Polymer modeling of 3D epigenome folding: application to Drosophila

Daniel Jost

Polymer modeling of 3D epigenome folding: application to *Drosophila*

Daniel Jost[1,*]

[1] University of Lyon, ENS de Lyon, Univ Claude Bernard, CNRS, Laboratoire de Biologie et Modélisation de la Cellule, 46 Allée d'Italie, 69007 Lyon, France

[*] corresponding author: daniel.jost@ens-lyon.fr ; phone: +33 4 72 72 80 00

Running head

Polymer modeling for epigenome folding

Abstract

Mechanistic modeling in biology allows to investigate, based on first principles, if putative hypotheses are compatible with observations and to drive further experimental works. Along this line, polymer modeling has been instrumental in 3D genomics to better understand the impact of key mechanisms on the spatial genome organization. Here, I describe how polymer-based models can be practically used to study the role of epigenome in chromosome folding. I illustrate this methodology in the context of *Drosophila* epigenome folding.

Keywords

Chromosome organization, polymer physics, epigenomics, simulations, coarse-graining,

Drosophila

1. Introduction

Compartmentalization is a ubiquitous feature of cellular function. In the nucleus, early observations revealed a non-random spatial organization of the genome within chromosome territories and a large-scale segregation between transcriptionally active -euchromatic- and silenced -heterochromatic- parts of the genome. Recent advances in genome-wide mapping and imaging techniques have strikingly improved the resolution at which genome folding can be analyzed. It has unveiled a multi-scale spatial organization with increasing evidence that 3D genome folding may result from and participate in genome function [1]. Analyses of genome-wide chromosome contact data by Hi-C technology [2] have demonstrated that the genome of higher eukaryotes is partitioned into hundreds of kbp-long domains, called topologically associating domains (TADs) [3,4] where local contacts are enriched compared to interactions with neighboring domains. These TADs are usually associated with transcriptional or epigenetic regulation: promoter-enhancer interactions mainly occur within the same TAD [5] and homogeneous chromatin domains (in particular constitutively repressed chromatin) tend to fold into single TADs [6,7]. At the Mbp-scale, contact maps display a cell-type specific checker-board like pattern [6,8,9] with a complex pattern of long-range contacts between TADs. In particular TADs are organized into (spatial) (A/B, active vs repressive) compartments associated to epigenomic regulation.

Investigations of the model organism *Drosophila melanogaster* have been instrumental in understanding the building blocks of genome folding. Briefly, for this organism, the 3 largest chromosomes (chr2, chr3 and chrX) are organized into separate territories in the well-known polarized Rabl-like organization [10] similar to yeast (all centromeres at one pole of the nucleus and telomeres pointing towards the other pole). In *Drosophila* and more generally in dipteran, homologous chromosomes are tightly paired at the molecular level in somatic cells and occupy the same territory [11,12]. Hi-C experiments in late embryos revealed the existence of

epigenetically-associated TADs that interact at larger scales to form segregated compartments [6]. Recently, Hi-C along fly embryogenesis have shown the dynamical, progressive formation of TADs and of their higher-order organization driven by transcriptional re-activation and epigenome restoration [13,14].

While our current understanding of the 3D chromosome organization is becoming more and more quantitative, the underlying physical or biochemical mechanisms driving such peculiar folding remain unclear and several processes have been shown to play major roles in organizing chromosomes [15]. In the past decade, in addition to many seminal experimental works, polymer-physics modeling has been instrumental in studying these processes [16]. In particular, we developed several polymer-based approaches to better understand the relation between epigenomic regulation and 3D chromosome organization [17] mostly in the context of fly 3D genomics.

In this chapter, I describe in detail the theoretical framework and its numerical implementation that we developed along the years to study epigenome folding and their application to *Drosophila* nuclear organization. I focus on the methodological part of our work and refer the reader to the original publications for a detailed description of the physical and biological implications.

2. Methods

2.1 Coarse-grained polymer model for chromosome

Simulating the precise dynamics of long polymer chains (chromosome size ranging typically from few Mbp to hundreds Mbp in higher eukaryotes) during prolonged time period (orders of hours for typical cell cycles) remains a critical numerical challenge. A standard strategy to reduce computation time is to develop coarse-grained models [18] where fine-scale individual entities like

atoms, base-pairs or nucleosomes are grouped together into larger monomers, hence decreasing the complexity of the system and the degrees of freedom. This reduction of complexity should not be however at the expense of a complete modification of the physical properties of the system.
Recently, we developed a coarse-graining (CG) strategy for polymers that conserves the topological regime and thus preserves thes large-scale structural and dynamical properties [19]. Indeed, chromosomes being confined, very long polymers, their characteristics are very different from simple isolated short chains [20]. In particular, their dynamics strongly depend on the ratio between their contour length L and the so-called entanglement length L_e [21] that measures the degree of topological constraints acting on the polymers. $(L/L_e)<<1$ (respectively $>>1$) implies weak (resp. strong) effects on the chain behavior. In the following, I describe this CG procedure applied to chromosome modeling.

Generically, let's first consider a reference fine-scale (FS) model describing the chromatin fiber (Fig. 1A): a chromosome is modeled as a semi-flexible self-avoiding chain composed of N monomers, each monomer being of size σ (in nm) and containing ν bp. The local rigidity of the fiber is characterized by the Kuhn length l_k (in nm) that defined the length below which the fiber can be considered as stiff, and the molecular crowding is quantified by the volumic base-pair density ρ (in bp/nm^3). For example, in the standard 30-nm fiber model $\sigma = 30$nm, $\nu = 3$kbp and l_k $=300$ nm [22]. Recent investigations based on *in vitro* dynamics of yeast-like chromatin lead more to $\sigma \sim 20$nm, $\nu = 1$kbp and l_k $=100$ nm [23]. ρ is equal to the ratio between the genome size and the average volume of nuclei. It is therefore more species or cell-type dependent and typically varies between 0.005 bp/nm^3 for haploid yeast to 0.015 bp/nm^3 in typical mammalian nuclei.
We aim now to coarse-grain this model into a polymer chain for which we note N', σ', l_k' and ρ' the corresponding parameters. The total genomic content being invariant ($N\nu = N'\nu'$), each CG monomer corresponds to $N/N' = \nu'/\nu$ FS monomers. To preserve the correct dynamics when

coarse-graining, we impose the bp density to be conserved ($\rho = \rho'$) and *(L/Le)=(L'/Le')* with $L = N\sigma$ (idem for *L'*) and $L_e = l_k \left(\frac{19}{(\rho/\nu)(\sigma/l_k)l_k^3}\right)^2$ [19,24] (idem for *Le'*). These two conditions lead to the simple relation

$$\left(\frac{l_k\sigma}{\nu}\right) = \left(\frac{l'_k\sigma'}{\nu'}\right) \qquad (1)$$

that constrains the CG parameters. In [19], we showed that any CG models verifying this relation predicts the correct structural and dynamical FS properties as long as the corresponding volumic fraction (the fraction of the available volume occupied by the chain) $\phi' \equiv \rho\omega'/\nu'$ is not too high (typically less than half of the maximum packing ratio of the system) with $\omega' \sim \sigma'^3$ the volume occupied by 1 CG monomer. We also remarked that higher ν' and ϕ' values lead to better numerical efficiencies.

While this CG strategy is general, when applied to a specific system, the question is how to choose the optimal coarse-graining parameters to maximize the numerical efficiency at a given desired resolution (see also the Note 3.1)? In the CG model, the spatial resolution is given by l_k' and the base-pair resolution by $(\nu' l_k'/\sigma')$. Therefore, the choice of parameters should be motivated by the maximization of ν' and ϕ' constrained to the typical experimental resolution of the system under study. For example, for ρ=0.015 bp/nm^3 (mammalian nuclei) and a desired resolution of 10kbp (typical of HiC maps), if monomers are modeled as spheres ($\omega' = \pi\sigma'^3/6$), then choosing ν'=4.1kbp, $\sigma' = 58$nm and $l_k' = 142$nm allows to maximize ϕ' up to ~37% (maximum packing ratio for spheres=74%) while satisfying the relation (1) and keeping a targeted resolution of 10kbp.

2.2 Block copolymer model for epigenome folding

Once parameters of the desired CG “null” model have been selected, one can then complement the model with various physical or biochemical interactions in order to investigate the impact of putative mechanisms on 3D chromosome organization. In this section, I describe how to implement and model the role of epigenome in chromatin folding and illustrate this for *Drosophila*.

As demonstrated by us and others [8,25–27], epigenomic information is tightly associated with 3D chromatin organization in particular at the compartment level where genomic regions sharing the same epigenomic content tend to cluster in space in the nucleus (Fig.1B). Moreover, there exists an increasing number of biochemical evidence showing that key chromatin-binding proteins or complexes (like HP1, PRC1 but also RNA polymerase), associated to specific epigenomic patterns, have the intrinsic capacity to oligomerize or phase-separate [28–33]. This suggests that epigenome, via such interactions, may actively contribute to shape the 3D nuclear organization.
At the unidimensional level, many (dozens) epigenomic features like post-translational modifications of histone tails, DNA methylation, chromatin-binding proteins have been measured along the genome. Statistical analyses of these different profiles have demonstrated that some of these markers are strongly correlated and that the local epigenomic content can be actually summarized by a handful (4 to 10) of different chromatin states [27,34–38]. In *Drosophila*, Filion et al [34] inferred few different types (Fig.1C): two euchromatic states associated with actively transcribed regions that, for simplicity, we usually merge into one single active state, one state associated with constitutive heterochromatin and enriched in H3K9me2/3 histone marks and HP1 proteins, one state associated with facultative heterochromatin and enriched in H3K27me3 histone marks and Polycomb-group (PcG) protein complexes (PRC1/2), and another heterochromatic state, the so-called black or null or quiescent chromatin, a prevalent form of

repressive chromatin not particularly enriched in known epigenomic markers. These different states segment the genome into contiguous chromatin domains of various sizes (Fig.1B).

Based on this genome partitioning, we developed a block copolymer model of epigenome folding (Fig.2A) [39]. Practically, we considered a CG null model with 10kbp-beads and assigned to each bead *i* its corresponding epigenomic state *e(i)* among the 4 considered types (active, HP1-like, PcG-like and black), consecutive beads belonging to the same chromatin domain forming one uniform block. To account for the interaction capacity of chromatin-binding proteins associated with these states, we assumed that monomers sharing the same state *e* can interact specifically with an energy $U_{e,e}$ if they are in close spatial proximity. For simplicity, we did not consider cross-state interactions ($U_{e,e'}$ =0 if $e \neq e'$) and assumed that the strengths of interactions are identical for every chromatin type ($U_{e,e} = U_{e',e'} \equiv U \ \forall e, e'$, see also the Note 3.3). Therefore, the dynamics of the chain is driven by the bending rigidity and excluded volume contributions of the null model plus epigenomic-mediated short-range interactions.

2.3 Lattice model and numerical simulations

To investigate the dynamical behavior of epigenome folding, we implemented the block copolymer on a lattice. More precisely, the CG polymeric chain is modeled as a self-avoiding walk on a face centered cubic (fcc) lattice of size *SxSxS* (Fig.2B) which has a high coordination number (=12). To account for the effect of contour length fluctuations on chain dynamics [40,41], particularly relevant in dense systems, we allow at maximum two monomers to be on the same lattice node if and only if they are consecutive along the chain [42]. Otherwise, due to excluded volume, two monomers cannot overlap on the same site.

The polymer stiffness of the CG model is accounted via a standard Kremer-Grest potential $H_{bend} = (\kappa/2)\sum_{i=1}^{N_l-1}(1 - cos\theta_i)$, where θ_i is the angle between bond vectors *i* and *(i+1)* (see

Fig.2B) and κ is the bending energy that is connected to $l_k{}'$ and σ' in our fcc lattice framework via the relation [19]

$$\frac{l_k{}'}{\sigma'} = \frac{1+x}{1-x} \qquad \text{with} \qquad x = \frac{12}{13}\left(\frac{1+2exp[-\kappa/2]-2exp[-3\kappa/2]-exp[-2\kappa]]}{1+4exp[-\kappa/2]+2exp[-\kappa]+4exp[-3\kappa/2]+exp[-2\kappa]}\right) \qquad (2)$$

For example, in *Drosophila* (ρ~0.009 bp/nm^3), we used to work with a CG null model defined as ν'=10kbp, $\sigma' = 115$nm and $l_k{}' = 274$nm at $\phi' \sim 1$ (the maximum packing ratio of this lattice model is 2) for a bp-resolution of 23kbp [19,42]. In this case, relation (2) leads to $\kappa = 1.5k_BT$. Molecular crowding and confinement are approximated by using periodic boundary conditions, the size of the box *S* being chosen such that the lattice volumic fraction $N'/(4S^3)$ is equal to the targeted ϕ' value.

Epigenomic-mediated interactions were modeled by a specific contact Hamiltonian $H_{epi} = U\sum_{i,j}\Delta_{e(i),e(j)}\delta_{i,j}$ with $U \equiv U_{e,e}$ the strength of interaction, $\Delta_{e(i),e(j)} = 1$ if monomers *i* and *j* have the same chromatin state (ie, *e(i)=e(j)*) ($\Delta_{e(i),e(j)} = 0$, otherwise), and $\delta_{i,j} = 1$ if *i* and *j* are in 3D contact (ie if they occupy nearest-neighbor lattice sites) ($\delta_{i,j} = 0$, otherwise).

Due to topological constraints, the large-scale organization of long polymers like chromosomes keeps a partial memory of the initial arrangement over long timescales [20,43]. Therefore, initial configurations should be carefully designed such that the polymer topology (presence or absence of knots) and the large-scales are consistent with biological observations. For *Drosophila*, we consider unknotted chromosomes in a Rabl-like configuration that we generate using the hedgehog algorithm [19,44] (Fig.2C): starting from a central, rod-like scaffold, configurations are iteratively grown by randomly inserting monomers at nearest-neighbor sites common to two already placed consecutive monomers.

The dynamics of the chain (Fig.2D) is then simulated using a kinetic Monte-Carlo (KMC) algorithm. Namely, one Monte-Carlo step (MCS) consists of *N'* trial moves where a monomer is randomly picked and a random trial new position is chosen among the nearest-neighbor (NN)

lattice sites (Fig.2B). The move is accepted: (i) based on a Metropolis criterion, ie with a probability equal to $min(1, exp(-[H_{tot}^{new} - H_{tot}^{old}]/(k_B T)))$ with $H_{tot} = H_{bend} + H_{epi}$ the total energy of the system and *old* and *new* refer respectively to the current and the trial conformations; (ii) and if the move does not break the chain connectivity (consecutive monomers along the chain should occupy the same or NN sites) and if the excluded volume criterion is still verified (two non-consecutive monomers cannot occupy the same site).

Time mapping between the simulation and real times is made by computing from the simulations the time evolution of the root-mean-squared displacement ($RMSD(\tau) \equiv \sqrt{< (r(t+\tau) - r(t))^2 >}$ with $r(t)$ the position at time *t*) which represents the average spatial displacement of a monomer after a given time lag τ and by comparing it to experimental measurements, where typically $RMSD_{exp}(\tau)[\mu m] \approx 0.1\tau^{1/4}$ with τ in sec [23]. For example, with the CG parameters given above for *Drosophila*, 1 MCS~20 msec [19].

The advantage of simulating polymers on lattice using KMC compared to standard off-lattice Molecular Dynamics approaches (see also the Note 3.2) is that, at the expense of discretizing the space, the implementation is very simple leading to numerically very efficient algorithms while still capturing the main generic features of polymer dynamics [45]. For example, simulating one trajectory of a 20Mbp-long chromosome (with the CG parameters given above) during 30 min of real time requires less than 30 CPU sec on a 3.2GHz CPU.

2.4 Comparison with experiments

Based on this simulation framework, we can then investigate the structural and dynamical properties of the block copolymer model. In this section, I focus on the methodologies we used to compare predictions to Hi-C data.

For a given parameter set, to estimate the corresponding *in silico* Hi-C map with good statistics, we run many (~ 500-1,000) independent trajectories, each corresponding to ~20h of real time, the typical cell cycle length in fly. Along one trajectory, at regular time interval (typically every 10 min), we store individual configurations. For each snapshot, we estimate an *in silico* single-cell Hi-C map by identifying the pairs of monomers whose relative distance is less than a given threshold r_c. r_c represents the 'chemical resolution of Hi-C data', ie the typical maximal spatial distance captured by Hi-C experiments. Its actual experimental value is not known and is certainly locus- and Hi-C-protocol-dependent, but typical values may lie between 150 and 250 nm. These single-cell maps can then be combined to obtain Hi-C maps for synchronized (averaged over a given time window along the cell cycle) or unsynchronized (averaged over all the snapshots) populations of cells.

In our application to epigenome folding in *Drosophila*, we vary the value of U, the strength of epigenomic interactions which is the only free remaining parameter of the model. For every value, we compute the corresponding unsynchronized Hi-C maps (Fig.3A) [19]. For weak U values, we observe that the system behaves as the CG null model. As U becomes stronger, blocks (ie the epigenomic domains) self-compact leading to the formation of TADs and long-range contacts between TADs of the same state appear leading to the creation of nuclear - micro-phased - compartments.

In order to compare predictions and experiments, one should first define the range where the developed model is expected to describe quantitatively the real system. The lower bound of this range is limited by the actual resolution of the model given by $(\nu' l_k'/\sigma')$ (=23kbp in our case). The upper bound depends on how well the large-scale organization was implemented in the initial configurations since the system will maintain a partial memory of this along the simulations. Without any strong knowledge or in absence of a systematic study of the impact of the large-scale

initial structure, it is better to limit the upper bound to ~5 Mbp, this scale reaching a metastable state, largely independent of the initial state, typically after a few hours of real time [19].

Once this range has been defined, several comparative scores can be computed. For example, one can simply estimate as a function of U the global Spearman correlation between the entries of the simulated and experimental Hi-C matrices (only keeping the entries between regions whose genomic distances lie in the range of comparison) for different r_c values (Fig.3B). Other indicators (eventually depending on the epigenomic state) may be estimated like, for example, correlation scores on the so-called expected decay plot (that characterizes the average contact frequency between loci separated by a given genomic distance) or on the 'observed/expected' matrices (obtained by dividing each entry of the original Hi-C map by the expected (average) value at the corresponding genomic distance). In our case, many measures suggest that Hi-C experiments in late embryos are compatible with $U \approx -0.1 k_B T$ and $r_c \sim 200 - 250$nm. The obtained scores are high, suggesting that epigenomic-driven interactions are main players of the chromosome organization in *Drosophila*. Interestingly, this corresponds to a situation where the system is still very dynamic [19]: interactions within and between TADs are stochastic and transient, leading to a plastic nuclear organization. However, there still exist many discrepancies between the predictions and the data suggesting missing ingredients that will drive further researches (see the Notes 3.3 and 3.4).

3. Notes

3.1 Regarding coarse-graining, our strategy allows to preserve the physical properties of the system by conserving the correct entanglement regime. However simpler strategies are often used in the field to make CG models: for example, by modeling chromatin as a simple flexible chain ($l_k' = \sigma'$) or by neglecting confinement and simulating isolated chains. Another strategy used to reduce the computation time is to

limit the study to only a small portion of the polymer. If not done properly, all these different strategies may modify the ratio *(L/L_e)* and therefore may modify the underlying physics of the null model [19].

3.2 Lattice modeling is a very efficient and powerful tool to simulate polymeric chains in various contexts. However, a main limitation of this implementation is that it is applicable as long as the involved energies (U or κ for example) are not too high. Indeed, discretization of the space implies that only a limited number of local moves are possible and that moves occur at the lattice resolution. This can be critical when energies are strong and when expected typical moves are only very small. In this case, one should better opt for off-lattice modeling like Molecular Dynamics.

3.3 While the correlation between the predicted epigenome folding of *Drosophila* and Hi-C experiments is remarkably high, discrepancies exist suggesting room for improvement of the model. One possibility is by allowing U to depend on the epigenomic state or by considering cross-epigenomic state interactions. For example, we recently showed that neglecting self-attraction between active regions may improve our description of the internal folding of active TADs [17]. Similarly, assuming that active monomers are repulsive ($U_{active} > 0$, ie have an effective bigger diameter, putatively mediated by transcription), HP1-like and Polycomb-like beads are self-attracting ($U_{hp1}, U_{PcG} < 0$) and black bead are neutral ($U_{black} = 0$) improves significantly the correlation with Hi-C data (Carrivain et al, under submission). Another possibility is to implement other mechanisms that might be of importance for shaping *Drosophila* nuclear organization. A main ingredient could be homologous pairing (maternal and paternal chromosomes are tightly paired [12,46]). While the underlying mechanism is still unknown, we recently showed that accounting for pairing by forcing homologous loci to occupy NN lattice sites may improve the description of intra-TAD

contacts [47]. Other possible important mechanisms that we did not explore yet include interactions with the nuclear lamina [48,49], formation of the nucleolus, or polymer loop extrusion mechanism [50,51] via cohesin or condensin II [46] [52].

3.4 The formalism of block copolymer used to study epigenome folding is not unique to *Drosophila* and after our seminal work in 2014 [39], us and other groups have used the same type of approach to model the impact of epigenomic-mediated interactions in mammals [53–55] (Sati et al, Mol. Cell, in press) but also in plant (Di Stefano et al, under submission).

Acknowledgements

I thank Noelle Haddad for providing unpublished figures from her PhD thesis, Cédric Vaillant for critical reading of the manuscript and Marco Di Stefano for fruitful discussions on the CG strategy. I acknowledge Agence Nationale de la Recherche (ANR-18-CE12-0006-03, ANR-18-CE45-0022-01) for funding.

Figure legends

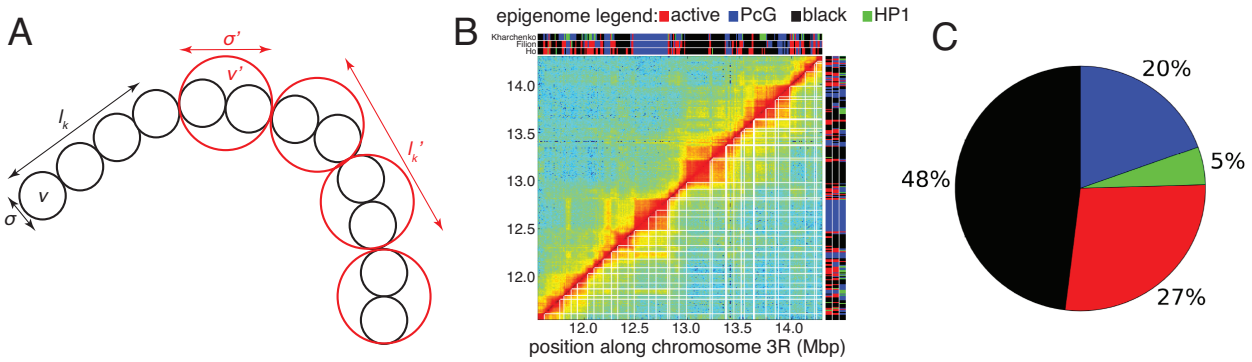

Figure 1. (A) Coarse-grained model: monomers of the fine-scale model (black circles) are lumped into bigger monomers (red circles). (B) Comparison between Hi-C experimental data of *Drosophila* chromosome 3R (from [6]) and 3 different partitionings of the genome into chromatin domains (from [27,34,36]). White lines in the bottom right represent the segmentation of the Hi-C map into TADs inferred by IC-Finder [25]. (C) Repartition of the different epigenomic state inferred by Filion et al [34] at the genome scale in *Drosophila*.

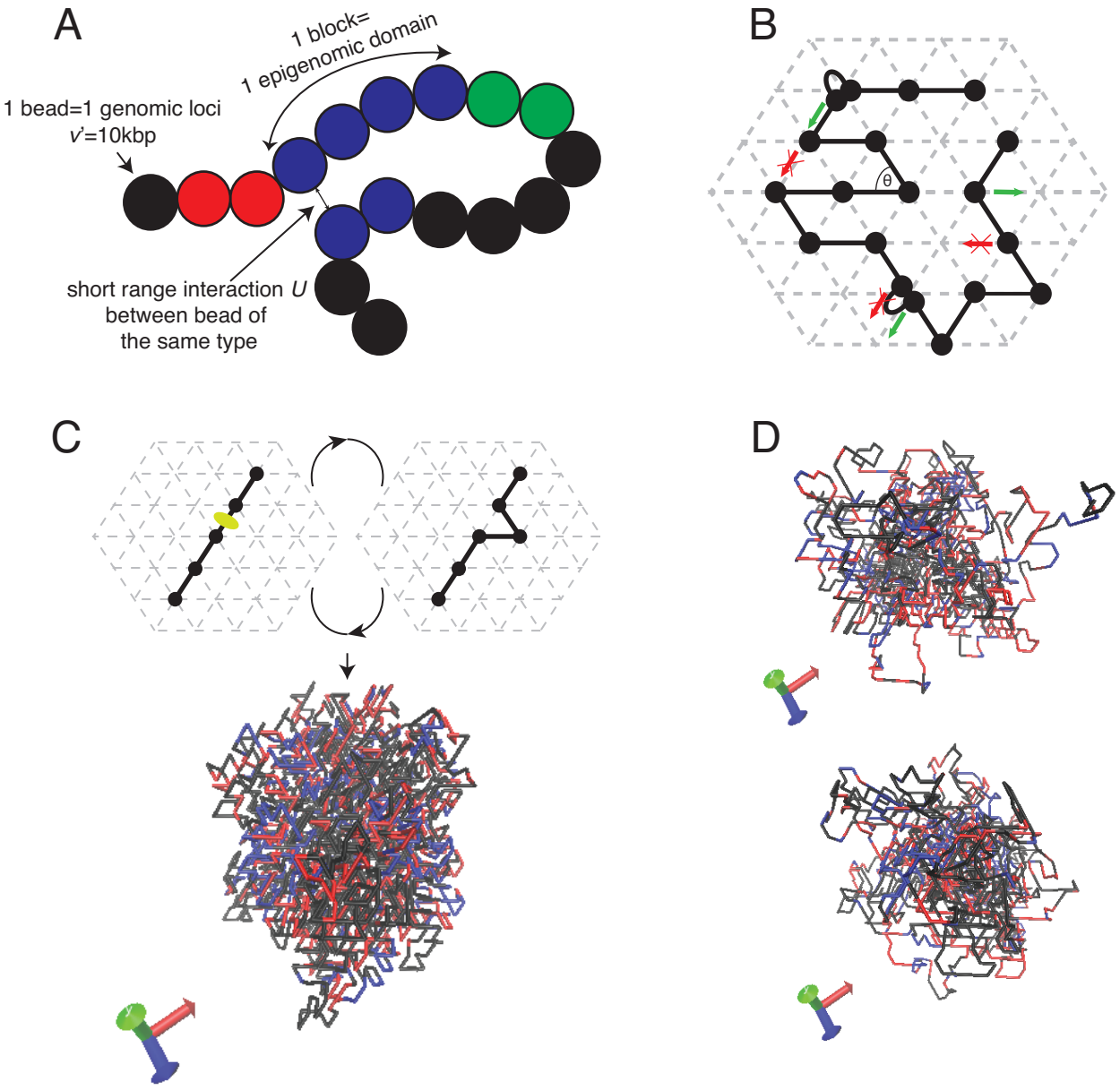

Figure 2. (A) Block copolymer model for epigenome folding with specific, short-range interactions between genomic loci sharing the same chromatin state. (B) 2D projection of a 3D fcc lattice (grey dashed lines). The solid line with full circles is a possible conformation of a polymeric chain. Semi-circular arcs indicate double-occupancy of consecutive monomers along the chain. Green (respectively red) arrows illustrate allowed (resp. forbidden) local moves. (C) The hedgehog algorithm to construct the initial configuration. Starting from a rod-like chain (left) containing few monomers, the structure is iteratively built by randomly choosing a link (green oval) and by

inserting a new monomer at a site closed to the two already placed monomers of the link (right) such that the new configuration still verifies lattice rules (excluded volume and connectivity). This step is repeated until the entire chain is grown (bottom). (D) Two examples of 3D configurations for a block copolymer model simulated for $U = -0.1k_BT$. Colors refer to the epigenomic state of the beads (same color code as in Fig.1B).

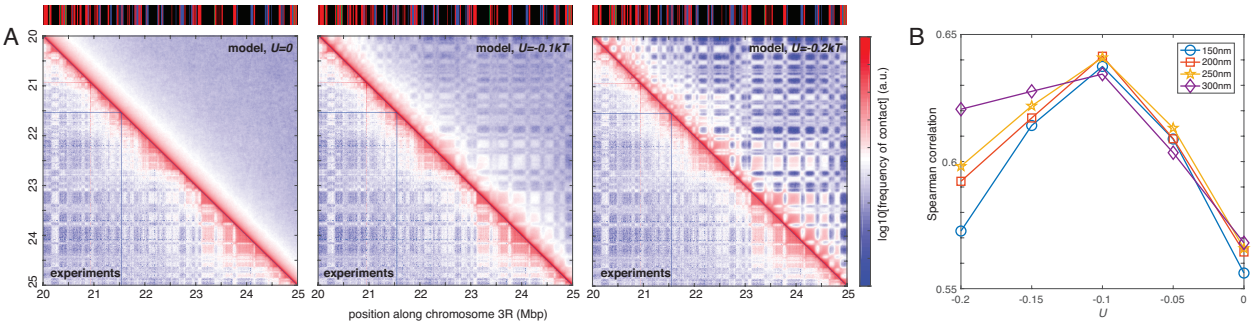

Figure 3. (A) Predicted Hi-C maps (triangular part on top right) for unsynchronized cells for different values of the strength of epigenomic interactions *U* ($r_c = 200$nm) and Hi-C experiments for late embryos from [6] (triangular part on bottom left). On the top, we showed the epigenomic state given by Filion et al [34] (same color code as in Fig.1B). (B) Spearman correlation computed between the entries of the predicted and experimental Hi-C maps. Predicted Hi-C maps were computed for different radius of contact r_c. We limit the computation to entries between loci separated by a genomic distance larger than 30kbp and lower than 5Mbp.

References

1. Jerković I, Szabo Q, Bantignies F, Cavalli G. Higher-Order Chromosomal Structures Mediate Genome Function. J Mol Biol. 2019. doi:10.1016/j.jmb.2019.10.014

2. McCord RP, Kaplan N, Giorgetti L. Chromosome Conformation Capture and Beyond:

Toward an Integrative View of Chromosome Structure and Function. Mol Cell. 2020;77: 688–708.

3. Nora EP, Lajoie BR, Schulz EG, Giorgetti L, Okamoto I, Servant N, et al. Spatial partitioning of the regulatory landscape of the X-inactivation centre. Nature. 2012;485: 381–385.

4. Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature. 2012;485: 376–380.

5. van Steensel B, Furlong EEM. The role of transcription in shaping the spatial organization of the genome. Nat Rev Mol Cell Biol. 2019;20: 327–337.

6. Sexton T, Yaffe E, Kenigsberg E, Bantignies F, Leblanc B, Hoichman M, et al. Three-dimensional folding and functional organization principles of the Drosophila genome. Cell. 2012;148: 458–472.

7. Hou C, Li L, Qin ZS, Corces VG. Gene density, transcription, and insulators contribute to the partition of the Drosophila genome into physical domains. Mol Cell. 2012;48: 471–484.

8. Rao SSP, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, et al. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. Cell. 2014;159: 1665–1680.

9. Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science. 2009;326: 289–293.

10. Hiraoka Y, Agard DA, Sedat JW. Temporal and spatial coordination of chromosome movement, spindle formation, and nuclear envelope breakdown during prometaphase in Drosophila melanogaster embryos. J Cell Biol. 1990;111: 2815–2828.

11. Fung JC, Marshall WF, Dernburg A, Agard DA, Sedat JW. Homologous chromosome pairing in Drosophila melanogaster proceeds through multiple independent initiations. J Cell Biol. 1998;141: 5–20.

12. AlHaj Abed J, Erceg J, Goloborodko A, Nguyen SC, McCole RB, Saylor W, et al. Highly structured homolog pairing reflects functional organization of the Drosophila genome. Nat Commun. 2019;10: 4485.

13. Hug CB, Grimaldi AG, Kruse K, Vaquerizas JM. Chromatin Architecture Emerges during Zygotic Genome Activation Independent of Transcription. Cell. 2017;169: 216–228.e19.

14. Ogiyama Y, Schuettengruber B, Papadopoulos GL, Chang J-M, Cavalli G. Polycomb-Dependent Chromatin Looping Contributes to Gene Silencing during Drosophila Development. Mol Cell. 2018;71: 73–88.e5.

15. Jost D, Vaillant C, Meister P. Coupling 1D modifications and 3D nuclear organization: data, models and function. Current Opinion in Cell Biology. 2017. pp. 20–27. doi:10.1016/j.ceb.2016.12.001

16. Tiana G, Giorgetti L. Modeling the 3D Conformation of Genomes. CRC Press; 2019.

17. Vaillant C, Jost D. Modeling the Functional Coupling between 3D Chromatin Organization

and Epigenome. Modeling the 3D Conformation of Genomes. 2019. pp. 21–56. doi:10.1201/9781315144009-2

18. Hyeon C, Thirumalai D. Capturing the essence of folding and functions of biomolecules using coarse-grained models. Nature Communications. 2011. doi:10.1038/ncomms1481

19. Ghosh SK, Jost D. How epigenome drives chromatin folding and dynamics, insights from efficient coarse-grained models of chromosomes. PLoS Comput Biol. 2018;14: e1006159.

20. Halverson JD, Smrek J, Kremer K, Grosberg AY. From a melt of rings to chromosome territories: the role of topological constraints in genome folding. Rep Prog Phys. 2014;77: 022601.

21. Pütz M, Kremer K, Grest GS. What is the entanglement length in a polymer melt? Europhysics Letters (EPL). 2000. pp. 735–741. doi:10.1209/epl/i2000-00212-8

22. Münkel C, Langowski J. Chromosome structure predicted by a polymer model. Physical Review E. 1998. pp. 5888–5896. doi:10.1103/physreve.57.5888

23. Socol M, Wang R, Jost D, Carrivain P, Vaillant C, Le Cam E, et al. Rouse model with transient intramolecular contacts on a timescale of seconds recapitulates folding and fluctuation of yeast chromosomes. Nucleic Acids Res. 2019;47: 6195–6207.

24. Uchida N, Grest GS, Everaers R. Viscoelasticity and primitive path analysis of entangled polymer liquids: from F-actin to polyethylene. J Chem Phys. 2008;128: 044902.

25. Haddad N, Vaillant C, Jost D. IC-Finder: inferring robustly the hierarchical organization of chromatin folding. Nucleic Acids Res. 2017;45: e81.

26. Zhu Y, Chen Z, Zhang K, Wang M, Medovoy D, Whitaker JW, et al. Constructing 3D interaction maps from 1D epigenomes. Nat Commun. 2016;7: 10812.

27. Ho JWK, Jung YL, Liu T, Alver BH, Lee S, Ikegami K, et al. Comparative analysis of metazoan chromatin organization. Nature. 2014;512: 449–452.

28. Canzio D, Liao M, Naber N, Pate E, Larson A, Wu S, et al. A conformational switch in HP1 releases auto-inhibition to drive heterochromatin assembly. Nature. 2013. pp. 377–381. doi:10.1038/nature12032

29. Larson AG, Elnatan D, Keenen MM, Trnka MJ, Johnston JB, Burlingame AL, et al. Liquid droplet formation by HP1α suggests a role for phase separation in heterochromatin. Nature. 2017;547: 236–240.

30. Strom AR, Emelyanov AV, Mir M, Fyodorov DV, Darzacq X, Karpen GH. Phase separation drives heterochromatin domain formation. Nature. 2017;547: 241–245.

31. Isono K, Endo TA, Ku M, Yamada D, Suzuki R, Sharif J, et al. SAM domain polymerization links subnuclear clustering of PRC1 to gene silencing. Dev Cell. 2013;26: 565–577.

32. Plys AJ, Davis CP, Kim J, Rizki G, Keenen MM, Marr SK, et al. Phase separation of Polycomb-repressive complex 1 is governed by a charged disordered region of CBX2. Genes Dev. 2019;33: 799–813.

33. Lu H, Yu D, Hansen AS, Ganguly S, Liu R, Heckert A, et al. Phase-separation mechanism for C-terminal hyperphosphorylation of RNA polymerase II. Nature. 2018;558: 318–323.

34. Filion GJ, van Bemmel JG, Braunschweig U, Talhout W, Kind J, Ward LD, et al. Systematic protein location mapping reveals five principal chromatin types in Drosophila cells. Cell. 2010;143: 212–224.

35. Ernst J, Kheradpour P, Mikkelsen TS, Shoresh N, Ward LD, Epstein CB, et al. Mapping and analysis of chromatin state dynamics in nine human cell types. Nature. 2011;473: 43–49.

36. Kharchenko PV, Alekseyenko AA, Schwartz YB, Minoda A, Riddle NC, Ernst J, et al. Comprehensive analysis of the chromatin landscape in Drosophila melanogaster. Nature. 2011;471: 480–485.

37. Roudier F, Ahmed I, Bérard C, Sarazin A, Mary-Huard T, Cortijo S, et al. Integrative epigenomic mapping defines four main chromatin states in Arabidopsis. EMBO J. 2011;30: 1928–1938.

38. Gerstein MB, Lu ZJ, Van Nostrand EL, Cheng C, Arshinoff BI, Liu T, et al. Integrative analysis of the Caenorhabditis elegans genome by the modENCODE project. Science. 2010;330: 1775–1787.

39. Jost D, Carrivain P, Cavalli G, Vaillant C. Modeling epigenome folding: formation and dynamics of topologically associated chromatin domains. Nucleic Acids Res. 2014;42: 9553–9561.

40. Rubinstein M. Discretized model of entangled-polymer dynamics. Physical Review Letters. 1987. pp. 1946–1949. doi:10.1103/physrevlett.59.1946

41. Newman ME, Strogatz SH, Watts DJ. Random graphs with arbitrary degree distributions and their applications. Phys Rev E Stat Nonlin Soft Matter Phys. 2001;64: 026118.

42. Hugouvieux V, Axelos MAV, Kolb M. Amphiphilic Multiblock Copolymers: From Intramolecular Pearl Necklace to Layered Structures. Macromolecules. 2009. pp. 392–400. doi:10.1021/ma801337a

43. Rosa A, Everaers R. Structure and dynamics of interphase chromosomes. PLoS Comput Biol. 2008;4: e1000153.

44. Imakaev MV, Tchourine KM, Nechaev SK, Mirny LA. Effects of topological constraints on globular polymers. Soft Matter. 2015;11: 665–671.

45. Olarte-Plata JD, Haddad N, Vaillant C, Jost D. The folding landscape of the epigenome. Phys Biol. 2016;13: 026001.

46. Rowley MJ, Jordan Rowley M, Lyu X, Rana V, Ando-Kuri M, Karns R, et al. Condensin II Counteracts Cohesin and RNA Polymerase II in the Establishment of 3D Chromatin Organization. Cell Reports. 2019. pp. 2890–2903.e3. doi:10.1016/j.celrep.2019.01.116

47. Pal K, Forcato M, Jost D, Sexton T, Vaillant C, Salviato E, et al. Global chromatin conformation differences in the Drosophila dosage compensated chromosome X. Nat Commun. 2019;10: 5355.

48. Tolokh IS, Kinney NA, Sharakhov IV, Onufriev AV. The Effect of Nuclear Envelope on Chromatin Architecture in Drosophila Melanogaster: Modeling of Three-Dimensional Interphase Chromosome Organization. Biophysical Journal. 2020. p. 551a. doi:10.1016/j.bpj.2019.11.3012

49. Ulianov SV, Doronin SA, Khrameeva EE, Kos PI, Luzhin AV, Starikov SS, et al. Nuclear lamina integrity is required for proper spatial organization of chromatin in Drosophila. Nat Commun. 2019;10: 1176.

50. Ghosh SK, Jost D. Genome organization via loop extrusion, insights from polymer physics models. Brief Funct Genomics. 2019. doi:10.1093/bfgp/elz023

51. Mirny L. Chromosome Organization by Loop Extrusion and Phase Separation. Biophysical Journal. 2019. p. 171a. doi:10.1016/j.bpj.2018.11.949

52. Rosin LF, Nguyen SC, Joyce EF. Condensin II drives large-scale folding and spatial partitioning of interphase chromosomes in Drosophila nuclei. PLoS Genet. 2018;14: e1007393.

53. Falk M, Feodorova Y, Naumova N, Imakaev M, Lajoie BR, Leonhardt H, et al. Heterochromatin drives compartmentalization of inverted and conventional nuclei. Nature. 2019;570: 395–399.

54. Di Pierro M, Cheng RR, Lieberman Aiden E, Wolynes PG, Onuchic JN. De novo prediction of human chromosome structures: Epigenetic marking patterns encode genome architecture. Proc Natl Acad Sci U S A. 2017;114: 12126–12131.

55. Qi Y, Zhang B. Predicting three-dimensional genome organization with chromatin states. PLOS Computational Biology. 2019. p. e1007024. doi:10.1371/journal.pcbi.1007024