



# Lattice simulation-based diffusion modelling of 3D chromatin structure

Qingzhu Yang<sup>a</sup>, Zhihua Zhang<sup>a,b,\*</sup>



<sup>a</sup> CAS Key Laboratory of Genome Sciences and Information, Beijing Institute of Genomics, Chinese Academy of Sciences, and China National Center for Bioinformatics, Beijing 100101, China

<sup>b</sup> School of Life Science, University of Chinese Academy of Sciences, Beijing, China

## ARTICLE INFO

### Article history:

Received 15 March 2022

Received in revised form 26 June 2022

Accepted 26 June 2022

Available online 29 June 2022

### Keywords:

Chromatin interaction

Structure

Polymer model

Diffusion

Preserve

## ABSTRACT

Eukaryotic nuclear genome is extensively folded in the nuclei, and the chromatin structure experiences dramatic changes, i.e., condensation and decondensation, during the cell cycle. However, a model to persuasively explain the preserved chromatin interactions during cell cycle remains lacking. In this paper, we developed two simple, lattice-based models that mimic polymer fiber decondensation from initial fractal or anisotropic condensed status, using Markov Chain Monte Carlo (MCMC) methods. By simulating the dynamic decondensation process, we observed about 8.17% and 2.03% of the interactions preserved in the condensation to decondensation transition, in the fractal diffusion and anisotropic diffusion models, respectively. Intriguingly, although interaction hubs, as a physical locus where a certain number of monomers inter-connected, were observed in diffused polymer models in both simulations, they were not associated with the preserved interactions. Our simulation demonstrated that there might exist a small portion of chromatin interactions that preserved during the diffusion process of polymers, while the interacted hubs were more dynamically formed and additional regulatory factors were needed for their preservation.

© 2022 The Author(s). Published by Elsevier B.V. on behalf of Research Network of Computational and Structural Biotechnology. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

## 1. Introduction

The nuclear genome DNA in human cells have a physical length about 2 m, which need to be folded into a nuclear with about  $10\mu\text{m}$  in diameter [1]. The folding of the genome influences many fundamental cellular processes such as gene expression, DNA replication, repair and recombination [2,3]. Abnormality of the folding may lead to severe physiologic consequence, such as cancer [2]. Chromatin is decondensed during interphase and recondensed during mitosis.

Pre-existence of enhancer-promoter interactions have long been reported, using the computational prediction methods, e.g. [4–6]. A pre-existing loop means an interaction which is already established for a pair of enhancer and promoter before the enhancer is activated [7]. The interaction remains unchanged with the transient activation of the enhancer [7]. There are evidences of pre-existing loops for a large number of genes, such as P53, FOXO3 and TNF- $\alpha$  [7], and in multiple cell lines, such as IMR90, HUVEC

and MCF7 [7]. The pre-existing loop allows the enhancer to dictate the expression profile of its target gene even before the target being activated. The cell specific formation of pre-existing loops also allows the transcription factor to affect a set of genes' expressions in a cell specific way [7]. The pre-existence of these loops implied that they were preserved in the cell cycle since small amount of regulations were involved in forming the loops. However, the mechanism of pre-existing loops passing through cell cycle is nowhere near fully investigated.

To model the dynamics of chromosomal structures, polymer based models have been developed [1,3,8]. For instance, the random walk model considers a chromosome as a string of monomers with no spatial volume [1,2,9]. The monomers are driven by a Brownian motion [1,2,9]. The self-avoiding walk model [1,2], which is built upon the random walk model, assigns a spatial volume for each monomer, forcing them not to overlap with each other. Another model is the fractal globule model [3], where the polymer is confined in a cage which is rapidly crumpled by compressing. Different types of polymer models with loops have been proposed, including the dynamical loop model [2,10] and the loop extrusion model [8]. The dynamical loop model is built upon the random walk model, supposing that there is a random chance of two monomers attaching with each other to form loop with a random life time when they move close enough with each other. At

\* Corresponding author at: CAS Key Laboratory of Genome Sciences and Information, Beijing Institute of Genomics, Chinese Academy of Sciences, and China National Center for Bioinformatics, Beijing 100101, China.

E-mail addresses: [yangqzhu@foxmail.com](mailto:yangqzhu@foxmail.com) (Q. Yang), [zhangzhihua@big.ac.cn](mailto:zhangzhihua@big.ac.cn) (Z. Zhang).

the end of the life time, the loop denatures and the two monomers become free again. The loop extrusion model is more systematic than the dynamical loop model, assuming that there are several “motors” with two paired partners banding on the chromosome. Every single partner is assumed to move along the chromosome independently with each other following a random walk without overlapping while still physically attached. When partners of a motor are moving apart from each other, a loop is extruded out. The loop extrusion model coincides well with the chromatin loop structure mediated by the CTCF and other chromatin modulators. It, together with some derivatives, is known as a powerful explanation of *in vivo* chromosome structure formulation [8]. However, few models have investigated the effect of transition of the condensation–decondensation states of the chromatin on genome organization during cell cycle.

In the present work, we aimed to address the question that to what extent the chromatin interactions preserved during the decondensation process of chromosome, ie. how much self-organizing might be involved in the formation of those preserved structure? We built two polymer models to mimic the dynamics of three-dimensional chromatin structures during the decondensation. The simulation was set to start from two condensed initial structures. The initial structures were modeled using fractal and anisotropic space filling curve. Markov Chain Monte Carlo (MCMC) method was performed to generate the chromatin structure. In our simulations, we found about 8.17% and 2.03% interactions preserved during the fractal diffusion and anisotropic diffusion process, respectively (see Table 1). Topological Associated Domains were not captured. Hubs were detected. The preserved interactions were not significantly enriched in hubs. The observation indicated that polymer thermodynamic motion did play a role in keeping the preserved interactions and hubs. However, the roles of thermodynamic motion in the two elements might be not the same. Moreover, complicated chromatin structures, such as TADs were probably not driven by thermodynamic motions. This investigation assessed the role of molecular dynamics in keeping chromatin structural elements which passed through the decondensation process and helped to understand of the mechanism about how they were kept.

## 2. Methods

### 2.1. Polymer model representation of chromatin fiber

We employed a polymer chain with  $N$  monomers in a cubic lattice to model the chromatin fiber. One lattice site was occupied by a single monomer. The distance between the consecutive monomer was  $\sigma$  (lattice gap). The lattice model was a generic, ramose macromolecular-like structure constructed on the lattice. Each monomer represented a structural unit, e.g., a TAD, or a nucleosome, while the links represented the genomic connections between the structural units. Two monomers were considered interacting if they were neighbors in the lattice space. The polymer

diffused by randomly moving from one vertex to a neighboring vertex on the lattice.

### 2.2. Scheme of the initial conformation

We modeled the chromatin dynamics of decondensation process using random walk. The chromatin fiber was diffused from an initial compacted stage, which was modeled using space filling curves (Fig. 1B). Two initial compact models were employed, the Peano curve [11], which represented fractal structure of chromatin suggested from Hi-C data [3], and anisotropic space filling curve, called Self-Avoiding Walk (SAW) curve, which represented another kind of chromatin dense structure (Fig. 1C). See more details for the generation of SAW curve in Fig. 1C. Both Peano and SAW curves were generated using in house code.

### 2.3. Monte Carlo simulation of decondensation

In our simulation, we used a standard Metropolis procedure to simulate the dynamics of the polymer chain. The procedure first proceeds by randomly choosing one of the monomers in the chain. Second, the chosen monomer is made to move. The following moves in the procedure are included:

1. Random rotation of terminal monomers by a rotation angle of  $90^\circ \cdot n$  ( $n = 1, 2, 3, \dots$ ) in a random direction (move 1 in Fig. 1A).
2. Random rotation of chosen monomer  $i$  by a rotation angle of  $30^\circ \cdot n$  ( $n = 1, 2, 3, \dots$ ) in a random direction around axis linking the  $i + 1$  monomer and  $i - 1$  monomer (move 2 in Fig. 1A).
3. Random rotation of 2 consecutive monomers (the  $i - 1$  and  $i$  monomers or the  $i$  and  $i + 1$  monomers) by a rotation angle of  $30^\circ \cdot n$  ( $n = 1, 2, 3, \dots$ ) in a random direction around axis linking the  $i - 2$  and  $i + 1$  monomers or the  $i - 1$  and  $i + 2$  monomers (move 3 in Fig. 1A).

We reject the moves that cause crashes between the monomers. We used the above moves to avoid the moves that go against the topology where one part of the chain passes through another.

Since each chromosome occupied a certain territory [12,13], we employ a boundary condition of the moves to get a configuration of a high density. We modeled an initial chromosome as a polymer fiber with 4096 monomers, which can be filled into a lattice cube of size 16. The decondensed chromosome territory was modeled as a lattice cube with size 31, and the initial chromosome was put at the central of the chromosome territory.

The monomers interact via a shifted Lennard-Jones potential.

$$U_{ij} = \begin{cases} 4\varepsilon \left[ \left( \sigma/r_{ij} \right)^{12} - \left( \frac{\sigma}{r_{ij}} \right)^6 + 1/4 \right], & r_{ij} \leq \sigma 2^{1/6} \\ 0, & r_{ij} > \sigma 2^{1/6} \end{cases} \quad (1)$$

where  $r_{ij}$  was the distance between two monomers. The parameters were given by  $\sigma = 1$ ,  $\varepsilon = K_B T$  where  $K_B$  and  $T$  respectively represent the Boltzmann constant and the absolute temperature [14].

The moves were accepted or rejected according to Metropolis algorithm. The probability of attempting to move to state  $i$  from state  $j$  is  $p_{i \rightarrow j}$

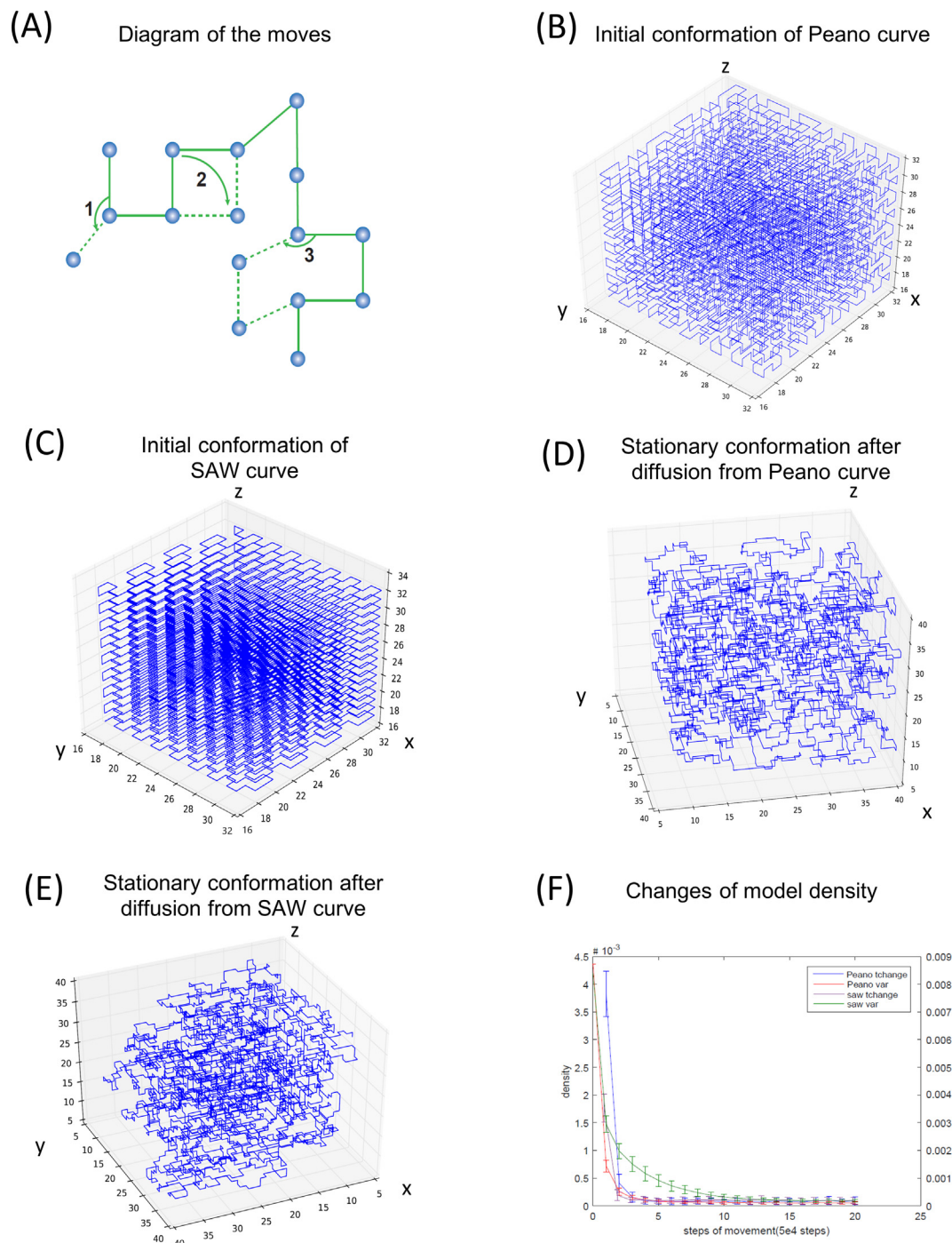
$$\begin{cases} p_{i \rightarrow j} = 1, E_j < E_i \\ p_{i \rightarrow j} = e^{-\left( \frac{E_j - E_i}{K_B T} \right)}, E_j < E_i \end{cases}, \quad (2)$$

where  $E_i$  and  $E_j$  are respectively the energy of state  $i$  and state  $j$ . The simulation was sustained until the stationary state is reached, when the density difference between the connecting every other  $5 \times 10^5$  steps of diffusion was sufficiently small to approximate a tenth of the density difference between the initial state and the first  $5 \times 10^5$  steps of diffusion. The density is calculated as the

**Table 1**

The preserved interactions averaged in 100 ensembles.

	Models diffusing from the Peano curve	Models diffusing from the SAW curve
D = Average of (ratio of Number of preserved pairs which interact in more than 10 models / All number of interaction pairs in stationary models) in 100 ensembles	8.17%	2.03%



**Fig. 1.** The process of the diffusion model from two initial conformations. (A) Diagram of the moves in Metropolis Monte Carlo simulation. The generation of the curve is shown. The initial conformation of two models, (B) Peano curve, and (C) SAW curve. This anisotropic curve SAW curve was generated as the following. We begin the curve from one monomer in a vertex labeled (0, 0, 0). Then the second monomer is added in +y direction at (0, 1, 0). We choose to extend the polymer layer by layer in +z direction by adding monomers. In each layer, the monomers are added one by one in (+y or -y) → (+x or -x) direction alternatively. If the polymer reaches the vertex of the layer and cannot extend without overlapping, it will extend in +z direction. The extension in another layer begins. For example, for the first four monomers, the polymer would be extended as positions (x, y, z) → (x, y + 1, z) → (x + 1, y + 1, z) → (x + 1, y, z). If the polymer grows to the desired length of 4096 monomers, the extension will stop. Examples of the final stationary conformation after diffusion from (D) Peano curve, and (E) SAW curve. (F) The dynamic changes of model density while diffusing. The line labeled “Peano tchange” represents the model density difference between the connecting every other  $5 \times 10^5$  steps of diffusion diffusing from Peano curve; The line labeled “saw tchange” represents the model density difference between the connecting every other  $5 \times 10^5$  steps of diffusion diffusing from SAW curve; The line labeled “Peano var” represents the variance of the model region density in the cube diffusing from Peano curve; The line labeled “saw var” represents the variance of the model region density in the cube diffusing from SAW curve.

ratio:  $\frac{\text{thenumberofmonomersintheregion}}{\text{thevolumeoftheregion}}$ . The volume of the region for calculating the density is set according to the method of the set of the best histogram in [15]. The simulation converged in no more than  $5 \times 10^6$  steps usually (Fig. 1F). We set the

maximum number of steps to  $7 \times 10^6$  to guarantee the convergence. The stationary states were further verified by checking the homogeneity of the density of the polymer in different regions in the cube at the end of the stage (Fig. 1F). The variance of the region

density in the cube is calculated for checking the homogeneity of the density of the polymer in different regions in the cube.

#### 2.4. Measurement of the contact probability

For a polymer chain with  $N$  monomers, the contact probability between two monomers at the stationary was estimated as the proportion  $P(s) = N_{actual}(s)/N_{possible}(s)$ , where  $N_{possible}(s)$  denoted the total number of monomer pairs with linear distance  $s$ , i.e.,  $N_{possible}(s) = N - s - 1$ , and  $N_{actual}(s)$  denoted the number of observed contacts for given distance  $s$ .

For a polymer chain with  $N$  monomers, the contact probability between two monomers at the stationary was estimated as the proportion  $P(s) = N_{actual}(s)/N_{possible}(s)$ , where  $N_{possible}(s)$  denoted the total number of monomer pairs with linear distance  $s$ , i.e.,  $N_{possible}(s) = N - s - 1$ , and  $N_{actual}(s)$  denoted the number of observed contacts for given distance  $s$ .

#### 2.5. Preserved interactions

To evaluate the preservation of the stationary conformations, we compared the simulated stationary conformation with the initial one. We simulate 100 diffused structures reaching stationary state from each initial conformation, respectively. To generate more structures, from two kinds of the diffused structures, we produce two kinds of 100 structure ensembles, using bootstrapping. One bootstrapping ensemble is generated as following: sampling with replacement from the 100 diffused structures for 100 times. Thus, each ensemble contains 100 structures.

A pair of two monomers was regarded as interacted, if the spatial distance is less than 2.5. The cut-off 2.5 was used because at this cut-off.

value the contact frequency of two monomers fit the power law relationship with the linear distance. A pair of interacting monomers, whose linear distance is more than 10, in the initial compacted stage is considered preserved, if in most of the simulations (i.e., more than  $\alpha\%$ ), the two monomers remained interacted after diffused into the stationary stage.

We defined the frequency of preserved interaction as proportion  $R = N(\text{preserved})/N(\text{all})$ , where  $N(\text{preserved})$  denotes the number of preserved interactions.  $N(\text{all})$  denotes the number of all interactions occurring in 100 structures. If two monomers, whose linear distance is more than 10, interacted at the stationary stage, the monomers interaction was included in the all interactions.

We compare the proportion and characteristics (such as linear distances) of preserved interaction observed in the polymer model with the random model. From each structure diffusing from the two initial conformations, we got the corresponding random structure. In the random structure, the monomers were randomly scattered, while the distribution of degrees and the number of edges of the interaction network were the same as the corresponding raw structure. Using Bootstrapping, 100 random structure ensemble is generated, respectively. One bootstrapping ensemble is generated as following: sampling with replacement from the 100 generated random structures for 100 times. Each ensemble contains 100 structures. We calculate the preserved interaction number in random structures. For the two models diffusing from Peano curve and SAW curve, we compare the number of preserved interactions in the 100 structure ensembles with the 100 random structure ensembles, using  $t$  test for different threshold of preserved interaction. For the two models, we also compare the frequency of preserved interaction in the 100 structure ensembles with the 100 random structure ensembles, using  $t$  test.

To characterize the observed preserved interactions in two models, we calculated and compared the distribution of the linear

distance between the interacted monomers of the preserved interactions and that of all interactions. For each model, we calculate the distribution of the linear distance of the preserved interactions and all interactions, respectively. To compare them, we perform a linear regression analysis to model the relationship between the probability for the preserved interaction and the overall interaction with the same linear distance.

#### 2.6. Hubs

The interaction hubs were defined as a physical locus where a certain number of monomers inter-connected at the stationary stage of the model. The number is 12 for the model diffusing from the Peano curve and 13 for the model diffusing from the SAW curve. The threshold 12 and 13 are determined by the degree distribution of interaction as shown in Fig. 4B and Fig. 4C. The enrichment of preserved interactions in the hub monomers was calculated by hypergeometric test.

### 3. Results

To investigate the dynamics of chromatin interactions during the decondensation process of chromosome, we did simulation analysis on polymer model of chromatin fibers on cubic lattice space. Two initial compacted chromosome models, the Peano curve and SAW curve, representing fractal and anisotropic models, respectively, were employed in the study (Fig. 1B C). Since the compacted conformations were of high energy, they were unstable and diffused following the thermodynamic rules [14]. We constrained the movement of the polymer in a larger but limited cubic lattice space to simulate the mutually exclusive chromosome territories in the nuclei. The stationary states were considered to be reached in the diffusion process when the changes of monomer distribution in the 3D space were sufficiently small between the diffusion steps. In our simulations, it would take about  $7 \times 10^6$  steps to reach the stationary state (Fig. 1F). For each initial model, we make Metropolis simulation 100 times for the analysis below.

#### 3.1. The stationary structures of simulated polymers caught the basic properties of decondensed chromosome.

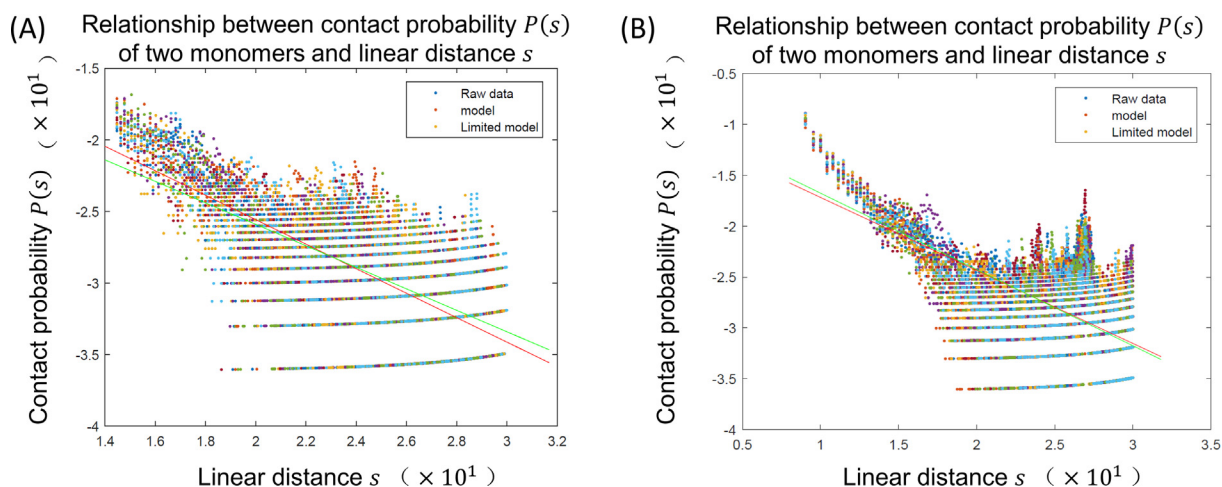
The distribution of contact frequencies (DCF) in resulted stationary structure of polymers with both the Peano and SAW model, were roughly the same as that of real decondensed chromosome. The DCF of the

polymers has a known form,  $P(s) \propto s^{-r}$  [12], in which the parameter  $r$  characterize the global structural property of the polymer chain. We compared the DCF between the stationary polymers models and the real chromosomes as indicated from Hi-C data [8,16]. The DCF in simulated polymers was calculated using a bin size 80. Comparing to the Hi-C data.

in mammalian ( $r \approx 0.75$ ) [8], the  $r$  is not grievously different to which of the stationary polymers (mean = 0.87, SD = 0.088, and mean = 0.71, SD = 0.099 for Peano and SAW models, respectively) (Fig. 2A and B). This similarity is remarkably high, considering the difference between the abstract lattice models and real chromatin structure in scale and form. Therefore, our abstract lattice model was capable to roughly mimic the chromatin dynamics.

#### 3.2. The random walk model reproduced the preserved interaction of the chromosome

We asked to what extent the monomer–monomer interactions at the initial compacted conformations remained to the terminal stationary states.. Specifically, in this paper, we took  $\alpha\%$  as thresh-



**Fig. 2.** The structure of diffused models. The relationship between contact probability  $P(s)$  of two monomers and linear distance  $s$  for the models diffused from (A) Peano curve, and (B) SAW curve.

old, remained interactions above the threshold were regarded as preserved (see methods for the details of  $\alpha$ , mentioned below means the same).

Preserved interactions were calculated for the simulated results. For the 100 simulated structures from the same initial compacted conformation, 100 bootstrap structure ensembles are sampled (see Methods for details).

For the structures diffusing from the Peano curve and SAW curve, we observed preserved interactions for different thresholds. For threshold  $\alpha\% = 10\%$ , we observe average 849.44 and 188.32 preserved interactions for 100 bootstrap ensembles diffusing from Peano curve and SAW curve, respectively (Fig. 3). Based on the results of the simulation, we calculated the preserved interactions the proportion of preserved interactions in the observed data is 8.17% or 2.03% (see Methods for details).

To investigate the preserved interaction predictions in our models, we compare the observed preserved interactions with the random model (see methods for the definition of the random model). First, the comparison of the number of the observed preserved interactions in our models and the preserved interactions in random model are performed. For threshold  $\alpha\% = 10\%$ , we observe averages of 0.88 and 2.58 preserved interactions for 100 structure ensembles from random models corresponding with models diffusing from Peano and SAW curves, respectively (Fig. 3C and Fig. 3D). The number of preserved interactions from Peano and SAW curves were much higher than the random models (both with  $p$  value  $< 1e-12$ ,  $t$  test). This result is conserved for different thresholds  $\alpha\%$  as shown in Fig. S1 and Fig. S2. Second, the comparison of the frequency of the observed preserved interactions (see methods for details) in our models and the random preserved interaction are performed. For the results of the 100 ensembles reaching stationarity from Peano curve, the average frequency of observed preserved interaction is 0.081 for threshold  $\alpha\% = 10\%$ . It is much higher than the random preserved interaction,  $7.6e-5$ . For the results of the 100 ensembles reaching stationarity from SAW curve, the average frequency of observed preserved interactions was 0.020 for threshold  $\alpha\% = 10\%$ , much higher than the corresponding value in the random model,  $2.5e-4$ . For the two initial conformations we compare the frequency of preserved interaction with the random models in the 100 bootstrap ensembles using  $t$  test as shown in Fig. 3E and Fig. 3F. For different thresholds  $\alpha\%$ , the preserved interaction frequencies of the models are higher than the random models, as shown in Fig. S3. These significant results imply that the preserved interactions are not generated randomly.

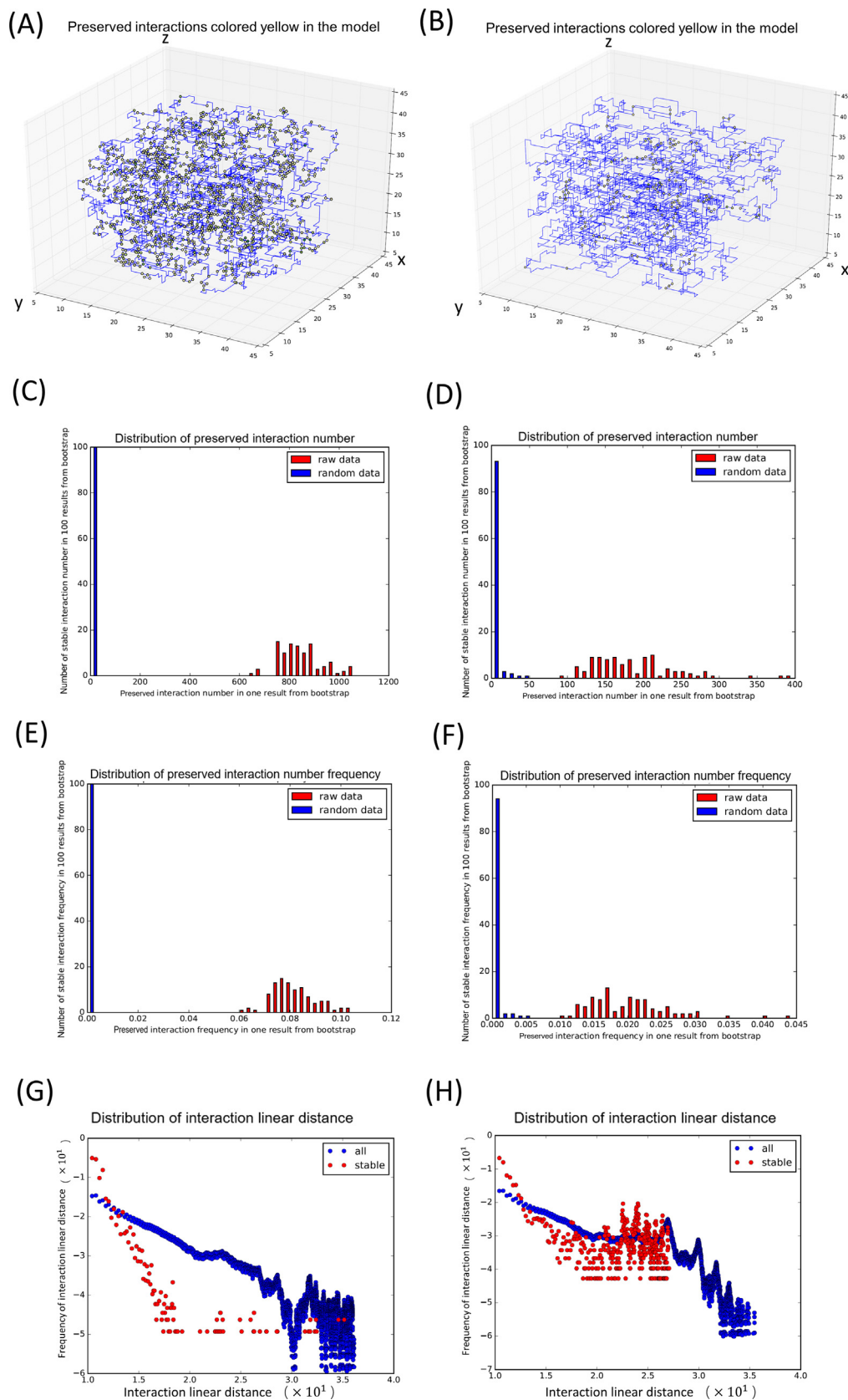
The preserved interactions tended to occur at loci with a smaller linear distance compared with other interactions. The distribution of the linear distance of the preserved interactions and all interactions at the stationary state were calculated for 100 ensembles, respectively (see Methods for the details of all interactions) (Fig. 3G and Fig. 3H). After performing the linear regression analysis for modelling the relationship between the probability for the same linear distance of the preserved interactions and all interactions, the coefficients of regression of the results from Peano and SAW curves were 5.4 and 4.6, respectively. The difference between linear distance distributions of preserved and all interactions, further indicated that the preserved interaction is different from the interaction of the rest part of the structure.

### 3.3. The preserved interactions reproduced part of the classic chromosome conformation hierarchies

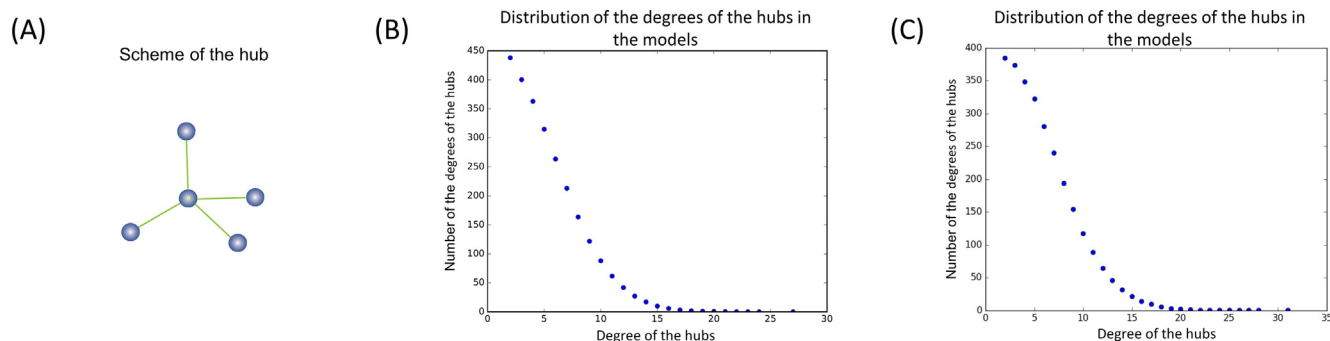
We inspected the relationship between the preserved interactions and the classic chromosome conformation hierarchies such as TADs and hubs.

First, there was no TAD or domain-like structure observed in simulated polymers. We applied a classical Hidden Markov model (HMM) [16] to the simulated polymers, however, none TAD or domain-like structure was reported in all the 100 simulations. This result implied that TAD structure may not automatically emerge from a homogenous initial stage, e.g. Peano or SAW curve, and additional factors and regulatory mechanism were essential to the formation of TADs.

Next, we asked whether the preserved interactions were enriched in interaction hubs. The interaction hubs were defined as a physical locus where a certain number of monomers interconnected. The number is 12 for the model diffusing from the Peano curve and 13 for the model diffusing from the SAW curve (see Methods for the details of the number). Though hubs existed our models, preserved interactions were not enriched there. Hub is important in the chromatin structure field because it describes the mechanism of various complex regulation patterns involving more than one regulation elements. The scheme is shown in Fig. 4A. We found hubs in the models diffusing from Peano curve and SAW curve. To analyze the enrichment of preserved interactions in the hub monomers, hypergeometric test is performed. We found for the models diffusing from the two conformation, preserved interactions are not enriched in the hub monomers by hypergeometric test. These findings implied that the formation of hubs may be related with polymer thermodynamical motion. But



**Fig. 3.** The preserved interactions. The preserved interactions colored yellow in the model diffusing from (A) the Peano curve, and (B) the SAW curve. The distribution of preserved interaction number in the models diffusing from (C) the Peano curve and (D) SAW curve. The preservation was defined that an interaction be found in more than 10 replicated simulations. Preserved interaction number frequency distribution in the models diffusing from (E) the Peano curve (F) the SAW curve. The distribution of linear distance of the preserved interactions and all interactions in the models diffusing from (G) Peano curve, and (H) the SAW curve. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)



**Fig. 4.** Hubs and the distribution of the degree. (A) The scheme of the hub. The distribution of the degrees of the hubs in the models reaching the stationarity diffusing from (B) the Peano curve and (C) the SAW curve.

the forces leading to the hubs and preserved interactions are probably different since they are not correlated with each other.

In conclusion, we demonstrated that preserved interaction can be observed even in the simplified compacted to decondensation process in a lattice space. While higher level structure of the complex polymer chain, such as TAD that observed in real Hi-C data, was absent in the simulation, suggested additional regulation was essential to formation of such structure.

#### 4. Discussion

In the problem of the preserved chromatin interactions, we tried to solve it with a brief model. Here we investigated the models to study this phenomenon from the viewpoint of the polymer diffusion from the initial condensed state. A series of methods both experimentally and computationally have been developed to investigate the chromatin organization. Chromatin conformation capture (3C) approach and related approaches, such as 4C, 5C and Hi-C approaches have also demonstrated various aspects of chromatin folding [3,17–19]. These approaches are based on nuclear ligation [20]. Analyzing the products of these approaches can estimate the contact probability  $P(s)$ , which represents the interaction frequency of pairs of loci with genomic distance  $s$  contact with each other. The  $P(s)$ , often displays scaling properties with the form of  $P(s) \propto s^{-\tau}$  [3,21].

We developed two models to mimic the process of chromosome decondensation from a compact configuration. The models were based on the fractal diffusion or anisotropic diffusion model. The models were simulated using MCMC methods to the stationary state. The similarity between the parameter  $\tau$  of the models and real chromatin were remarkably high. The resulted structural ensembles showed the existence of preserved interactions, indicating that these preserved interactions may be formed by thermodynamics forces. Furthermore, TAD cannot be detected from the thermodynamical model, implying a different mechanism of them. Hubs could be formed by the diffusion since they were found in the resulted structures, yet were not correlated with the preserved interactions, suggesting the two were from a different mechanism. These implied parts of chromosome structure, such as preserved interactions and hubs may be correlated with polymer thermodynamical motion.

We could find from the above derivation that the mechanism of forming of TAD and hubs are far beyond the thermodynamics motion. Further, the thermodynamics derived preserved interactions has no relation with TAD and hubs although hubs may also be formed by thermodynamics motion. They may need some other factors in cell, for example, transcription factor and so on.

On the other hand, the models proposed in this paper still could be improved. Causing the chromatin is composed of DNA and pro-

tein factors, such as CTCF and cohesion. Future work of the models could encompass the integration the different factors binding to the DNA.

#### 5. Availability

<https://cn.mathworks.com/matlabcentral/fileexchange/28529-3d-peano-space-filling-curve>  
<https://github.com/Wind20226/Modelling-of-3D-chromatin-structure>

#### Funding

This work was supported by the National Key R&D Program of China (2018YFC2000400 to ZZ) and the National Natural Science Foundation of China (no: 31501081 to Q.Y.).

#### Author contributions

Q.Y. and Z.Z. conceived this work. Q.Y. and Z.Z. contributed to methodology, software and formal analyses. Q.Y. applied data visualization, and data curation. Q.Y. and Z.Z. wrote the paper.

#### Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

#### Acknowledgements

We thank Dr. Bingxiang Xu for providing helpful advice and support for this project.

#### Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.csbj.2022.06.057>.

#### References

- [1] Bohn M, Heermann DW, van Driel R. Random loop model for long polymers. *Phys Rev E Stat Nonlin Soft Matter Phys* 2007;76(5 Pt 1):051805.
- [2] Rosa A, Zimmer C. Computational models of large-scale genome architecture. *Int Rev Cell Mol Biol* 2014;307:275–349.
- [3] Lieberman-Aiden E et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* 2009;326(5950):289–93.
- [4] Rao SS et al. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* 2014;159(7):1665–80.

- [5] Ay F et al. Statistical confidence estimation for Hi-C data reveals regulatory chromatin contacts. *Genome Res* 2014;24(6):999–1011.
- [6] Zhou Y et al. Modeling and analysis of Hi-C data by HiSIF identifies characteristic promoter-distal loops. *Genome Med* 2020;12(1):69.
- [7] Jin F et al. A high-resolution map of the three-dimensional chromatin interactome in human cells. *Nature* 2013;503(7475):290–4.
- [8] Sanborn AL et al. Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. *Proc Natl Acad Sci U S A* 2015;112(47):E6456–65.
- [9] Sachs RK et al. A random-walk/giant-loop model for interphase chromosomes. *Proc Natl Acad Sci U S A* 1995;92(7):2710–4.
- [10] Zhang Y, Heermann DW. Loops determine the mechanical properties of mitotic chromosomes. *PLoS ONE* 2011;6(12):e29225.
- [11] Peano G. Sur une courbe, qui remplit toute une aire plane. *Math Ann* 1890;36(1):157–60.
- [12] Rosa A, Everaers R. Structure and dynamics of interphase chromosomes. *PLoS Comput Biol* 2008;4(8):e1000153.
- [13] Über CR. Zellteilung. *Morph. Jb* 1885;10:214–330.
- [14] Rubinstein M, Colby RH. *Polymer physics*. Oxford University Press; 2003.
- [15] Scott DW. On optimal and data-based histograms. *Biosmetrika* 1979;66(2):605–10.
- [16] Rao SSP et al. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* 2014;159(7):1665–80.
- [17] Simonis M et al. Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture–on-chip (4C). *Nat Genet* 2006;38(11):1348–54.
- [18] Simonis M, Kooren J, de Laat W. An evaluation of 3C-based methods to capture DNA interactions. *Nat Methods* 2007;4(11):895–901.
- [19] Dostie J et al. Chromosome Conformation Capture Carbon Copy (5C): a massively parallel solution for mapping interactions between genomic elements. *Genome Res* 2006;16(10):1299–309.
- [20] Cullen KE, Klädde MP, Seyfred MA. Interaction between transcription regulatory regions of prolactin chromatin. *Science* 1993;261(5118):203–6.
- [21] Barbieri M et al. Complexity of chromatin folding is captured by the strings and binders switch model. *Proc Natl Acad Sci U S A* 2012;109(40):16173–8.