Homopolymer with intrachain reversible bonds as a model of chromatin large-scale spatial organization and dynamics

A.Petrov, P.Kos, and A.Chertovich
Faculty of Physics, Lomonosov Moscow State University, 119991 Moscow, Russia
(Dated: November 14, 2018)

We study structure of a homopolymer, which beads can form pairwise reversible bonds. If a large number of reversible bonds is formed, chain folds into a long-living structure, similar to metastable crumpled globule. For such structure we observe anomalously fast diffusion on a long time-scale, in agreement with experimental data in chromatin. Moreover, behavior of the contact probability and spatial distance between beads is very similar to the observed dependencies from Hi-C and FISH experiments in different organisms. We suggest a new universal model of large-scale chromatin structure with one free parameter - fraction of formed reversible bonds.

INTRODUCTION

Since the invention of Hi-C technique [1] much research was done to characterize spatial chromatin structure by the probability of finding two loci spatially close to each other, depending on the genomic distance between the loci, $P(s)$. It was revealed, that chromatin is packed hierarchically: on a scale of several thousands base pairs domain-like structures are formed, they called topologically-associated domains (TADs) [2]. In most species on a larger scale TADs form compartments, which segregate into active and inactive parts [2–4]. Moreover, $P(s)$ dependencies appeared to be significantly distinct for different species and different cell types [1, 4–7] (Supplementary Fig. S1), indicating different spatial organization of chromatin.

Structure of a single topologically-associated domain (TAD) [2] in mammals is successfully described by loop-extrusion model [8], and there is evidence of fractal globule model applicability for chromatin structure description on submegabase scale [1, 2, 9, 10]. However, there are deviations from fractal globule model predictions on the scale ≈ 10 Mbp and larger [2, 7]. Moreover, unexplained transition from fractal globule dynamics to fast anomalous subdiffusion on a long time-scale is observed in mammals [11, 13]. It is also a challenge to obtain a long-living fractal globule state: collapse of a chain, induced by the artificial gravitation-like potential [13], formed during collapse of a long chain in poor solvent, quickly transforms into globular state [10], and fractal globule state, observed in SBS model [7], exists in a narrow range of binder concentrations, and is, therefore, unstable to fluctuations. Attempts to describe large-scale chromatin structure were described in the work on the block copolymer models with different types of interactions [7, 17, 18]; however, dynamical behavior of the aforementioned models remains obscure. Therefore, the problem of large-scale chromatin structure description remains unsolved.

In this work we studied collapse of a long homopolymer chain, placed in athermal solvent (see "Methods" section). Beads of the chain can form pairwise reversible bonds, which induce the collapse. Varying the number of formed reversible bonds, we observed different states of the chain: swollen state with small number of formed reversible bonds, Gaussian chain, globular state with Gaussian statistics and a long-living non-equilibrium state, formed by large number of bonds and similar to metastable crumpled-globule with $P(s) \propto s^{-1}$. We studied dynamic behavior of the system and showed, that transition from fractal globule dynamics to much faster anomalous subdiffusion occurred on a long time-scale in the aforementioned non-equilibrium state. In our work we also compared our results with experimental data for different species and cell types, and propose a novel approach for description of large-scale chromatin structure.

METHODS

We used dissipative particle dynamics (DPD) method to perform simulations. DPD is a mesoscale molecular dynamics with explicit solvent and soft potential for the conservative force. We do not describe this method in detail, more information about DPD can be found in Ref. [19]. We studied a homopolymer chain (length $N = 1000$ beads) in athermal solvent (which corresponds to zero Flory-Huggins parameter, $\chi$, or, equivalently, self-avoiding walk (SAW) in 3D space). Volume interaction coefficients were chosen to satisfy the criterion $\chi = 0$, and to make chain non-phantom [20]. The simulation box with impermeable boundaries was chosen to be a cube $18 \times 18 \times 18$ DPD units, so the polymer concentration was $n \approx 0.06$ (DPD density $\rho = 3$ particles per unit volume). Pairwise reversible bonds can be formed with certain probability between any non-neighboring beads, which are spatially closer, than cutoff radius, $R_c = 1$. We define $N_{stp}$ as the number of DPD steps between calculation of formation/breaking of bonds, and it was equal to 200 steps in our simulations. The probability of breaking a bond $\beta$ was set constant $\beta = 0.004$ for all simulations. The average lifetime of a bond is defined as the
ratio $N_{\text{stp}}/\beta$. Therefore, the average bond lifetime was equal to 50000 DPD steps. The only variable parameter was the probability of bond formation. This parameter sets the time-averaged number of bonds formed inside the chain. We denote fraction of beads, which formed reversible bonds, as $f$, and further name this quantity as “fraction of bonds” for simplicity. We studied dynamics of the system by the time dependency of mean-squared displacement, $MSD(t)$. We distracted MSD of the center of mass from the $MSD(t)$ of the chain to study the details of dynamics inside the structure (detailed information in Supplementary).

To set the simulation time, we estimated Rouse reptation time in good solvent for the whole chain $\tau_R$ and for a subchain, lying between the boundaries $\tau_{\text{Rs}}$. For Rouse model, $MSD(t) \propto t^{2/3 + d_F}$, where $d_F$ is the fractal dimension of the chain [9]. For chain in athermal solvent $d_F \approx 5/3$, therefore, $MSD(t) \propto t^{6/11}$ and the Rouse reptation time of the chain in good solvent is $\tau_R = \tau_0 N^{2.2}$. From $MSD(t)$ measurements we obtained $\tau_0 \approx 700$ - number of DPD time steps, sufficient for a bead reptation on its size (bead size was estimated from $R(s)$ dependency as $R(1) \approx 0.6$). Hence, we have $\tau_R \approx 2.8 \times 10^9$ DPD steps and $\tau_{\text{Rs}} \approx 1.8 \times 10^8$ DPD steps. Number of simulation steps was set to $1.5 \times 10^8$ DPD steps. This value was chosen to be much larger than the average bond lifetime and of the same order of magnitude as $\tau_{\text{Rs}}$. Initial conformation, from which the simulations started, was generated as a confined random walk inside the impermeable simulation box. Therefore, the initial structure was close to the equilibrium state of a confined chain in good solvent. Hence, by choosing the almost equilibrated initial conformation and sufficient simulation time, we avoided unnecessary effects of the initial conformation on the studied final state (Supplementary Fig. S5). After evolution of the system during $1.5 \times 10^8$ DPD steps, we performed simulations for $0.5 \times 10^8$ DPD steps to calculate $P(s)$, $R(s)$ and $MSD(t)$. Each dependency was averaged over a period of time equal to 1000 bond lifetimes. This choice ensures, that every bead in the chain with length $N = 1000$ was able to form a bond with $N - 1$ other beads in the chain.

We performed simulations of a longer confined chain $N = 20000$ (Supplementary Fig. S9) as well to analyze the role of chain length and subchain overlapping (details in Supplementary). Simulation box size was set to $52 \times 52 \times 52$ to keep the polymer concentration $c \approx 0.06$ as for the chain length $N = 1000$. The other parameters: solvent quality, simulation time, $\rho$, $R_c$, $N_{\text{stp}}$ and $\beta$ were set as for the chain length 1000.

Additionally, we performed simulations of the free (non-confined) chain ($N = 1000$) to obtain information about the effect of impermeable boundaries on the formed states. Periodic boundary conditions were applied and adjusted to make $P(s)$ of the free chain without reversible bonds ($f = 0$) close to the theoretical prediction $P(s) \propto s^{1.93}$, simulation box size was set to $38 \times 38 \times 38$ (Supplementary Fig. S10). Solvent quality, simulation time, $\rho$, $R_c$, $N_{\text{stp}}$ and $\beta$ was set as for the confined chain.

We also compare the collapse of a chain, induced by reversible bonds formation, with coil-globule transition in the confined chain length $N = 1000$, caused by the change of solvent quality (Supplementary Fig. S6). Flory-Huggins parameter was calculated from DPD coefficients according to the Ref. [19]. In particular, we calculated $P(s)$, $R(s)$ (average spatial distance between beads, separated by $s$ beads along the chain) and $MSD(t)$ dependencies, as well as analyzed time evolution of contact domains in the globule in poor solvent ($\chi \approx 1.4$), which radius of gyration equals to the radius of gyration of the chain with high fraction of bonds ($f = 0.97$) to compare the characteristics of the two compact states.

**RESULTS**

![Graph showing $R(s)$ dependencies. Snapshots of swollen and compact states are included.](image)

First we discuss the swollen chain state. Swollen chain in the impermeable box breaks into linearly uncorrelated subchains, lying between the boundaries. The concentration chosen in the simulation ensures (details in Supplementary), that dilute solution of overlapping subchains was formed. For simplicity we will further apply formalism, applicable for semi-dilute solutions, although the concentration in the system is only slightly higher, than the overlap concentration (see Supplementary for the discussion of the effect of subchain overlapping strength). There are three characteristic scales in the swollen state (Fig. 1, black solid lines): concentration in the system is only slightly above the overlap concentration ($\chi \approx 1.4$), which radius of gyration equals to the radius of gyration of the chain with high fraction of bonds ($f = 0.97$) to compare the characteristics of the two compact states.
transition occurs at \( f \approx 0.97 \), than in the globule in poor solvent. Due to the finite size of the globule in poor solvent, transition from \( P(s) \propto s^{-1.5} \) to plateau is too wide and the equilibrium dependency \( P(s) \propto s^{-1.5} \) is observed only for small \( s \leq 10 \). Due to large \( P(s) \) error values for large \( s \), error bars are partially removed to not to obstruct the plot.

![Diagram](image.png)

**FIG. 2:** \( P(s) \) dependencies. For small and intermediate lengths \( s \) contact probability is higher in chain with \( f \approx 0.97 \), than in the globule in poor solvent. Due to the finite size of the globule in poor solvent, transition from \( P(s) \propto s^{-1.5} \) to plateau is too wide and the equilibrium dependency \( P(s) \propto s^{-1.5} \) is observed only for small \( s \leq 10 \). Due to large \( P(s) \) error values for large \( s \), error bars are partially removed to not to obstruct the plot.

If we increase the fraction of bonds \( f \), coil-globule transition occurs at \( f \approx 0.3 \) (see Supplementary Fig. S10). Chain with this fraction of bonds remains effectively Gaussian, and can be described by equilibrium dependencies for ideal chains \( P(s) \propto s^{-1.5} \) and \( R(s) \propto s^{0.5} \) (Fig. 2 Supplementary Fig. S10, \( f = 0.3 \) case). Chain with \( f \approx 0.3 \) has Rouse chain dynamics due to the small enough length of the chain and weakness of entanglement effects, \( MSD(t) \propto t^{0.5} \) (Fig. 4 Supplementary Fig. 10, \( f = 0.3 \) case). After increase of the fraction of bonds up to \( f \approx 0.5 \) chain forms a Gaussian globule state, as shown in \( P(s) \) dependency for the free chain (Supplementary Fig. S10). The aforementioned Gaussian chain \( P(s) \) dependency remains on the scale, less than the Gaussian subchain length in the globule \( s \propto N^{2/3} \approx 100 \) for chain length \( N = 1000 \). Interestingly, dynamics, observed in yeast, is close to the Rouse dynamics [24]. \( P(s) \) and \( R(s) \) on 0.1 - 1 Mbp scale in yeast have Gaussian behavior [9]. Similar \( R(s) \) scaling was observed in murine chromatin [29].

If the fraction of bonds \( f \approx 0.6 \) and higher is formed, the chain transfers into a kinetically "frozen" state, which is out of equilibrium and similar to the fractal globule [1]. This statement is confirmed by \( P(s) \propto s^{-1} \) and \( R(s) \propto s^{1/3} \) dependencies characteristic of the fractal globule (Fig. 2 Supplementary Fig. S10 and S9). Analysis of contact maps (Fig. 3), comparison with globule in poor solvent, as well as visual inspection of conformations confirm this statement as well. However, the observed state is not strictly equivalent to the idealized fractal globule, which is characterized by \( P(s) \propto s^{-1} \) and \( R(s) \propto s^{1/3} \) on all length scales. On small length scales chain tends to form more elongated structures, than in true fractal globule. This leads to more steep \( P(s) \) and \( R(s) \) dependency for small lengths \( s \) (Fig. 2 Supplementary Fig. S10 and S9). Moreover, chain with high fraction of bonds tends to form maximally compact state with linear size \( \propto N^{1/3} \) due to the surface tension, as in equilibrium globule. Therefore, due to more sparse conformations on small scale and the finite chain size, in the observed state self-similar fractal structure is formed not on all length scales, but within "fractal domains" on length scale less than the maximally compact globule size \( s < N^{2/3} \approx 100 \). On larger scale \( s \) the domains become partially mixed, which results in plateau in \( P(s) \), \( R(s) \) and \( MSD(t) \) dependencies. The similar state, crumpled globule [13], forms during coil-globule transition in poor solvent [16]. However, crumpled structure exists much longer in case of the state, formed during collapse by reversible bonds, than the crumpled globule, formed in poor solvent. In the latter state crumples quickly merge into a single globule, which then reaches equilibrium. In our simulations the "crumpled" structure does not disappear after equilibration during \( 1.5 \times 10^8 \) DPD steps (3000 average bond lifetimes). In comparison, it was shown [16], that globule with \( N = 10000 \) reaches equilibrium dependency \( R(s) \) for \( s \) up to 200 beads in \( 2 \times 10^8 \) DPD steps, which is at least 100 times faster, than for the state with high fraction of formed reversible bonds. We suppose, that the self-similar fractal structure does not disappear so long due to "caging" of beads: after breaking
of bond free beads are constrained in the "crumple" and mostly can form bonds only with the surrounding beads inside the "crumple". Therefore, the structure remains long-living.

Naturally, fractal domains are also regions of high contact frequency. To analyze the structure of the chain with high fraction of bonds in detail, we used Armatus software [26] to analyze contact maps (details of contact maps generation in Supplementary Fig. S8) and determined the borders of the contact domains. This analysis shows how borders of domains change with time for the chain with \( f = 0.97 \) and for the globule in poor solvent, equilibrated during \( 1.5 \times 10^8 \) DPD time steps, chosen as a reference structure (see Fig. 3). It is clear, that domain borders remain unperturbed much longer in the latter case, than in the globule in poor solvent. In the latter state the domains start to change their location along the chain on the time-scale, approximately two orders of magnitude shorter, than in the chain with \( f = 0.97 \). This shows the absence of long-living crumpled structure in equilibrated globule, as expected. On the other side, large and long-living (compared to the globule in poor solvent) contact domains exist in the chain with high fraction of bonds \((f = 0.97)\). We also point out highly non-homogeneous sizes of domains in the chain with \( f = 0.97 \). We used the parameter \( \gamma = 0 \). The same analysis for 4 more structures with \( f = 0.97 \), simulated from different initial conformations, is given in Supplementary Fig. S8 to demonstrate the persistence of the "crumpling" effect.

Self-similar structure of the fractal domains is stabilized by saturating bonds and not by topological constraints [27], so it leads to several non-trivial effects. First, it is the deviation of the system dynamics from the fractal globule dynamics [9] (Fig. 4, Supplementary Fig. S10 and S9). Second, large number of bonds leads to formation of the self-similar structure even on the length scale \( s \approx 10 \) (Fig. 2 Supplementary Fig. S10 and S9). In contrast to the unkotted globule, in which self-similar fractal structure forms due to the effect of topological constraints on the scale, several times larger, than the entanglement length [27]. We note, that the aforementioned \( P(s) \propto s^{-1} \) behavior is observed in mammals and Drosophila (Supplementary Fig. S11 and S7). Also in the work [2] the transition from \( R(s) \propto s^{0.33} \) to \( R(s) \propto s^{0.2} \) was observed on the \( \approx 7Mbp \) scale (Fig. 1). Similar \( R(s) \) scaling was observed in human male fibroblast cells [25].

As it was mentioned above, the non-equilibrium state, formed by reversible bonds, has a specific dynamic behavior (Fig. 4 Supplementary Fig. S10 and S9). We observe ballistic regime, universal for all chains, in the shortest time-scale. On the larger scale, we observe motion of beads inside the fractal domains. Therefore, for \( f = 0.97 \) case, the dynamics is similar to the fractal globule dynamics with \( MSD \propto t^{0.4} \) (Fig. 4 Supplementary Fig. S10 and S9). As time increases, we

---

**FIG. 3:** Distance maps of structures, obtained after equilibration during \( 1.5 \times 10^8 \) DPD steps (see "Methods" section) are shown above the domain borders maps (black and white plots). A row of black stripes shows positions of domain borders along the chain after \( t \in [10^5 \ldots 10^7] \) DPD steps passed after equilibration. a) Chain with high fraction of reversible bonds, \( f = 0.97 \). b) Globule in poor solvent. Characteristic domain size and lifetime \( \tau \) is shown on both plots.
start to observe motion of long-living fractal domains. Therefore, on this time-scale, beads participate in collective motion as a part of a domain. Due to the relatively large sizes of domains and correlated motion of beads inside the domains, we observe rapid subdiffusion with $MSD \propto t^{0.8}$, in contrast to the Rouse dynamics of swollen states and the globule in poor solvent (Fig. 4, Supplementary Fig. S10 and S9). Note, that the observed subdiffusion is slower, than Brownian motion, because during the average lifetime of a domain, its motion is constrained due to the presence of bonds between the domains and the chain structure of the polymer. Interestingly, $MSD(t)$ error sufficiently increases in this regime, indicating highly non-homogeneous motion. The existence of domains in the time-scale of the fast subdiffusion is confirmed by contact maps analysis (Fig. S3 Supplementary Fig. S8). Structure with domains resembles the pearl-like state, observed during collapse by irreversible bonds formation [29]. This motion structure differs from the fractal globule, induced by topological constraints, where much slower anomalous subdiffusion is observed with $MSD \propto t^{0.4}$ on a long time-scale [30]. We note, that previously unexplained transition from fractal globule subdiffusion with $MSD \propto t^{0.4}$ to much faster anomalous subdiffusion on a relatively long time-scale was observed in mammals [11–13].

As we mentioned in Introduction, the structure of a single TAD is successfully described by different models. Moreover, $P(s)$ scaling inside TADs is radically different, than $P(s)$ on larger scale [4, 32]. We propose a novel model for large-scale chromatin organization, choosing TAD as a basic interacting unit. We describe chromatin as a coarse-grained flexible homopolymer chain in athermal solvent. One bead in our model chain roughly represents a single TAD. The reversible bonds are treated as a long-living inter-TAD contacts (we consider them as pairwise interactions and neglect triple and more complex inter-TAD contacts). There are several biological mechanisms underlying this interactions, for example interactions through insulator proteins, such as cohesin [8, 34]. We emphasize the reversible pairwise nature of the aforementioned interactions, so we introduce this type of bonds in our model. By changing the number of pairwise reversible inter-TAD contacts, the model reproduces different states of chromatin with different static and dynamic properties. $P(s)$, $R(s)$ and $MSD(t)$ dependencies obtained for these states are observed experimentally on large scales in chromatin for different species. Each of the states exists in a rather wide range of the fraction of bonds. The reversible bonds are treated as a single TAD. The reversible bonds are treated as a long-living inter-TAD contacts (we consider them as pairwise interactions and neglect triple and more complex inter-TAD contacts). There are several biological mechanisms underlying this interactions, for example interactions through insulator proteins, such as cohesin [8, 34]. We emphasize the reversible pairwise nature of the aforementioned interactions, so we introduce this type of bonds in our model. By changing the number of pairwise reversible inter-TAD contacts, the model reproduces different states of chromatin with different static and dynamic properties. $P(s)$, $R(s)$ and $MSD(t)$ dependencies obtained for these states are observed experimentally on large scales in chromatin for different species. Each of the states exists in a rather wide range of the fraction of bonds, therefore, the model is resistant to fluctuations of parameters. We can also change the size of the plateau in $R(s)$, $P(s)$ and $MSD(t)$ by adjusting the chain length and the impermeable box size.

Interesting, that pairwise reversible interactions can not cause phase transition in gas, however, in polymer chains this type of interactions leads to the first-order coil-globule phase transition [35]. In general, pairwise reversible interactions are abundant in living systems:
sulfide bridges and hydrogen bonds play important role in protein folding and DNA double-stranded structure formation. There are a lot of mechanisms of pairwise reversible interaction in chromatin as well, depending on the type of chromatin [28].

We can not reproduce ensemble-averaged Hi-C maps due to the homopolymer nature of the model, because we neglected the difference between interactions of different types of chromatin [28]. Moreover, rather simple confinement conditions in our model can not reproduce effects, connected with sophisticated lamina structure, as well as reproduce effects, connected with lamina structure, is the matter of future research.

Acknowledgments

We thank M.V. Tamm for fruitful discussions and comments, A.A. Gavrilov for providing DPD computer code and A.A. Galitsyna for the experimental data and useful remarks. The research of P.I.K. is supported in part by Skoltech Systems biology fellowship and grant of the Foundation for the advancement of theoretical physics “Basis”. The reported study was funded by RFBR according to the research project # 18-29-13041.

References

[1] E. Lieberman-Aiden, N. L. Van Berkum, L. Williams, M. Imakaev, T. Ragoczy, A. Telling, I. Amit, B. R. Lajoie, P. J. Sabo, M. O. Dorschner, et al., science 326, 289 (2009).
[2] S. Wang, J.-H. Su, B. J. Beliveau, B. Bintu, J. R. Moffitt, C.-t. Wu, and X. Zhuang, Science , aaf8084 (2016).
[3] J. D. Halverson, J. Smrek, K. Kremer, and A. Y. Grosberg, Reports on Progress in Physics 77, 022601 (2014).
[4] T. Sexton, E. Yaffe, E. Kenigsberg, F. Bantignies, B. Leblanc, M. Hoichman, H. Parrinello, A. Tanay, and G. Cavalli, Cell 148, 458 (2012).
[5] Y. Zhang, R. P. McCord, Y.-J. Ho, B. R. Lajoie, D. G. Hildebrand, A. C. Simon, M. S. Becker, F. W. Alt, and J. Dekker, Cell 148, 908 (2012).
[6] H. Wong, H. Marie-Nelly, S. Herbert, P. Carrivain, H. Blanc, R. Koszul, E. Fabre, and C. Zimmer, Current biology 22, 1881 (2012).
[7] M. Barbieri, M. Chotalia, J. Fraser, L.-M. Lavitas, J. Dostie, A. Pombo, and M. Nicodemi, Proceedings of the National Academy of Sciences 109, 16173 (2012).
[8] G. Fudenberg, M. Imakaev, C. Lu, A. Goloborodko, N. Abdelnour, and L. A. Mirny, Cell reports 15, 2038 (2016).
[9] M. V. Tamm, L. I. Nazarov, A. A. Gavrilov, and A. V. Chertovich, Physical review letters 114, 178102 (2015).
[10] S. Shinkai, T. Nozaki, K. Maeshima, and Y. Topgash, PLoS computational biology 12, e1005136 (2016).
[11] I. Bronstein, Y. Israel, E. Kepten, S. Mai, Y. Shav-Tal, E. Barkai, and Y. Garini, Physical review letters 103, 018102 (2009).
[12] K. Burnecki, E. Kepten, J. Janczura, I. Bronshtein, Y. Garini, and A. Weron, Biophysical journal 103, 1839 (2012).
[13] I. Bronshtein, E. Kepten, I. Kanter, S. Berezin, M. Lindner, A. B. Redwood, S. Mai, S. Gonzalez, R. Foisner, Y. Shav-Tal, et al., Nature communications 6, 8044 (2015).
[14] L. A. Mirny, Chromosome research 19, 37 (2011).
[15] A. Y. Grosberg, S. K. Nechaev, and E. I. Shakhnovich, Journal de physique 49, 2095 (1988).
[16] A. Chertovich and P. Kos, The Journal of chemical physics 141, 134903 (2014).
[17] S. V. Ulianov, E. E. Khrameeva, A. A. Gavrilov, I. M. Flyamer, P. Kos, E. A. Mikhailova, A. A. Penin, M. D. Logacheva, M. V. Imakaev, A. Chertovich, et al., Genome research 26, 70 (2016).
[18] D. Jost, P. Carrivain, G. Cavalli, and C. Vaillant, Nucleic acids research 42, 9553 (2014).
[19] R. D. Groot and P. B. Warren, The Journal of chemical physics 107, 4423 (1997).
[20] P. Nikunen, I. Vattulainen, and M. Karttunen, Physical Review E 75, 036713 (2007).
[21] M. Daoud, J. Cotton, B. Famoux, G. Jannink, G. Sarma, H. Benoit, C. Duplessix, C. Picot, and P. De Gennes, Macromolecules 8, 804 (1975).
[22] P. Heun, T. Larote, K. Shimada, P. Furrer, and S. M. Gasser, Science 294, 2181 (2001).
[23] J. Vazquez, A. S. Belmont, and J. W. Sedat, Current Biology 11, 1227 (2001).
[24] G. G. Cabal, A. Genovesio, S. Rodriguez-Navarro, C. Zimmer, O. Gadal, A. Lesne, H. Buc, F. Feuerbach-Fournier, J.-C. Oliovo-Marin, E. C. Hurt, et al., Nature 441, 770 (2006).
[25] L. S. Shopland, C. R. Lynch, K. A. Peterson, K. Thornton, N. Kepper, J. von Hase, S. Stein, S. Vincent, K. R. Molloy, G. Kreth, et al., The Journal of cell biology 174, 27 (2006).
[26] D. Filippova, R. Patro, G. Duggal, and C. Kingsford, Algorithms for Molecular Biology 9, 14 (2014).
[27] M. V. Imakaev, K. M. Tchourine, S. K. Nechaev, and L. A. Mirny, Soft matter 11, 665 (2015).
[28] C. Münkell, R. Eils, S. Dietzel, D. Zink, C. Mehring, G. Wedemann, T. Cremer, and J. Langowski, Journal of molecular biology 285, 1053 (1999).
[29] V. F. Scolari, G. Mercy, R. Koszul, A. Lesne, and J. Mozziconacci, arXiv preprint arXiv:1804.04099 (2018).
[30] P. R. Langer-Safer, M. Levine, and D. C. Ward, Proceedings of the National Academy of Sciences 79, 4381 (1982).
[31] I. Williamsonson, S. Berlivet, R. Eskeland, S. Boyle, R. S. Illingworth, D. Paquette, J. Dostie, and W. A. Bickmore, Genes & development 28, 2778 (2014).
[32] A. L. Sanborn, S. S. Rao, S.-C. Huang, N. C. Durand, M. H. Huntley, A. I. Jewett, I. D. Bochkov, D. Chinnappan, A. Cutkosky, J. Li, et al., Proceedings of the National Academy of Sciences 112, E6456 (2015).
SUPPLEMENTARY MATERIALS

Calculation of $P(s)$, $R(s)$ and $MSD(t)$

To calculate $P(s)$ we counted $N_i$ - the number of beads, which have less distance to the $i$-th bead, than cutoff radius = 1.5, and then averaged $N_i$ over all measurements:

$$P(s) = \frac{1}{N-s} \frac{1}{5 \times 10^7} \sum_{t_0=15 \times 10^7}^{20 \times 10^7} \sum_{i=1}^{N-s} N_i(t_0)$$

Here, $15 \times 10^7$ is the simulation time, measured in DPD steps. We time-averaged the $P(s)$ dependency during 1000 average bond existence times, $5 \times 10^7$ DPD steps (see "Methods" section).

In $R(s)$ we calculated average spatial distance between beads, separated by $s$ beads along the chain. Averaging over all measurements was performed as for the $P(s)$ dependency:

$$R(s) = \frac{1}{N-s} \frac{1}{5 \times 10^7} \sum_{t_0=15 \times 10^7}^{20 \times 10^7} \sum_{i=1}^{N-s} ((x_i(t_0) - x_{i+s}(t_0))^2 + (y_i(t_0) - y_{i+s}(t_0))^2 + (z_i(t_0) - z_{i+s}(t_0))^2)^{1/2}$$

We calculated $MSD(t)$ as follows: we distracted displacement of the center of mass of the chain (COM) from displacement of a bead, and then averaged this displacement over all measurements:

$$MSD(t) = \frac{1}{N} \frac{1}{5 \times 10^7 - t} \sum_{t_0=15 \times 10^7}^{20 \times 10^7-t} \sum_{i=1}^{N} ((x_i(t+t_0) - x_{COM}(t+t_0))^2 + ((y_i(t+t_0) - y_{COM}(t+t_0))^2 + ((z_i(t+t_0) - z_{COM}(t+t_0))^2 - ((x_i(t_0) - x_{COM}(t_0))^2 + ((y_i(t_0) - y_{COM}(t_0))^2 + ((z_i(t_0) - z_{COM}(t_0))^2$$

Effect of subchain overlapping

Let us roughly represent the chain in good solvent in a simulation box with impermeable boundaries as a self-avoiding walk in 3D space (first we neglect the effects, connected with screening of volume interactions). In this approximation length of a subchain, lying between boundaries, is given as $N_{subchain} \approx \left(\frac{D}{a}\right)^{5/3}$ ($D$ - linear size of the confinement volume, $a$ - mean distance between beads). In the case of the chain length $N = 1000$ we chose $D = 18$ (see "Methods" section), and figured out the bond length $a \approx 0.6$ from the $R(s)$ dependency ($R(s = 1) \approx 0.6$, raw $R(s)$ dependencies not shown). Hence, the overlap concentration of subchains is $c^* \approx a^{-3} N^{-4/5}_{subchain} \approx 0.05$. As we discussed in "Methods" section, the polymer concentration was set to $n \approx 0.06$. Therefore, $n$ is of the order of the overlap concentration $n \approx 0.06 = O(c^*)$ in this approximation. However, we neglected effects, connected with volume interactions screening. It is known, that screening effect is maximal for systems with small polymer concentration, if chains strongly overlap. This occurs in the so-called semi-dilute solution state. In this state, subchain length is of the order of $(\Phi^{1/4} (\frac{D}{a})^2)$ (see "Results" section), therefore, $c^* \approx a^{-3} N^{-4/5}_{subchain} \approx a^{-2} D^{-8/5} n^{-1/5} \approx 0.047$. Here $\Phi \propto a^3 n$ - volume fraction of a polymer. Hence, we estimated the overlap concentration $c^*$ in two extreme cases - without any screening of volume interactions and with strong screening in semi-dilute solution, and obtained similar results. Hence, subchains, lying between the boundaries of the impermeable simulation box, do overlap, but semi-dilute solution state is not reached, because the condition for semi-dilute solution formation $c^* << n << 1$ is not satisfied.

For the confined chain length $N = 20000$, the condition for semi-dilute solution state formation $c^* << n << 1$ is, of course, satisfied due to the much larger subchain length. However, dynamic and static behavior (Fig. 10) is analogous to those characteristics for the chain length $N = 1000$. Moreover, simulations of the free chain show (Fig. 10), static and dynamic scalings for different $f$ are similar to the confined chains, except the plateau presence in swollen states. Hence, we can conclude, that overlapping regime does not affect strongly the formation of states, discussed in the paper.

Additional plots
FIG. 5: $P(s)$, $R(s)$ and $MSD(t)$ dependencies for the chains with $f = 0.97$, simulated from different initial conformations - 4 structures are simulated from different random walk conformations, and 1 structure was simulated from the initial conformation, which was initially equilibrated for $1.5 \times 10^8$ DPD steps in athermal solvent without reversible bonds formation. We show, that no significant difference between dependencies is observed for the chains with $f = 0.97$. Therefore, there is no effect of initial conformation on resulting states for any fraction of bonds $f$. a) $P(s)$ dependencies. b) $R(s)$ dependencies. c) $MSD(t)$ dependencies.
FIG. 6: (a) Dependency of the radius of gyration of the chain ($N = 1000$) on the fraction of reversible bonds. Coil-globule transition occurs at $f \approx 0.3$. (b) Dependency of the radius of gyration of the chain ($N = 1000$) on the solvent quality (Flory-Huggins parameter $\chi$). Phase transition occurs at $\chi \approx 0.9$. 
FIG. 7: Distribution of reversible bonds in the chain length $N = 1000$ with $f = 0.97$. Every point is the average number of reversible bonds formed between beads, lying $s$ beads along the chain far from each other, divided by the total number of formed reversible bonds in the chain.
FIG. 8: a)-d) Domain structure evolution as determined by Armatus software in chains with high fraction of bonds, $f = 0.97$, simulated from different initial conformations (Fig. (d) - from initial conformation, which was equilibrated without bonds formation, see "Methods" section). A row of black vertical stripes shows positions of domain borders along the chain after $t \in [10^3 - 10^7]$ DPD steps passed after equilibration. e)-h) Distance maps of structures, obtained after system equilibration for $1.5 \times 10^8$ DPD steps starting from different initial conformations (Fig. (h) - from initial conformation, which was equilibrated without bonds formation, see "Methods" section). Contact maps for Armatus analysis were calculated from distance maps according to the law $p_{ij} = 1/r_{ij}^3$, where $p_{ij}$ - contact probability, $r_{ij}$ - distance, obtained from the distance maps.
FIG. 9: a) $P(s)$ for chains with different fraction of bonds ($f$), $N = 20000$. b) $R(s)$ for chains with different fraction of bonds ($f$), $N = 20000$. c) $MSD(t)$ for chains with different fraction of bonds ($f$), $N = 20000$. Due to much stronger subchain overlapping regime than in the case of the confined chain length $N = 1000$ (see "Effect of Subchain Overlapping" section), dynamics in the swollen state agrees with Rouse dynamics for chains in $\Theta$ solvent due to volume interactions screening.
FIG. 10: a) $P(s)$ for the free chain ($N = 1000$) with different fraction of bonds ($f$). $P(s)$ of the free chain without bonds ($f = 0$) is close to the theoretical prediction for self-avoiding walk in 3D space $P(s) \propto s^{1.93}$. b) $R(s)$ dependencies for the free chain ($N = 1000$) with different fraction of bonds $f$. c) $MSD(t)$ dependencies for the free chain ($N = 1000$) with different fraction of bonds $f$. 
Contact probability (normalized)
Genomic distance ($10^4$ bp)

Danio 8hpf
Danio 24hpf
Drosophila DmBG3-c2
Drosophila Kc167
Drosophila OSC
Drosophila S2
Homo A549
Homo HEK293
Homo K562
Mus B(bonemarrow)
Mus Ba(bonemarrow)
Mus E14Tg2a(embryos)
Mus PSC(embryo)

FIG. 11: The figure shows the experimental contact probability dependencies for different species and cell lines. Blue dashed line corresponds to the equilibrium gauss globule and red dashed line corresponds to the crumpled (fractal) globule. Well-known dependence of the contact probability on genomic distance $P(s) \sim s^{-1}$ for Homo sapiens actually is not strictly enforced. For example, contact probability depends on considering scales and it works for another biological species as well. Moreover, scalings for different cell lines of the same species could significantly differ from each other on the same scale. Apparently minor protocol changes in Hi-C experiment, different equipment and different data processing have a significant impact on the final results. To plot this figure we used publicly available Hi-C datasets re-analyzed with distiller software with standard parameters (https://github.com/mirnylab/distiller-nf). All replicates, if available, were merged together. The data for Homo sapiens was retrieved for: K562 cell line (GEO ID GSE63525, https://www.ncbi.nlm.nih.gov/pubmed/25497547), A549 GSE105600, https://www.ncbi.nlm.nih.gov/pubmed/22955616, HEpG2 GSE105381, https://www.ncbi.nlm.nih.gov/pubmed/22955616, Drosophila melanogaster cell lines Kc167, Dm3, OSC, S2: GSE69013 https://www.ncbi.nlm.nih.gov/pubmed/26518482; Danio rerio embryos GSE105013 https://www.ncbi.nlm.nih.gov/pubmed/29972771; Mus musculus data from GSE96611 https://www.ncbi.nlm.nih.gov/pubmed/29335546 and from GSE96107 https://www.ncbi.nlm.nih.gov/pubmed/29053968. This analysis was performed by Aleksandra Galitsyna.