Manuscript submitted to **Biophysical Journal**

Article

The intriguing dynamics of chromatin folding and assembly

Z. Yang¹, W. Miao², and M. R.K. Mofrad^{*}

^{*} Correspondence: mofrad@berkeley.edu

Z. Yang and M. R.K. Moftad

ABSTRACT We investigate the dynamics of chromatin folding based on the "strings and binders" (SBS) model with molecular dynamics simulation. SBS model is a coarse-grained model considering a self-avoiding chain interacting with diffusive binders. By introducing transition among different categories of beads with specific transition cycles and transition probabilities, our model is capable of introducing different dynamics quantitatively during the folding process, thus capturing variety of phenomena related to chromatin dynamics. Firstly, roles of dynamics in the process of chromatin folding were examined. We discovered that there is a minimum gyration of chromatin under varying characteristic times of transition which indicates neither dramatically dynamic nor static folding process is optimal for chromatin to reach stable states with relatively low free energy. Secondly, it is noticeable that when beads transit from or into others in distinct dynamics, the equilibrium concentrations are distinct as well. As a consequence, the distribution of chromatin loop length is relevant to the dynamics of binders which can be modified by complex such as Wings apart-like protein homolog (Wapl) and SCC2/SCC4 cohesin loader complex (SCC2/SCC4). Finally, our model is able to reproduce contact matrices of both wild type HAP1 cell and Δ WAPL HAP1 cell obtained from Hi-C technology with a relatively high accuracy. Our model recapitulate the accumulating contacts at the corners of TADs and vanishing short-range contacts along the diagonal, manifesting the difference of chromatin structures before and after eliminating WAPL.

STATEMENT OF SIGNIFICANCE

Our model includes reciprocal transition among beads in SBS model to introduce different dynamics in chromatin folding process. Our model is able to examine the roles of dynamics in chromatin folding, reveal the loop length variation due to the concentration imbalance caused by distinct dynamics and reproduce contact matrices of both wild type and WAPL-deficient cells. Our research work provides a model to investigate the dynamics of chromatin folding quantitatively and displays its significance of revealing multiple experimental results using computational tools.

INTRODUCTION

Three-dimensional structure of genomes inside the nucleus is a long-standing interest of biologists(1), which is essential to gene regulation(2–4). However, the mechanism behind folding and unfolding of the chromatin inside the nucleus is still under heated debate. In order to shed light on the formation of 3D chromatin structure, chromosome conformation capture (3C)-based technologies which reveal the interaction between gene loci by measuring spatial affinity in vitro have been developed in recent decades(5–8). Thanks to the rapid development of experimental technologies, the roles of multiple regulators such as cohesin, CCCTC-binding factor (CTCF), Wings apart-like protein homolog (Wapl), SCC2/SCC4 cohesin loader complex (SCC2/SCC4) in the process of assembly and disassembly of 3D chromatin architecture have been revealed (9–19). The collaboration among those regulators results in distinct dynamics of the chromatin structures(20–23).

The ring-shaped protein complex cohesin plays essential roles in the spatial organization of chromosome. Cohesin acts as a loop-extruding factor, stretching progressively larger loops to allow long-range interaction inside the chromatin. The extruding process is inclined to stalling at the boundary of topologically associated domains (TAD) due to the interaction with boundary elements. The folding process of chromatin through loop extrusion is dramatically dynamic with frequent binding and unbinding events considering both CTCF and cohesin(23). In addition, Wapl known as the antagonist of loop-extruding factors imposes an impact on the dynamic structure of chromatin by releasing cohesin from the chromatin to restrict the loop length extension(16, 21). On the other hand, SCC2/SCC4 (known as the activator) is capable of promoting the processivity of cohesin, increasing the probability of long-range interactions(16, 19). The antagonists and activators together further influence the dynamics of chromatin structure.

With the boom of experimental data such as Hi-C contact matrices, computational models based on polymer physics have been built to inversely produce the 3D structure of chromatin(24). Nicodemi et al. developed a simple polymer physics model named "strings and binders switch" (SBS) model to investigate the architectural pattern of chromatin(25, 26). Barbieri et al.(27) further utilized the model to recapitulates various experimental observation including the scaling properties of chromatin folding, the fractal state of chromatin, the processes of domain formation and looping out. One recent study(28) implemented the SBS model to explain large scale behavior of chromatin. Combined with the PRISMR algorithm developed in another work(29), the model based on polymer physics is capable of reproducing quantitative Hi-C map in a relatively high accuracy. However, these models only raised general discussion of polymer physics insights into chromatin 3D architecture and discussed little about dynamics of binders and binding sites. Here, we report a model which enables us to investigate dynamics of chromatin folding process specifically by considering quantified transition among beads of different species. The model is a straightforward extension of SBS model and the transition follows the basic rules of chemical kinetic equation. By taking the dynamic transition quantitatively into consideration, it is possible to explain the biological significance of binding and unbinding events of those key factors including CTCF and cohesin. Furthermore, we are able to prove the functions of dynamics regulator such as Wapl by comparing experimental results with our simulations.

MATERIALS AND METHODS

SBS model with distinct binders and binding sites

Our polymer model is based on a polymer physics model known as "String and Binders Switch"(SBS) model(25–27) in which a chromatin fiber is represented by a coarse-grained self-avoiding (SAW) polymer chain interacting with randomly sparse binders. In our model, we basically separated beads of polymer chain into 2 categories and binders into 3 categories according to interaction. More specifically, as is shown in Figure 1(a), green beads of polymer chain are sites having a larger attraction with binders named as "binding sites", whereas blue beads within the polymer are repulsive or weakly attractive to binders named as "regular DNA". Binding sites in our model represents DNA sequences bound with CTCF where binders reside for a longer period and regular DNA refers to DNA sequences that are not binding sites. In addition, there are three kinds of binders dynamically converting reciprocally in the model. These binders differ in the interaction with polymer chain which influence their processivity of extruding chromatin loops. The orange beads are cohesin binding with DNA chains in a regular state(without interference of activators and antagonists) which are named as regular binders (see Figure 1). Apart from regular cohesin, the cohesin activated by protein like SCC2/SCC4(16, 19) is shown as red beads in Figure 1(a) and yellow beads are representative of those cohesins released from DNA chain induced by antagonists such as Wapl. We name them as strong binders and weak binders separately.

As far as the interaction is concerned, we follow the original form of potential used in Chiariello's work except the parameters of interaction (ϵ, r_c) between beads of strings and diffusive binders(28). Under different discussions of dynamics, we varied the interaction (ϵ, r_c) and details about parameters are included under different topics in Supporting Material. The general settings used in all topics are stated below. In the model, all the beads have an identical diameter σ . Considering the interaction within the string, two consecutive beads are connected with a finitely extensible FENE bond(shown in Eq. 1 where $R_0 = 1.6\sigma$ and $K = 30K_B T/\sigma^2$) and the Weeks-Chandler-Andersen potential between any two beads of the string(shifted Lennard-Jones potential shown in Eq. 2 where $\epsilon = 1.0K_B T, r_c = 1.0\sigma$) is included to account for excluded volume effect. In addition, the beads of the string interact with diffusive binders by the Lennard-Jones potential. From Eq. 2, it is noticeable that the interaction is determined by an energy scale ϵ and a cut-off distance of r_c . When r_c is larger than $2^{\frac{1}{6}}\sigma$, the interaction between the bead on the chain and binder is attractive, and when r_c is smaller than $2^{\frac{1}{6}}\sigma$, the interaction is repulsive. Moreover, a larger r_c leads to a longer-range interaction. In our model, we set r_c of the strong binders, regular binders and weak binders as $1.5\sigma, 1.3\sigma$ and 1.0σ . As a consequence, strong binders and regular binders are attractive in different processivities whereas weak binders interact with beads of the string through repulsion. The energy scale ϵ for the interaction between binders and beads of the chain will be specified separately in different cases.

$$E_{FENE} = \frac{1}{2}KR_0^2 \ln(1 - (\frac{r}{R_0})^2) \quad (1)$$

where K is the strength of the spring and R_0 is distance constant representing the max length of FENE spring.

$$V_{LJ} = \begin{cases} 4\epsilon[(\frac{\sigma}{r})^{12} - (\frac{\sigma}{r})^6 - (\frac{\sigma}{r_c})^{12} + (\frac{\sigma}{r_c})^6] & \text{if } r \leq r_c \\ 0 & \text{if } r > r_c \end{cases} \quad (2)$$

where r_c is the cut off distance for the Lennard-Jones potential, ϵ is the energy scale of the potential.

Dynamics quantification

To introduce different dynamics in chromatin folding, we extended SBS model within the Molecular Dynamics(MD) Simulation by allowing reciprocal transition within binders or binding sites with certain transition rate k. Transition rate k is defined by transition probability $P_{i \rightarrow j}$ and transition cycle T as $k = \frac{P_{i \rightarrow j}}{T}$. Transition probability $P_{i \rightarrow j}$ is defined as the probability of transition from category i to category j. The transition cycle is the period of transition. Two parameters are utilized to define transition rate due to the implement in MD simulations. The transition scheme and related notation are displayed in Figure 1(b). Single transition from category i to category j observes Eq. 3. The characteristic time scale of the transition is $\tau = \frac{1}{k}$, which is equal to the mean time of transition time, thus being named as period.

$$\frac{N_i P_{i \rightarrow j}}{T} = -\frac{dN_i}{dt}_{i \rightarrow j} = kN_i \quad (3)$$

where N_i is the number of category i. Here dN_i only refers to the change of number from category i into category j.

Figure 1: Schematic of the proposed dynamic model (a)Representation of binders and binding sites. The green beads are binding sites and the blue beads are regular DNA. The red, orange and yellow beads are strong, regular, weak binders separately. (b)Representation of transition of binders and binding sites. Transition takes place within the SAW chain or diffusive binders. The notations of transition probabilities are based on the type of beads.

In dynamic equilibrium, the concentrations of binders and binding sites are determined by the transition rates. With fixed transition cycle, the relation between concentration of beads and transition probability satisfies simple chemical kinetic equation. Eq. 4 and Eq. 5 display the relation considering binders and binding sites separately.

$$\begin{bmatrix} P_{A \rightarrow B} + P_{A \rightarrow C} & -P_{B \rightarrow A} & -P_{C \rightarrow A} \\ -P_{A \rightarrow B} & P_{B \rightarrow A} + P_{B \rightarrow C} & -P_{C \rightarrow B} \\ -P_{A \rightarrow C} & -P_{B \rightarrow C} & P_{C \rightarrow A} + P_{C \rightarrow B} \end{bmatrix} \begin{bmatrix} C_A \\ C_B \\ C_C \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad (4)$$

$$\begin{bmatrix} P_{D \rightarrow E} & -P_{E \rightarrow D} \\ -P_{D \rightarrow E} & P_{E \rightarrow D} \end{bmatrix} \begin{bmatrix} C_D \\ C_E \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad (5)$$

where C_A, C_B, C_C, C_D, C_E are concentrations of strong binders, regular binders, weak binders, strong binding sites and weak binding sites. $P_{A \rightarrow B}, P_{A \rightarrow C}, P_{B \rightarrow A}, P_{B \rightarrow C}, P_{C \rightarrow A}, P_{C \rightarrow B}$ are transition probability indicated in Figure 1(b), satisfying $P_{A \rightarrow B} + P_{A \rightarrow C} < 1, P_{B \rightarrow A} + P_{B \rightarrow C} < 1, P_{C \rightarrow A} + P_{C \rightarrow B} < 1$.

Apart from the transition probability, the transition period T plays an important role in determining the dynamics as well. A larger period of transformation leads to a less dynamic self-assembly process according to the Eq. 3. Although it is possible to obtain same transition rate k by varying both transition probability and transition period, we fix either transition probability or the period in order to reduce ambiguity when comparing dynamics.

RESULTS AND DISCUSSION

Role of dynamics

We investigated the role of dynamics in self-assembly of chromatin as dynamics is an integral part of chromatin folding, which influences various genome functions including RNA transcription and DNA replication (30). However, the general structures of chromatin under distinct dynamics are barely discussed. Figure 2(a) shows a typical example of the difference between a "static" chromatin structure (yellow line) and a "dynamic" chromatin structure (red line). More specifically, "static" refers to the situation where all transition probabilities $P_{i \rightarrow j}$ are zero, thus characteristic time scale τ (defined in Materials and Methods) being infinite, whereas "dynamic" refers to the situation where τ is finite. As the yellow line in Figure 2(a) indicates, without dynamic transition, the chromatin falls into a metastable state (gyration around 16) quickly and is stuck all over the simulation. However, a dynamic structure of chromatin is more likely to jump out of a metastable state and fall into a state with lower free energy as the gyration shown by red line in Figure 2(a) drops more than once. The snapshot of chromatin structure in the middle is a metastable state (See Figure 2 middle panel). From the perspective of statistic mechanics, a more dynamic structure is capable of exploring more states of phase space in equal times and the probability for the system to occupy the state with a lower free energy is exponentially higher. In other words, the dynamics of self-assembly plays an important role in forming a more condensed structure of chromatin.

Figure 2: Role of dynamics(period $\tau = 0.36s, 1.11376s, 3.6s, 11.376s, 36s, 113.76s, 360s, 1137.6s, 3600s$) (a)Typical example of self-assembly showing the role of dynamics which is characterized by the gyration of chromatin filament. (b)Equilibrium gyration under different periods considering dynamics of binders and binding sites separately.

To further compare distinct dynamics quantitatively, we calculated the end gyrations of chromatin under different dynamics. The end gyration is an average gyration over last 2% frames of the simulation and the value is normalized according to the initial gyration. In order to vary the dynamics of folding process, we fixed the transition probability $P_{i \rightarrow j}$ and varied the period τ . The specific details of dynamics setting are presented in Supporting Material. We examined dynamics of binders and binding sites separately by setting transition rate of the other category to zero (setting the category static) when considering one category. For example, there is no transition between binders when we change the dynamics of binding sites and vice versa. In both cases, as the Figure 2(b) displays, there is a minimum of end gyration as the period τ varies. When period τ is small (dynamics is dramatic), though the chromatin is able to explore more states in phase space, it is more likely to jump from a state with lower free energy to a state with higher free energy as the times of transition accumulate. When the period τ is large (dynamics is minor), the situation is close to a "static" self-assembly. Consequently, the end gyration is slightly smaller than that without transition. All in all, neither dramatic dynamics or minor dynamics is optimal for chromatin to fold into a condensed structure.

Despite there is a minimum end gyration for cases of both binders and binding sites, the time scales when the minimum arises are different. As the Figure 2(b) reveals, the minimum of end gyration for binders(at $\tau = 36s$) arises at a larger period than that of binding sites(at $\tau = 11.376s$). In addition, we noticed that when the dynamics is dramatic, the end gyration for binders is even higher than that without dynamics. In other words, the dramatic dynamics ruins the process of self-assembly for binders. However, considering the binding sites, even with the period $\tau = 0.36s$, the chromatin is still more condensed than that without dynamics. These two features indicate that the transition between binding sites is more adaptable to huge

Z. Yang and M. R.K. Mofrad

dynamics than binders considering forming a close state of chromatin. The reason is that one binder (strong or regular) is always surrounded by several binding sites (DNA sequences bound with CTCF) to form many-body contact(28) given that the concentration of binding sites are higher. Even if partial binding sites are transforming into regular DNA (DNA sequences bound without CTCF), the local structure is still maintained. Therefore, transition between binding sites needs to be more dynamic to impose the same impact on the structure of chromatin as binders.

The conclusion is capable of explaining the distinct dynamics of CTCF and cohesin(23). The residence time of CTCF is ~ 1 -2min while the residence time of cohesin is ~ 22 min. Though given the coarse-grained nature of proposed model, any connection from MD simulation to real time is difficult to be established, the relative relation showing distinct dynamics of binders and binding sites is still conserved.

Furthermore, according to Figure 2(b), when the period is large or the dynamics is minor, the end gyrations for binders and binding sites become similar. The result manifests that in minor dynamics, dynamics itself becomes the unique factor in determining the conformation of chromatin. The reason is that when the dynamics is minor, both transitions of binders and binding sites have little impact on the structure of chromatin. As a result, instead of the category of beads that make transition, the transition frequency exerts predominant influence on the process of self-assembly.

Loop length variation

In the last section, we investigated the impacts of overall dynamics on self-assembly of chromatin as the dynamics for different binders and binding sites were set the same. However, in general cases, the dynamics of different binders(weak, regular and strong) or different binding sites(weak and strong) are disparate. As a consequence, the concentrations of different types of beads in dynamics equilibrium are distinct according to Eq 4, thus causing variation of loop length of chromatin. More specifically, when one class of binders or binding sites is more dynamic, the concentration of the class in dynamic equilibrium is generally lower. For instance, considering i th class, when the class is more dynamic, $P_{i \rightarrow j}$ is larger(j represents other classes except i th class). According to the Eq. 4 and Eq. 5, the concentration of i th class in equilibrium is consequently lower.

In Figure 3, we displayed the impact of distinct dynamics of binders on chromatin loop length. In order to compare different dynamics, we varied the transition probabilities of different binders while fixing the periods. Combination of larger transition probabilities for i th class of binders D_l is ($P_{i \rightarrow j} = 0.2$, $P_{i \rightarrow k} = 0.2$) where $j, k \neq i$. Meanwhile, combination of smaller transition probabilities D_s is ($P_{i \rightarrow j} = 0.1$, $P_{i \rightarrow k} = 0.1$). In our simulations, we compared three different situations. For reference, We set D_l as transition probabilities to all three classes of binders, which leads to identical concentration in equilibrium. In other two cases, we changed the combination of transition probabilities from D_l to D_s for strong binders and weak binders separately. Therefore, according to Eq. 4, the concentration ratios varied from 1:1:1 to 1:3:1 and 1:1:3(regular binders: strong binders: weak binders) respectively.

As a result of the distinct dynamics, we revealed the variation of distribution of loop length in Figure 3(a). The method we determined loop length is declared in Supporting Material. When the strong binders are more static, the concentration of strong binders in equilibrium increases while the concentrations of other two classes decrease, thus the overall curve of distribution of loop length being lifted comparing with the reference curve. When the weak binders are less dynamic, the distribution curve displayed an opposite variation. Figure 3(b) quantitatively provided the mean loop length under different dynamics. The results are corresponding to the distribution shown in Figure 3(a).

The conclusion that distinct dynamics are capable of varying the loop length indicates that the activity of protein related to cohesin such as Wapl and SCC2/SCC4 is able to control the structure of chromatin in nucleus. Therefore, self-assembly of chromatin is likely to be regulated by the activating or prohibiting certain proteins.

Reproduction of Hi-C contact matrices of wild-type cell and Wapl-deficient cell

Our models were able to reproduce Hi-C contact matrices of wild-type cell and Wapl-deficient cell by varying dynamics of different binders. We selected segment(142.5Mb - 144.1Mb) of chromosome 2 of HAP1 cell whose Hi-C contact matrices has been detailed investigated (16, 31) and widely used for modeling and building new methods (32, 33).

In order to reproduce the process of folding of the chromatin segment we selected, the binding sites(CTCF sites) were assigned based on the ChIP-seq data(data accessible at NCBI GEO database(16), accession GSE95015) instead of using Hi-C map to optimize the positions of binding sites. Combined with information about TADs indicated by Hi-C contact map shown in Figure 4 (a) and (b), we split 9 categories of regular DNA and 10 types of binding sites to interact with 7 different binders(each binder interacts with one TAD). Details about the positions of binding sites and the methods we separated the beads are presented in Supporting Material. In this part, to simplify the situation, we only considered regular binders(type B) and weak binders(type C) because these two binders are sufficiently capable of representing binding(regular binders) and unbinding(weak binders) states given that Wapl is a release factor of binders (16). According to recent discoveries about the chromatin folding

Figure 3: Loop length variation due to distinct dynamics of binders(concentration ratio(regular binders: strong binders, weak binders) = 1:1:1, 1:3:1, 1:1:3) (a)Distribution of loop length under different concentration ratios. (b)Mean loop length under different concentration ratios.

mechanism(10, 12, 20, 33, 34), chromatin loops are extruded to form chromosomal domains until loop extrusion process is delayed by binding sites(extrusion model). In other words, the binder is capable of attaching itself to any part of the chromatin chains during the extrusion. However, there is a larger chance that the binder is bound with binding sites when counting the contact frequency. As a result, we assigned the binders to be attractive to both binding sites and regular DNA but the attraction is stronger between binders and binding sites. In addition, we only considered the intra-TAD interaction except TAD 5. As a result, there is basically no contacts outside the triangles in Figure 4 (a). The reasons are as follows: firstly, the inter-TAD interaction is much weaker than the intra-TAD interaction; secondly, the interaction we neglected is in the regions close to the boundaries of matrices which is relatively difficult to assign because it is more likely to be influenced by segments of chromatin beyond the segment we chose. These two reasons also explained why we included TAD 5 because the interaction is relatively strong with accumulating contacts at its corner after deleting Wapl and it is considerably independent of other segments due to the fact that it is in the middle of the map.

Figure 4 (a) and (b) revealed that erasing Wapl in cells changed the contacts distribution of chromatin by increasing the contacts at the corner of TADs and decreasing intra-TAD contact frequency(16). In order to realize the influence caused by deleting Wapl, we decreased the transition probability from regular binders to weak binders by 9 times compared with wide type cells. We assumed that the binding events were not influenced by deleting Wapl. As a consequence, the transition probability from weak binders to regular binders remain the same when simulating chromatin folding in Δ Wapl cell. Details about the simulation parameters including transition probabilities are declared in Supporting Material.

According to Figure 4 (a) and (c), the model we constructed for chromatin folding of the segment of chromosome 2 of HAP1 wild type cell show great accuracy in predicting the contacts in the triangle considering we do not use any statistical methods or optimization methods for parameters. In order to demonstrate the accuracy, we used Pearson correlation r to compare the similarity of simulated contact matrix and experimental Hi-C contact matrix(29). Though we used a extremely simplified model, Pearson correlation between two maps(Figure 4 (a) and (c)) is approximately 0.92. Furthermore, through decreasing the dynamics of regular binders, we are able to simulate the contacts matrices in Δ Wapl cells. As the Figure 4 (d) displays, contacts accumulate at the corner of TADs. The reason why contacts become more frequently is that when regular binders are more static, they are more likely to be stuck at binding sites due to a larger attraction. The Pearson correlation between the matrices shown in Figure 4 (b) and (d) is approximately 0.9. It is noticeable that for TAD 2 and TAD 6, the accumulating contacts at the corner are not obvious in our simulations, the main reason for that is due to that we did not include the inter-TAD interaction near the boundaries. In order to further compare the variation after deleting Wapl, we also exhibited the difference matrices of the experiment and simulation in Figure 4 (e) and (f) by subtracting maps of wide type cell from maps of Wapl-deficient cell. As the TADs near the boundaries of matrices are more difficult to predict due to the influence outside the segment we chose, we here only selected the proportion of matrices containing TAD 3, 4 and 5 to compare the difference matrices. Both matrices revealed the decrease of contacts along the diagonals(short-range interaction) and increase of contacts at the corner of TADs(long-range interaction).

Z. Yang and M. R.K. Mofrad

However, some may argue that despite the accumulating contacts at the corner after deleting Wapl, intra-TAD contact frequency also increases while the experimental data shows an opposite result. There are actually several reasons for the difference. Firstly, Figure 4 (a) and (b) indicates that after Wapl is eliminated, the inter-TAD contact frequency increases which can cause the decrease of intra-TAD contacts owing to the normalization of total contact numbers. Nevertheless, we did not consider the inter-TAD interaction near the boundaries in our models. Secondly, we assumed that removing Wapl was not capable of influencing the binding events so the transition probability from weak binders to regular binders did not change. As a result, the concentration of regular binders in equilibrium is much too larger in Δ Wapl cell than in wild type cell, leading to the increase of contacts intra TADs. Though the increase of intra-TAD contact frequency can be explained by the increase of regular binders, another question is raised: Is it also the reason for the accumulating contacts at the corner of TADs? To demonstrate that the increase of concentration of regular binders are not responsible for the increase of contacts at the corner, we further generated the difference matrices with fixed concentration of binders. The results included in Supporting Material prove that dynamics variation is the key of the accumulating contacts at the corner.

Figure 4: Contact matrices comparison (a), (b) Experimental Hi-C contact matrices of wild type cell and $\Delta Wapl$ cell (data accessible at NCBI GEO database(16), accession GSE95015). The triangles in (a) are TADs we took into consideration. (c), (d) Simulated contact matrices for wild type cell and $\Delta Wapl$ cell. Contact matrices (including (a), (b), (c), (d)) are normalized to 120000 contacts. (e), (f) Experimental and simulated matrices showing difference of contacts between two categories of cells

Z. Yang and M. R.K. Moftad

CONCLUSION

Here we discussed about the dynamics of chromatin folding based on SBS model using MD simulations. Through varying periods and transition probabilities, dynamics of binders and binding sites were controlled and compared quantitatively in our simulations. We investigated the roles of dynamics during the process of chromatin folding and revealed that neither dramatically dynamic nor static folding is favorable for chromatin to assemble to a condensed state. In addition, we noticed that distinct dynamics is capable of causing a difference of concentration of beads involved with the folding process in equilibrium. As a result, the difference of concentration leads to chromatin loop length variation. Last but not the least, contact matrices were reproduced with relatively high accuracy based on a simply model. By decreasing the transition probabilities, difference between contact matrices of wild type cell and Wapl-deficient cell was also predicted by our simulations. All in all, our dynamic model for chromatin folding is able to reveal variety of phenomena related to dynamics of the process both qualitatively and quantitatively.

SUPPORTING MATERIAL

Supporting Material is available at ...

Z. Yang and M. R.K. Mofrad

AUTHOR CONTRIBUTIONS

Z.Y. and M.R.K.M. designed the research. Z.Y. carried out all simulations, analyzed the data. Z.Y. and M.R.K.M. wrote the manuscript.

ACKNOWLEDGMENTS

REFERENCES

1. Hu, M., K. Deng, Z. Qin, and J. S. Liu, 2013. Understanding spatial organizations of chromosomes via statistical analysis of Hi-C data. *Quantitative Biology* 1:156–174.
2. Fraser, P., and W. Bickmore, 2007. Nuclear organization of the genome and the potential for gene regulation. *Nature* 447:413.
3. Dekker, J., 2008. Gene regulation in the third dimension. *Science* 319:1793–1794.
4. Misteli, T., 2007. Beyond the sequence: cellular organization of genome function. *Cell* 128:787–800.
5. Dekker, J., K. Rippe, M. Dekker, and N. Kleckner, 2002. Capturing chromosome conformation. *science* 295:1306–1311.
6. Simonis, M., P. Klous, E. Splinter, Y. Moshkin, R. Willemsen, E. De Wit, B. Van Steensel, and W. De Laat, 2006. Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture–on-chip (4C). *Nature genetics* 38:1348.
7. Dostie, J., T. A. Richmond, R. A. Arnaout, R. R. Selzer, W. L. Lee, T. A. Honan, E. D. Rubio, A. Krumm, J. Lamb, C. Nusbaum, et al., 2006. Chromosome Conformation Capture Carbon Copy (5C): a massively parallel solution for mapping interactions between genomic elements. *Genome research* 16:1299–1309.
8. Lieberman-Aiden, E., N. L. Van Berkum, L. Williams, M. Imakaev, T. Ragozy, A. Telling, I. Amit, B. R. Lajoie, P. J. Sabo, M. O. Dorschner, et al., 2009. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *science* 326:289–293.
9. Sofueva, S., E. Yaffe, W.-C. Chan, D. Georgopoulou, M. V. Rudan, H. Mira-Bontenbal, S. M. Pollard, G. P. Schroth, A. Tanay, and S. Hadjur, 2013. Cohesin-mediated interactions organize chromosomal domain architecture. *The EMBO journal* 32:3119–3129.
10. Gassler, J., H. B. Brandão, M. Imakaev, I. M. Flyamer, S. Ladstätter, W. A. Bickmore, J.-M. Peters, L. A. Mirny, and K. Tachibana, 2017. A mechanism of cohesin-dependent loop extrusion organizes zygotic genome architecture. *The EMBO journal* 36:3600–3618.
11. Rao, S. S., S.-C. Huang, B. G. St Hilaire, J. M. Engreitz, E. M. Perez, K.-R. Kieffer-Kwon, A. L. Sanborn, S. E. Johnstone, G. D. Bascom, I. D. Bochkov, et al., 2017. Cohesin loss eliminates all loop domains. *Cell* 171:305–320.
12. Fudenberg, G., M. Imakaev, C. Lu, A. Goloborodko, N. A. Abdennur, and L. A. Mirny, 2016. Formation of Chromosomal Domains by Loop Extrusion. *Cell reports* 15 9:2038–49.
13. Kim, T. H., Z. K. Abdullaev, A. D. Smith, K. A. Ching, D. Loukinov, R. D. Green, M. Q. Zhang, V. V. Lobanenko, and B. Ren, 2007. Analysis of the Vertebrate Insulator Protein CTCF-Binding Sites in the Human Genome. *Cell* 128:1231–1245.
14. Zuin, J., J. R. Dixon, M. van der Reijden, Z. H. Ye, P. Kolovos, R. W. W. Brouwer, M. P. C. van de Corput, H. J. G. van de Werken, T. A. Knoch, W. F. J. van Ijcken, F. G. Grosveld, B. Ren, and K. S. Wendt, 2014. Cohesin and CTCF differentially affect chromatin architecture and gene expression in human cells. *Proceedings of the National Academy of Sciences of the United States of America* 111 3:996–1001.
15. Gandhi, R., P. J. Gillespie, and T. Hirano, 2006. Human Wapl Is a Cohesin-Binding Protein that Promotes Sister-Chromatid Resolution in Mitotic Prophase. *Current Biology* 16:2406–2417.
16. Haarhuis, J. H., R. H. van der Weide, V. A. Blomen, J. O. Yáñez-Cuna, M. Amendola, M. S. van Ruiten, P. H. L. Krijger, H. G. M. Teunissen, R. H. Medema, B. van Steensel, T. R. Brummelkamp, E. de Wit, and B. D. Rowland, 2017. The Cohesin Release Factor WAPL Restricts Chromatin Loop Extension. *In Cell*.
17. Tedeschi, A., G. Wutz, S. Huet, M. Jaritz, A. Wünsche, E. Schirghuber, I. F. Davidson, W. W. Tang, D. A. Cisneros, V. Bhaskara, T. Nishiyama, A. Vaziri, A. Wutz, J. Ellenberg, and J. G. Peters, 2013. Wapl is an essential regulator of chromatin structure and chromosome segregation. *Nature* 501:564–568.

Z. Yang and M. R.K. Mofrad

18. Lopez-Serra, L., G. Kelly, H. M. Patel, A. Stewart, and F. Uhlmann, 2014. The Scc2/Scc4 complex acts in sister chromatid cohesion and transcriptional regulation by maintaining nucleosome-free regions. *In Nature Genetics*.
19. Petela, N. J., T. G. Gligoris, J. B. Metson, B.-G. Lee, M. Voulgaris, B. Y.-K. Hu, S. Kikuchi, C. Chapard, W. Chen, E. Rajendra, M. Srinivasan, H. Yu, J. Löwe, and K. Nasmyth, 2018. Scc2 Is a Potent Activator of Cohesin's ATPase that Promotes Loading by Binding Scc1 without Pds5. *In Molecular cell*.
20. Wutz, G., C. Várnai, K. Nagasaka, D. A. Cisneros, R. R. Stocsits, W. H. Tang, S. Schoenfelder, G. Jessberger, M. Muhar, M. J. Hossain, N. Walther, B. Koch, M. Kueblbeck, J. Ellenberg, J. Zuber, P. Fraser, and J.-M. Peters, 2017. Topologically associating domains and chromatin loops depend on cohesin and are regulated by CTCF, WAPL, and PDS5 proteins. *The EMBO journal* 36 24:3573–3599.
21. Kueng, S., B. Hegemann, B. H. Peters, J. J. Lipp, A. Schleiffer, K. Mechtler, and J. G. Peters, 2006. Wapl Controls the Dynamic Association of Cohesin with Chromatin. *Cell* 127:955–967.
22. Kanke, M., E. Tahara, P. J. H. I. Veld, and T. Nishiyama, 2016. Cohesin acetylation and Wapl-Pds5 oppositely regulate translocation of cohesin along DNA. *The EMBO journal* 35 24:2686–2698.
23. Hansen, A. S., I. Pustova, C. Cattoglio, R. Tjian, and X. Darzacq, 2017. CTCF and cohesin regulate chromatin loop stability with distinct dynamics. *In eLife*.
24. Zickler, D., and N. Kleckner, 1999. Meiotic chromosomes: integrating structure and function. *Annual review of genetics* 33:603–754.
25. Nicodemi, M., B. Panning, and A. Prisco, 2008. A thermodynamic switch for chromosome colocalization. *Genetics* 179:717–721.
26. Nicodemi, M., and A. Prisco, 2009. Thermodynamic pathways to genome spatial organization in the cell nucleus. *Biophysical journal* 96:2168–2177.
27. Barbieri, M., M. Chotalia, J. Fraser, L.-M. Lavitas, J. Dostie, A. Pombo, and M. Nicodemi, 2012. Complexity of chromatin folding is captured by the strings and binders switch model. *Proceedings of the National Academy of Sciences* 109:16173–16178.
28. Chiariello, A. M., C. Annunziatella, S. Bianco, A. Esposito, and M. Nicodemi, 2016. Polymer physics of chromosome large-scale 3D organisation. *In Scientific reports*.
29. Bianco, S., D. G. Lupiáñez, A. M. Chiariello, C. Annunziatella, K. Kraft, R. Schöpflin, L. Wittler, G. Andrey, M. Vingron, A. Pombo, S. Mundlos, and M. Nicodemi, 2018. Polymer physics predicts the effects of structural variants on chromatin architecture. *Nature Genetics* 50:662–667.
30. Nozaki, T., D. F. Hudson, S. Tamura, and K. Maeshima, 2018. 4 - Dynamic Chromatin Folding in the Cell. *In* C. Lavelle, and J.-M. Victor, editors, *Nuclear Architecture and Dynamics*, Academic Press, Boston, volume 2 of *Translational Epigenetics*, 101 – 122. <http://www.sciencedirect.com/science/article/pii/B9780128034804000041>.
31. Ramani, V., X. Deng, R. Qiu, K. L. Gunderson, F. J. Steemers, C. M. Disteche, W. S. Noble, Z. Duan, and J. Shendure, 2017. Massively multiplex single-cell Hi-C. *Nature methods* 14:263.
32. Zhou, J., J. Ma, Y. Chen, C. Cheng, B. Bao, J. Peng, T. Sejnowski, J. Dixon, and J. Ecker, 2018. HiCluster: A Robust Single-Cell Hi-C Clustering Method Based on Convolution and Random Walk. *bioRxiv* 506717.
33. Sanborn, A. L., 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., 2015. Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. *Proceedings of the National Academy of Sciences* 112:E6456–E6465.
34. Ganji, M., I. A. Shaltiel, S. Bisht, E. Kim, A. Kalichava, C. H. Haering, and C. Dekker, 2018. Real-time imaging of DNA loop extrusion by condensin. *Science* 360:102–105.