimizes Δ measured in the main text.



Figure S5: (a) The dependence of modularity Q_m of the community generated by Louvain algorithm on A_{sp} at different L_{ld} for a -50-50- polymer. (b) The dependence of Q_m on L_{ld} at $A_{sp} = 170$.

Effect of interaction parameters



Figure S6: The dependence of the (a) order parameter Δ and (b) difference in order parameter between systems with and without LD, $\Delta - \Delta_0$, on L_{ld} for phase separation of a -50-50-polymer driven by either hydrophobicity (black) or effective attraction between *p* beads (red). Results for $A_{sp}=A_{pp}=150$ (blue) are also presented for reference.

Fig. S6 compares the order parameter Δ for phase separation driven either by hydrophobicity (with $A_{sp} = 150$ and $A_{pp} = 130$) or by effective attraction between p beads (with $A_{sp} = 160$ and $A_{pp} = 150$). Based on the χ -parameter theory for polymer solutions, the Flory–Huggins parameter for *s* and *p* beads χ_{sp} can be approximated as:

$$\chi_{\rm sp} = 2\alpha \overline{\rho} (A_{sp} - \frac{A_{ss} + A_{pp}}{2}), \tag{1}$$

with α depending on the pair correlation function and $\overline{\rho}$ the average bead-density. If we roughly assume α as a constant and $\rho_s = \rho_p = \overline{\rho}$ under segregation, the above two systems are of the same χ_{sp} . Indeed, very similar $\Delta - L_{ld}$ relations are obtained. For both systems, the minimum values of Δ are within the range 320 kb< L_{ld} <520 kb. This indicates the change of repulsive parameters A_{sp} and A_{pp} exert very limited impact on the position of critical LD length L_{ld}^* . Since the cutoff radius for measuring contact pairs $r_c = 1$ is invariant, Δ at the same L_{ld} is larger for system with effective attraction between p beads. By comparing the variation $\Delta - \Delta_0$ (Δ_0 denoting the degree of phase separation at $L_{ld} = 0$), as shown in Fig. S6(b), we find Δ is slightly more sensitively dependent on L_{ld} for phase separation driven by hydrophobicity of p beads.

In this work, we fix all repulsive parameters at $A_{ij}=150$ and exclusively vary the prairie-solvent repulsion A_{sp} . The phase separation is thus driven by the hydrophobicity of the prairie beads. If A_{sf} is also varied, the effective Flory-Huggins parameter χ_{eff} for *f* and *p* beads¹²⁻¹³ can be estimated as:

$$\chi_{\rm eff} = \frac{V_{f,p} (\delta_f - \delta_p)^2}{RT},\tag{2}$$

where $V_{f,p}$ denotes the average molar volume of the beads, δ_f and δ_p denote the Hildebrand solubility parameter for *f* and *p* beads respectively. We thus expect larger degree of phase separation Δ for larger difference in repulsive parameter $|A_{sp}-A_{sf}|$. Fig.S7(a-b) below presents the dependence of Δ on A_{sf} for a single -50-50- polymer. For any given L_{ld} , Δ decreases with A_{sf} at fixed A_{sp} =160, which is in accord with the prediction of Eq.2. Additionally, the critical loop length L_{ld}^* =440 kb is unchanged for A_{sf} =140, as shown in Fig.S7(b). This again indicates that L_{ld}^* is an intrinsic feature of the polymer determined by the length of F/P segments rather than interaction parameters. For the two-polymer system, as illustrated by Fig.10(g-h), Ω also decreases with A_{sf} at fixed A_{sp} . For $A_{sf} < 150$, *f* beads tend to extend within the solvent. This creates an effective "repulsion" between *f* beads that prevents inter-polymeric contacts of *p* beads. For $A_{sf} > 150$, both *f* and *p* beads are collapsed so that Ω gets much smaller. Very similar Ω - L_{ld} relation is also observed for both $A_{sf} = 140$ and $A_{sf} = 150$, with the same critical length obtained at around $L_{ld}^* = 520$ kb.



Fig.S7: (a-b) The dependence of Δ on (a) A_{sf} and (b) L_{ld} for a single -50-50- polymer. The prairie-solvent repulsion is fixed at $A_{sp}=160$. (c-h) The dependence of Ω , P^{same} and P^{diff} on (c,e,g) A_{sf} and (d,f,h) L_{ld} for two -50-50- polymers at fixed $A_{sp}=170$.

Fig.S8 below presents the local densities of forest, ρ_f , prairie, ρ_p , and solvent beads, ρ_s , as functions of genomic distance *S* for a single -50-50- polymer at various $L_{\rm ld}$. The phase separation is driven either by (a-c) hydrophobicity or (d-f) effective attraction between *p* beads. The local densities of forest beads in proximity to polymeric beads are higher at loop length $L_{\rm ld}$ =440 kb for both cases. This is because *f* beads are more likely to get buried inside the collapsed globule for $L_{\rm ld} \sim L_{\rm ld}^*$. Consequently, slightly lower local densities of *p* beads in proximity to *p*, and lower densities of *s* beads in proximity to *f* are observed for $L_{\rm ld}$ =440 kb. The density fluctuation is also larger at $L_{\rm ld}$ =440 kb.



Fig.S8: (a-f) The local densities of (a,d) forest ρ_f (b,e) prairie ρ_p and (c,f) solvent beads ρ_s as a function of genomic distance *S* for a single -50-50- polymer at various $L_{\rm ld}$. The phase separation is driven either by (a-c) hydrophobicity or (d-f) effective attraction between *p* beads. (g-i) the mean densities of (g) *p* in proximity to *f*, $\overline{\rho_{fp}}$, (h) *p* in proximity to *p*, $\overline{\rho_{pp}}$, and (i) *s* in proximity to all polymeric beads , $\overline{\rho_s}$, as function of $L_{\rm ld}$ with two sets of interaction parameters.

For better comparison of local densities with two sets of repulsive parameters, in Fig.S8 (g-i) we present the mean densities of p in proximity to f, $\overline{\rho_{fp}} p$ in proximity to p, $\overline{\rho_{pp}}$, and s in proximity to all polymeric beads, $\overline{\rho_s}$, as functions of $L_{\rm ld}$. As expected, larger $\overline{\rho_{pp}}$ for phase separation driven by effective attraction and larger $\overline{\rho_s}$ for phase separation driven by hydrophobicity are obtained for any given $L_{\rm ld}$. $\overline{\rho_{fp}}$ peaks at the critical length $L_{\rm ld}^*$ =440 kb. However, $\overline{\rho_s}$ is minimized at a lower loop length $L_{\rm ld}$ = 320 kb. This is in consistency with the minimum of the unit solvation energy $u_{\rm sol}$ at $L_{\rm ld}$ = 320 kb shown in Fig.S12(c).

We also measure the order parameter Δ for polymers with different bond stiffness *K* for connected beads, as shown in Fig. S9. When *K* is decreased from 200 to 100, the Δ - L_{ld} relation is barely changed for $L_{ld} < 160$ kb. Although Δ becomes larger at smaller *K* for $L_{ld} > 160$ kb, the critical LD length that minimizes Δ is always located at $L_{ld}^* \sim 520$ kb for any given *K*. If *K* is further decreased to 50, a slightly smaller critical length $L_{ld}^* \sim 320$ kb is obtained (data not shown). However, the rule of topological constraints (i.e. the chain cannot cross itself) is also severely violated for polymers with $K \leq 50$.



Figure S9: The dependence of Δ on L_{ld} for -50-50- polymers of different bond stiffness K for connected beads.

The results indicate that the magnitude of the critical length L_{1d}^* is not an artifact caused by the choice of interaction parameters, but rather an intrinsic feature of the polymer determined by the length of F/P segments.

Shape of collapsed polymers

Fig.S10(a) presents the radius of gyration as a function of the length for loop domains L_{ld} . For a single -50-50- polymer, a small peak for the gyration radius of the polymer $R_{f,p}$ is observed at L_{ld} =440 kb. For the two-polymer system, the peak can no longer be discerned due to the large variance of $R_{f,p}$. To remove the effect of forest beads on collapsing behavior, the gyration radius of prairie beads R_p is measured in Fig.S10(b). The peak values of R_p are obtained at L_{ld} =440 kb and L_{ld} =520 kb respectively for the single and two-polymer system, which agree well with the critical length L_{ld} * for phase separation determined in the main text.



Fig.S10: The dependence of (a) gyration radius of the polymer $R_{f,p}$ and (b) gyration radius of p beads R_p on L_{ld} for a single -50-50- polymer (black) and two -50-50- polymers (red).

Fig.S11 presents $R_{f,p}$ and Δ as a function of L_{ld} for a -50-50- polymer with $A_{sp}=A_{sf}=160$ (i.e. fully collapsed polymer). Here $R_{f,p}$ reduces to 6.0 and Δ fluctuates around 0.0, both of which are barely changed with L_{ld} . This indicates the degree of collapsing does not depend on the elasticity of the entire chain, but is determined by the combined effect of segmental rigidity and sequence randomness. This also explains why L_{ld} and $L_{f,p}$ are correlated.



Fig.S11: The dependence of (a) $R_{f,p}$ and (b) Δ on L_{ld} for a fully collapsed -50-50- polymer with $A_{sf}=A_{sp}=160$.

The aspheric parameter Λ of the crumpled polymers can be defined as:¹⁴

$$\Lambda = 1 - 3 \frac{\langle \lambda_1 \lambda_2 + \lambda_1 \lambda_3 + \lambda_2 \lambda_3 \rangle}{\langle (\lambda_1 + \lambda_2 + \lambda_3)^2 \rangle},\tag{3}$$

with λ_1 , λ_2 and λ_3 denoting the eigenvalues of the gyration tensor. By definition, $\Lambda \in [0,1]$ measures the degree of asphericity of the crumpled polymer. We find the peak value of Λ also locates at around the critical length L_{ld}^* .

Potential energy for $-n_f - n_p$ - polymers

In the main text, we find the critical LD length that minimizes phase separation L_{ld}^* depends sensitively on $L_{f,p}$. In general, the number of LDs that cross-link *f* and *p* can be minimized when $L_{f,p} = n(L_{ld} - 1)$ (n=1, 2, 3 ... is a positive integer). However, we do not observe any periodic peaks of Δ (with $L_{f,p} = n(L_{ld} - 1)$) in Fig.3(a) of the main text. This means the sensitive dependence of L_{ld}^* on $L_{f,p}$ is an effect of elasticity rather than the pre-determined LD contacts.

The competition between elastic energy and surface energy cannot be determined directly by measuring the corresponding energy terms. This is because the polymeric beads significantly

overlap each other even in the unfolded state (obtained at $A_{sp}=150$) due to the bonded interactions between loop boundaries.



Figure S12: The dependence of (a) unit bond energy u_B , (b) unit polymeric energy u_{poly} and (c) unit solvation energy u_{sol} on L_{ld} for a single -50-50- polymer at A_{sp} =160. (d-f) The variations of u_B , u_{poly} and u_{sol} compared to the unfolded state (obtained at A_{sp} =150) for the same system.

Fig. S12(a-c) presents the dependence of unit bond energy $u_{\rm B} = U_{\rm B}/N_{\rm poly}$, unit polymeric energy $u_{\rm poly} = U_{\rm poly}/N_{\rm poly}$ and the unit solvation energy $u_{\rm sol} = U_{\rm sol}/N_{\rm poly}$ on $L_{\rm ld}$ at $A_{\rm sp}$ =160, where $U_{\rm B}$, $U_{\rm poly}$, $U_{\rm sol}$ and $N_{\rm poly}$ respectively denote the total bond energy, the total energy between polymeric beads, the total energy between solvent and polymeric beads and the number of polymeric beads. Since forest beads are more likely to become buried inside the nonspherically collapsed globule for $L_{\rm ld} \sim L_{ld}^*$, a large dip for $u_{\rm sol}$ is obtained at $L_{\rm ld}$ =320 kb. Both $u_{\rm B}$ and u_{poly} reach the peak values at L_{ld} =80 kb and gradually decrease with L_{ld} , following the bonds formed between nonadjacent beads at loop boundaries, the number of which also reaches a peak at L_{ld} =80 kb and decreases with L_{ld} . To provide a further test, in Fig. S12(d-f) we compare the variations of u_B , u_{poly} and u_{sol} from the unfolded state (obtained at A_{sp} =150) to collapsed state (obtained at A_{sp} =160). We find that u_B is barely changed, and the magnitude of δu_B no longer depends on L_{ld} . Interestingly, the variation of polymeric energy δu_{poly} depends sensitively on L_{ld} and reaches its minimum at L_{ld} =440 kb. This result indicates the polymer is of the smallest degree of collapse at $L_{ld}=L_{ld}^*$ due to elastic repulsions. The variation in solvation energy δu_{sol} slightly decreases with L_{ld} and levels off around -2.0 for $L_{ld}>L_{ld}^*$, which means polymer could achieve greater surface energy in a less collapsed state.



Fig.S13: The dependence of (a) unit polymeric energy u_{poly} and (b) unit solvation energy u_{sol} on L_{ld} for two -50-50- polymers at $A_{sp}=170$. (c-d) The variations of u_{poly} and u_{sol} compared to unfolded state (obtained at $A_{sp}=150$) for the same system.

Next we compare the dependence of unit polymeric energy u_{poly} and unit solvation energy u_{sol} on L_{ld} for the two-polymer system in Fig.S13. Similarly, we find the peak for u_{poly} at L_{ld} =80 kb and the dip for u_{sol} at L_{ld} =320 kb, which demonstrates that these energy terms are determined by

the initial configurations (i.e. L_{ld}). Fig.S13(c-d) compares the variations of u_{poly} and u_{sol} from the unfolded state (obtained at $A_{sp}=150$) to the fully collapsed state (obtained at $A_{sp}=170$). Interestingly, both the variation of polymeric energy δu_{poly} and variation of solvation energy δu_{sol} reach their maximum/minimum at $L_{ld}=L_{ld}^*=520$ kb. This result confirms that the twopolymer system reaches its smallest degree of collapse at $L_{ld}^*=520$ kb, which is slightly larger than that for a single -50-50- polymer (i.e. $L_{ld}^*=440$ kb). By comparing Fig.S12(f) with Fig.S13(d), we find this change to be attributable to the depletion effect: the decrease in solvation energy from phase separated state to unfolded state is significantly larger for the twopolymer system than that for a single polymer, especially for $L_{ld} > L_{ld}^*$.

Determination of Δ from Hi-C data

Similar to the measurement of Δ in our simulations, experimentally the magnitude of Δ can be determined as:

$$\Delta = \frac{\sum_{S=1}^{L/2} \langle (p_S^{\text{same}} - p_S^{\text{diff}}) / (p_S^{\text{same}} + p_S^{\text{diff}}) \rangle}{L/2},$$
(3)

where p_S^{same} and p_S^{diff} respectively denote the average probability of contact in same and different types at curvilinear distance S = i - j along the chain. Both p_S^{same} and p_S^{diff} are derived from the ICE normalization of the original counts from Hi-C data. We thus obtain smaller degree of phase separation for Chr1 and Chr2 taken from LV and H1 (with smaller difference $\delta \varepsilon = |\varepsilon - \varepsilon^*|$) than that taken from HegG2 (Table 2 of main text), which is consistent with Δ for Chr1+Chr2 determined in simulations. We do not expect an accurate prediction of Δ determined from Hi-C data based on our simulations, since the effects of confinement, lamina-associated domains (LADs), nucleolus and interactions with other chromosomes are not taken into account.

Further discussions on P^{same} and P^{diff}

The cross-linking between two polymers works against intra-polymer compactness. As illustrated in Fig.S14(a-b), for $L_{ld} <520$ kb, while the probability of inter-chromosomal contact, P^{diff} , decreases with L_{ld} , the probability of intra-chromosomal contact P^{same} is significantly increased. The relation is reversed for $L_{ld} >520$ kb, as loops formed by LDs become looser.

Fig.8 of the main text indicates that Ω decreases monotonically with A_{sp} for human chromosome 1 and 2 (Chr1+Chr2) taken from embryonic stem cells⁴ (H1). However, as indicated by Fig.S14(c-d), for Chr1+Chr2 both P^{same} and P^{diff} increase with A_{sp} . The decrease of Ω with A_{sp} is thus caused by a faster growth rate for P^{diff} compared to that for P^{same} . This growth rate of P^{diff} becomes even larger for polymers without LD.



Figure S14: (a-b) The dependence of (a) the probability of intra-chromosomal contact P^{same} and (b) the probability of inter-chromosomal contact P^{diff} on L_{ld} for two -50-50- polymers at $A_{\text{sp}} = 160$. (c-d) The dependence of (c) P^{same} and (d) P^{diff} on A_{sp} for Chr1+Chr2 with LD boundaries taken from H1 or without LD at $A_{\text{sp}} = 150$.

Fig. S15 presents normalized number of contacting pairs $N_{cr}/(N_{\alpha} + N_{\beta})$ as a function of L_{ld} at $A_{sp} = 170$, with N_{cr} , $N_{\alpha} = 5706$ and $N_{\beta} = 6006$ respectively denoting number of contacting pairs across two -50-50- polymers, bead number for polymer α and β . For $L_{ld} < 320$ kb N_{cr} decreases significantly with L_{ld} , while for $L_{ld} > 320$ kb N_{cr} slightly increases. This turning point at around $L_{ld} \sim 320$ kb is again very close to the critical LD length ($L_{ld}^* = 440$ kb) that impedes phase separation for a single polymer.



Figure S15: Normalized number of contacting neighbours for two -50-50- polymers, $N_{cr}/(N_{\alpha} + N_{\beta})$, as a function of L_{ld} at $A_{sp} = 170$.

Effect of heterogeneous loop sizes

Due to the effect of loop entropy, variance of the loop length for human chromosomes can stabilize the phase transitions¹⁵. In absence of variance, the loop entropy $S_{lp} \sim N_{lp} \log(L_{lp})$, where N_{lp} and L_{lp} respectively denote the number and length of the loops. By contrast, for system with sufficiently small variance, the entropy $S_{lp}^{dis} \sim \sum_{i}^{N} \log(L_{i}) = N_{f,p} \sum_{i}^{N} \frac{1}{N} \log(L_{i}) = N_{f,p} \langle \log(L_{i}) \rangle < N_{f,p} \log(\langle L_{i} \rangle)$. This gives $S_{lp}^{dis} < S_{lp}$. Therefore, introducing variance will stabilize the formation of clusters. To check this, we compare the gyration radius of *p* beads R_{p} as a function of A_{sp} for polymer either with identical or different lengths of loop domains in Fig.S16. To remove the effect of segmental rigidity, the polymer is composed entirely of *p* beads. The transition from a crumple globule into unfolded chain takes place at round $A_{sp} = 153$ for polymers both of fixed

 $L_{ld} = 440$ kb (i.e. $\sigma = 0$) or with disordered distributions of L_{ld} in Gamma distributions with $\sigma = 12$. However, a smoother transition is obtained for the latter one in the presence of disorder.



Figure S16: Comparison of the R_g - A_{sp} relation for polymer with identical or different length of loop domains. The polymer is composed entirely of p beads.

We expect giant loops can protrude far from chromosome territories and intermingle with fibers from neighboring chromosomes in simulations. Fig.S17 presents the dependence of interchromosomal coordinate number density $n_{cr} = C_{cr}/N_{ld}$ (where C_{cr} and N_{ld} respectively denote the inter-chromosomal coordinate number and the number of beads for each loop) on L_{ld} for the two-polymer system with LD boundaries determined from Hi-C data of H1 (i.e. human stem cells). As expected, n_{cr} is positively correlated with L_{ld} , which can be approximated by linear fitting with a slope $k = n_{cr}/L_{ld} = 0.72 \text{ Mb}^{-1}$. These results provide evidence that, for two chromosomes in close contact, longer loops from one chain are more likely to get in touch with the neighboring chromosome.



Fig.S17: The dependence of inter-chromosomal coordinate number density n_{cr} on loop length L_{ld} for LD boundaries determined from Hi-C data of H1. The cutoff radius for measuring contact pairs $r_c=1.0$.

References

(1) Liu, S.; Zhang, L.; Quan, H.; Tian, H.; Meng, L.; Yang, L.; Feng, H.; Gao, Y. Q. From 1d sequence to 3d chromatin dynamics and cellular functions: A phase separation perspective. *Nucleic Acids Res* **2018**, *46*, 9367–9383.

(2) Quan, H.; Yang, Y.; Liu, S.; Tian, H.; Xue, Y.; Gao, Y. Q. Chromatin structure changes during various processes from a DNA sequence view. *Curr Opin Struct Biol* **2020**, *62*, 1–8.

(3) Zhou, R.; Gao, Y. Q. A DNA Sequence Based Polymer Model for Chromatin Folding. *International Journal of Molecular Sciences* **2021**, *22*, 1328.

(4) Schmitt, A. D.; Hu, M.; Jung, I.; Xu, Z.; Qiu, Y.; Tan, C. L.; Li, Y.; Lin, S.; Lin, Y.; Barr, C. L.;

Others, A compendium of chromatin contact maps reveals spatially active regions in the human genome. *Cell reports* **2016**, *17*, 2042–2059.

(5) Yardımcı, G. G.; Ozadam, H.; Sauria, M. E.; Ursu, O.; Yan, K. K.; Yang, T.; Chakraborty, A.; Kaul, A.; Lajoie, B. R.; Song, F., et al. Measuring the reproducibility and quality of Hi-C data. *Genome Biol.* 2019, 20, 57.

(6) Su, J.-H.; Zheng, P.; Kinrot, S. S.; Bintu, B.; Zhuang, X. Genome-scale imaging of the 3D organization and transcriptional activity of chromatin. *Cell* **2020**, *182*, 1641–1659.

(7) Liu, L.; Shi, G.; Thirumalai, D.; Hyeon, C. Chain organization of human interphase chromosome determines the spatiotemporal dynamics of chromatin loci. *PLoS computational biology* **2018**, *14*, e1006617.

(8) Shinkai, S.; Nozaki, T.; Maeshima, K.; Togashi, Y. Dynamic nucleosome movement provides structural information of topological chromatin domains in living human cells. *PLoS computational biology* **2016**, *12*, e1005136.

(9) Baumgärtner, A. Statistics of self-avoiding ring polymers. J Chem Phys 1982, 76, 4275-4280.

(10) Morlot, J.-B.; Mozziconacci, J.; Lesne, A. Network concepts for analyzing 3d genome structure from chromosomal contact maps. *EPJ Nonlinear Biomedical Physics* **2016**, *4*.

(11) Blondel, V. D.; Guillaume, J.-L.; Lambiotte, R.; Lefebvre, E. Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment* **2008**, *2008*, P10008.

(12) Huggins, M. L. Solutions of long chain compounds. J. Chem. Phys, 1941, 9, 440-440.

(13) Lu, H.; Du, S. A phenomenological thermodynamic model for the chemo-responsive shape memory effect in polymers based on Flory–Huggins solution theory. *Polym. Chem.* **2014**, *5*, 1155-1162.

(14) Szleifer, I.; Kramer, D.; Ben - Shaul, A.; Gelbart, W. M.; Safran, S. A. Molecular theory of curvature elasticity in surfactant films. *J. Chem. Phys*, **1990**, 92, 6800-6817.

(15) Negri, M.; Gherardi, M.; Tiana, G.; Cosentino Lagomarsino, M. Spontaneous domain formation in disordered copolymers as a mechanism for chromosome structuring. *Soft Matter* **2018**, 14, 6128–6136.