Effective Model of Loop Extrusion Predicts Chromosomal Domains

Martina Crippa

Department of Physics, Università degli Studi di Milano,

via Celoria 16, 20133 Milano, Italy and

Department of Applied Science and Technology, Politecnico di Torino,

Corso Duca degli Abruzzi 24, 10129 Torino, Italy

Yinxiu Zhan

Friedrich Miescher Institute for Biomedical Research,

Maulbeerstrasse 66, 4058 Basel, Switzerland

Guido Tiana*

Department of Physics and Center for Complexity and Biosystems,

Università degli Studi di Milano and INFN,

via Celoria 16, 20133 Milano, Italy

(Dated: September 4, 2020)

Abstract

An active loop-extrusion mechanism is regarded as the main out–of–equilibrium mechanism responsible for the structuring of megabase-sized domains in chromosomes. We developed a model to study the dynamics of the chromosome fibre by solving the kinetic equations associated with the motion of the extruder. By averaging out the position of the extruder along the chain, we build an effective equilibrium model capable of reproducing experimental contact maps based solely on the positions of extrusion–blocking proteins. We assessed the quality of the effective model using numerical simulations of chromosomal segments and comparing the results with explicit-extruder models and experimental data.

I. INTRODUCTION

Chromosomes display a hierarchical structure of domains during cellular interphase [1, 2]. In mammals, the level of topological associating domains (TADs), at the mega–base scale, constitutes the most important level in the hierarchy for their role in controlling gene expression. The folding of TADs has been described at a molecular level by an active loop–extrusion mechanism [3], where a protein complex extrudes chromatin loops and it can be stopped by proteins bound to chromosome (for a review, see refs. [4, 5]).

The cohesin protein complex has been suggested to extrude the chromatin fiber, keeping close in space the two chromosomal segments at which it is bound at a given time (see Fig. 1). The extrusion activity can be stopped by CTCF proteins bound to chromatin, thus stabilizing the contact between the CTCF-bound chromosomal regions. In fact, enrichment in CTCF has been observed in loci pivoting strong contacts [6]. Cells lacking either CTCF [7] or cohesin [8] display a reduced structuring of TADs. Recently, microscopy experiments using biochemically reconstructed systems showed that cohesin can extrude chromatin in an ATP–dependent way [9, 10].

An interesting feature of CTCF is that it is directional, in the sense that it can bind asymmetrically to chromatin in both directions and can stop efficiently cohesin only if it oriented towards it, but not those oriented opposite to it [11]. This directionality arises because CTCF is not simply a barrier to the motion of cohesin, but interacts with it in

---

* guido.tiana@unimi.it

a specific way. CTCF binds to DNA in a non-palindromic way [12] and a segment in its terminal segment interacts strongly with a specific domain of cohesin [13]. The directionality of CTCF seems to be at the basis of "corner peaks" observed in contact maps of mammalian cells at the scale of 100 kbp [3, 14].

Several polymeric models have been employed to describe the conformation and the dynamics of chromatin at the length scale of TADs [15]. They usually describe the chromosomal segments as a chain of beads interacting with some contact potential. Simulations of polymer chains including an additional degree of freedom that specify the position of extruders along the chain were shown to produce contact maps which are qualitatively similar to the experimental ones [3, 11, 16]. Polymeric simulations of a diffusing extruder produced realistic contact maps also without energy consumption [17, 18], even if experimental data suggest that ATP hydrolysis is a key ingredient for extrusion [9, 10]. A mechanistic model for ATP–dependent translocation of cohesin is described in ref. [19].

In the present work we studied the active, out–of–equilibrium dynamics of cohesin along the chromatin fiber and we built an effective model in which the position of cohesin along the chain is averaged out. In this way, we obtained a polymer model controlled by an effective potential whose equilibrium state reflect the distribution of conformations in cellular nuclei.

The reason why to build such an effective model is two–fold. On one side, it can be helpful to better understand the physics that controls the conformational properties of chromosomes. In fact, the reduction of the degrees of freedom maps chromosomes into systems that can be studied by standard polymer theory. On the other side, it can be a predictive tool to generate contact maps of chromosomal regions based on the position of CTCF, in a computationally more efficient way than explicit-extruder models.

II. MOTION OF THE EXTRUDER

Some of the numerical parameters that are necessary to build the model are known, see Appendix A. In particular, the diffusion coefficient of cohesin in the nucleoplasm is much larger than that of chromatin loci on the TAD length scale, suggesting that one can assume cohesin to be well-mixed in cellular nucleus. Moreover, the time scale associated with extrusion is slightly smaller than that associated with the motion of the polymer chain on the TAD length scale. Even if this difference is marginal, we tested the assumption that
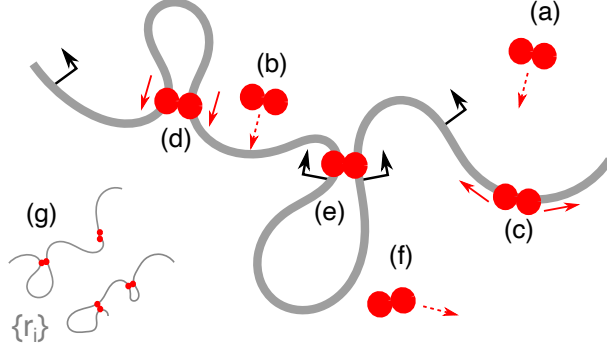
3

FIG. 1. A sketch of the loop–extrusion mechanism. Cohesin diffuses in the nucleus (a) and can be loaded onto the chromosome at a random position (b). From here, it starts to run on the chromosome (c), extruding its two strings and thus forming a loop (d). CTCF proteins bound to chromosomes towards the running cohesin (black arrows) can stop tits motion (e). At any point, cohesin has a probability to detach from the chromosome (f). (g) The position of cohesin along the chain define the interactions that contribute to determining the conformations $\{r_i\}$ of the chromosome.

the distribution of cohesin along the chain can be regarded as stationary. We compared the results of the effective model with both the experimental data and polymer model in which cohesin in simulated explicitly, thus without making in this case any assumption on its probability distribution.

Let's assume that the extruder can only walk towards the ends of the chain, that it walks with constant rate in a fixed direction and that it cannot overcome a CTCF molecule. Let's define the binary quantities σ_i^+ and σ_i^- that assume the values 1 if site i contains a CTCF molecule oriented forward and backward, respectively. We also define

$$\tilde{\delta}_i^{\pm} \equiv 1 - \delta_{\sigma_i^{\pm},1} \tag{1}$$

that assumes the value 0 in the sites with a CTCF molecule oriented in the specified direction and thus it is able to stop the motion of the extruder in that direction; it takes the value 1 otherwise.

The rate equation that describes the amount $p_{i,j}(t)$ of extruder linking sites i and j of

the chromosomes is

$$\frac{dp_{i,j}}{dt} = k_{\text{on}}\delta_{|i-j|,1} - k_{\text{off}}p_{i,j} + k\tilde{\delta}^-_{i+1}p_{i+1,j}-$$

$$- k\tilde{\delta}^-_i p_{i,j} + k\tilde{\delta}^+_{j-1}p_{i,j-1} - k\tilde{\delta}^-_j p_{i,j}, \tag{2}$$

where k_{on} is the loading rate of the extruder on the chromosome, k_{off} the detachment rate and k the advancement rate. The stationary distribution can be obtained setting to zero the time derivative for every i and j, that is

$$p_{ij} = \frac{k_{\text{on}}\delta_{|i-j|,1} + k\tilde{\delta}^-_{i+1}p_{i+1,j} + k\tilde{\delta}^+_{j-1}p_{i,j-1}}{k_{\text{off}} + k(\tilde{\delta}^-_i + \tilde{\delta}^+_j)}. \tag{3}$$

This equation can be solved recursively, exploiting the fact that p_{ij} depends only on the probabilities p_{kl} such that $i < k < l < j$.

An important approximation that we implicitly did in Eq. (2) is that multiple extruders do not interact with each other by excluded volume when they walk on the chromosome.

A. Chromosome without CTCF

The simplest case is that in which the extruder can walk freely on the chromosome in absence of CTCF, as described in Fig. 2(a).

In this case, Eq. (3) becomes

$$p_{i,j} = \frac{k_{\text{on}}\delta_{|i-j|,1} + kp_{i+1,j} + kp_{i,j-1}}{k_{\text{off}} + 2k}. \tag{4}$$

Starting from the case $i, i+1$ at which the extruder can bind, one can write iteratively

$$p_{i,i+1} = \frac{k_{\text{on}}}{k_{\text{off}} + 2k} \equiv p_0,$$

$$p_{i,i+2} = \frac{kp_{i+1,i+2} + kp_{i,i+1}}{k_{\text{off}} + 2k} = \frac{2k}{k_{\text{off}} + 2k}p_0,$$

$$p_{i,i+3} = \frac{kp_{i+1,i+3} + kp_{i,j+2}}{k_{\text{off}} + 2k} = \left(\frac{2k}{k_{\text{off}} + 2k}\right)^2 p_0$$

$$\dots$$

$$p_{i,j} = \left(\frac{2k}{k_{\text{off}} + 2k}\right)^{j-i-1} p_0, \tag{5}$$

where use is made of the translational invariance $p_{i+n,j+n} = p_{ij}$ and the boundary condition $p_{i,i} = 0$.
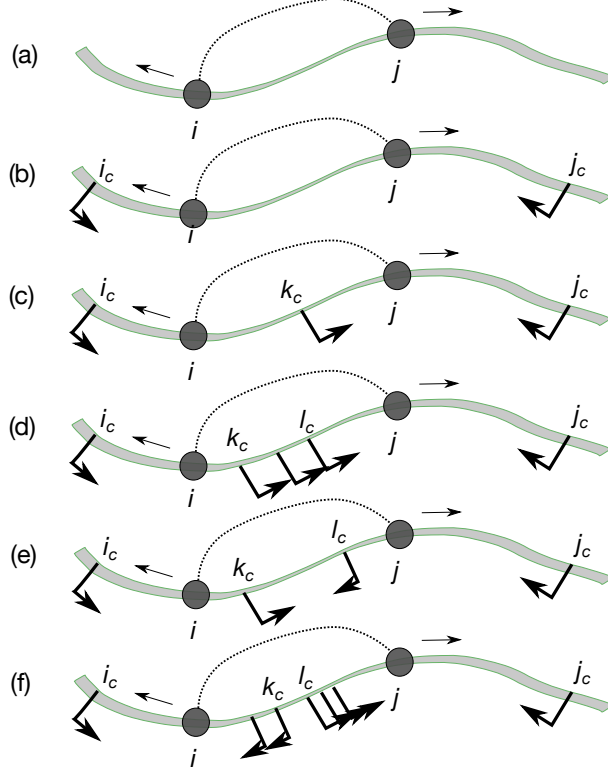
FIG. 2. A sketch of the different ways in which cohesin can run, according to the position of CTCF. The pairs of bead indicate the position of cohesin at the two chromosomal sites it encloses. (a) A chromosomal segment where the extruder can move freely. (b) The case in which the extruder is constrained by two convergent CTCF molecules (i.e., $\sigma_{i_c}^+ = 1$ and $\sigma_{j_c}^- = 1$). (c) The case in which a further CTCF molecule prevents the motion of the extruder in one direction (i.e., $\sigma_{i_c}^+ = 1$, $\sigma_{k_c}^+ = 1$ and $\sigma_{j_c}^- = 1$). (d) Is the case similar to the previous one, with multiple aligned CTCF. (e) The case with convergent CTCF in between. (f) The case of several divergent CTCF in between, the inner being at positions k_c and l_c, respectively.

B. Contacts between sites within convergent CTCF

Consider a chromosome segment bordered by convergent CTCF at sites i_c and j_c, as in Fig. 2(b). The value of p_{ij} with $i_c < i < j < j_c$ depends only on the amount of extruder in the interval from i to j, so for $i_c < i < j < j_c$ Eq. (5) still holds.

Equation (3) can now be written as

$$p_{ij} = \frac{k_{\text{on}}\delta_{|i-j|,1} + k(1 - \delta_{i+1,i_c})p_{i+1,j} + k(1 - \delta_{j-1,j_c})p_{i,j-1}}{k_{\text{off}} + k(2 - \delta_{i,i_c} - \delta_{j,j_c})}. \tag{6}$$

The amount of extruder in sites containing a CTCF molecule can be found from Eqs. (6)

6

and (5). For example, the term

$$p_{i,j_c} = \frac{k p_{i+1,j_c} + k p_{i,j_c-1}}{k_{\text{off}} + k},$$ (7)

where p_{i,j_c-1} is that of Eq. (5) and we iterate on p_{i+1,j_c}. We get from Eqs. (7) and (6)

$$p_{j_c-1,j_c} = \frac{k_{on}}{k_{\text{off}} + k}$$

$$p_{j_c-2,j_c} = \frac{k}{k_{\text{off}} + k} \left[\frac{k_{on}}{k_{\text{off}} + k} + p_0 \right]$$

$$p_{j_c-3,j_c} = \frac{k_{on}}{k_{\text{off}} + k} \left(\frac{k}{k_{\text{off}} + k} \right)^2 + \left(\frac{k}{k_{\text{off}} + k} \right)^2 p_0 +$$

$$+ \frac{k_{on}}{k_{\text{off}} + k} \left(\frac{k}{k_{\text{off}} + 2k} \right)^2 p_0$$

$$p_{j_c-n,j_c} = \frac{k_{on}}{k_{\text{off}} + k} \left(\frac{k}{k_{\text{off}} + k} \right)^{n-1} +$$

$$+ \sum_{l=1}^{n-1} \left(\frac{k}{k_{\text{off}} + k} \right)^{n-l} \left(\frac{k}{k_{\text{off}} + 2k} \right)^{l-1} p_0.$$ (8)

The general form of p_{j_c-n,j_c} contains a geometric sum that gives

$$p_{j_c-n,j_c} = \frac{k_{on}}{k_{\text{off}} + k} \left(\frac{k}{k_{\text{off}} + k} \right)^{n-1} +$$

$$+ \frac{k_{\text{off}} + 2k}{k} \left(\frac{k}{k_{\text{off}} + k} \right)^{n-1} p_0 - \left(\frac{k}{k_{\text{off}} + 2k} \right)^{n-2} p_0$$ (9)

By symmetry, the same expression is valid for p_{i_c,i_c+n}. The probability associated with both CTCF sites obeys by Eq. (6) the relation

$$p_{i_c,j_c} = \frac{k p_{i_c+1,j_c} + k p_{i_c,j_c-1}}{k_{\text{off}}},$$ (10)

that can be evaluated substituting Eq. (9) in it.

C. Contacts across a CTCF site

Consider a segment from i_c to j_c closed by convergent CTCF molecules, with a further CTCF molecule at position k_c with $i_c < k_c < j_c$ and, for instance, directed upward (i.e., $\sigma_{k_c}^+ = 1$), as in Fig. 2(c).

Pairs of sites on the same side with respect to k_c displays the same probabilities as described above, that is Eqs. (4), (9) and (10). Pairs interspersed by CTCF molecules,

i.e. $i < k_c < j$, are affected by the fact that the two sites cannot be reached evenly from extruders coming from all parts of the segment (i, j).

Let's use again an iterative approach, starting from

$$p_{k_c-1,k_c} = p_0. \tag{11}$$

The probabilities involoving site $k_c + 1$ obey

$$p_{k_c-n,k_c+1} = \frac{k}{k_{\text{off}} + 2k} p_{k_c-n+1,k_c+1} =$$
$$= (2^n - 1) \left(\frac{k}{k_{\text{off}} + 2k} \right)^n p_0. \tag{12}$$

Similarly, those involving site $k_c - 1$ are given by Eq. (5),

$$p_{k_c-1,k_c+m} = \left(\frac{k}{k_{\text{off}} + 2k} \right)^m p_0. \tag{13}$$

For any pair of sites across k_c, the probability obeys the iterative relation

$$p_{k_c-n,k_c+m} = \frac{k}{k_{\text{off}} + 2k} \left(p_{k_c-n+1,k_c+m} + p_{k_c-n,k_c+m-1} \right). \tag{14}$$

One can look for solutions in the form

$$p_{k_c-n,k_c+m} = a_{n,m} \left(\frac{k}{k_{\text{off}} + 2k} \right)^{n+m-1} p_0, \tag{15}$$

that, substituted in Eq. (14), gives the iterative relation

$$a_{n,m} = \frac{k}{k_{\text{off}} + 2k} (a_{n-1,m} + a_{n,m-1}), \tag{16}$$

starting from $a_{n,1} = 2^n - 1$ (Eq. 12) and $a_{1,m} = 1$ (Eq. 13).

Solving the iterative problem making use of a bivariate generating function (see Appendix B), one obtains

$$p_{k_c-n,k_c+m} = \binom{n+m-1}{m} {}_2F_1(1, 1-n, 1+m, -1) \cdot$$
$$\cdot \left(\frac{k}{k_{\text{off}} + 2k} \right)^{m+n-1} p_0, \tag{17}$$

where ${}_2F_1$ is the Gaussian hypergeometric function.

8

D. Contacts across several CTCF sites

Consider now the case of a pair of sites i and j separated by more than a CTCF molecule, with various orientations, like in Figs. 2(d)-(e). If both orientations are present, like in Fig. 2(e), then $p_{i,j} = 0$ because no extruder can bind to any pair of sites $q, q+1$ with $i < q < j$ and reach sites i and j.

For sites i and j separated by two CTCF sites (at positions k_c and l_c) with the same alignment, as in Fig. 2(d) one can follow the same strategy as that of Sect. II C.

Analogously to Eq. (11), the starting point is the probability p_{k_c-1,l_c} that in the present case is given by Eq. (13) because sites $k_c - 1$ and l_c fall in the case of Fig. 2(c), that is

$$p_{k_c-1,l_c} = \left(\frac{k}{k_{\text{off}} + 2k} \right)^{l_c - k_c} p_0. \tag{18}$$

From here, an iterative relation analogous to Eq. (14) holds, that is

$$p_{k_c-n,l_c+m} = \frac{k}{k_{\text{off}} + 2k} \left(p_{k_c-n+1,l_c+m} + p_{k_c-n,l_c+m-1} \right), \tag{19}$$

whose solution is the same as that of Eq. (17),

$$p_{k_c-n,l_c+m} = \binom{n+m-1}{m} {}_2F_1(1, 1-n.1+m, -1) \cdot$$
$$\cdot \left(\frac{k}{k_{\text{off}} + 2k} \right)^{m+n+l_c-k_c-1} p_0, \tag{20}$$

with the difference that the iterative propagation is applied to Eq. (18) instead that to p_0 only.

This solution can be easily extended to the case in which between the two sites of interest there is an arbitrary sequence $\{k_{c1}, k_{c2}, ..., k_{cN}\}$ of CTCF sites aligned in the same direction. In this case, one can apply the propagator of Eq. (20) to $p_{k_{c1,c(N-1)}}$ obtaining

$$p_{k_{c1}-n,k_{cN}+m} = \binom{n+m-1}{m} {}_2F_1(1, 1-n.1+m, -1) \cdot$$
$$\cdot \left(\frac{k}{k_{\text{off}} + 2k} \right)^{m+n+cN-c1-1} p_0, \tag{21}$$

thanks to the fact that ${}_2F_1(1, 0, 1+m, -1) = 1$.

The most problematic case is that of two sites i and j separated by diverging CTCF molecules, like in Fig. 2(f). Calling k_c and l_c, respectively, the inner sites, we know that

$$p_{k_c-1,l_c+1} = \frac{k}{k_{\text{off}} + 2k} [p_{k_c-1,l_c} + p_{k_c,l_c+1}] \tag{22}$$

9

that can be easily evaluated using Eq. (20). However, the exact solution of this case for generic values of n and m would require the summation of terms in the form of Eq. (20), that we are not able to do. For this reason we resort to an approximation, writing

$$p_{k_c-n,l_c+m} = p_{k_c,l_c+m} \left(\frac{k}{k_{\text{off}} + k} \right)^n + p_{k_c-n,l_c} \left(\frac{k}{k_{\text{off}} + k} \right)^m , \qquad (23)$$

in which the probabilities at the right–hand side are given by Eq. (20). This corresponds to the assumption that the extruders that can reach sites i and j are only those that after reaching sites k_c and $l_c + m$ walks n steps on to reach $k_c + n$, and those that do the same thing from sites $k_c + n$ and l_c, making m steps from the former. Equation (23) is exact for $n = m = 1$ and is expected to underestimate the true probability for large n and m, probability that is anyway low in this limit. Moreover, under the same assumptions, the resulting probability does not change if multiple CTCF sites are aligned in the two directions, as in Fig. 2(f).

III. EFFECTIVE MODEL

From the knowledge of the stationary distribution of extruder, we built an effective polymeric model in which the degrees of freedom of the extruder are averaged out. In other words, we started from a model which is surely out of equilibrium because the motion of cohesin does not obey the condition of detailed balance, we showed that the distribution of cohesin along the chain has a stationary distribution and investigated if there is an effective potential that displays that distribution at equilibrium, through Boltzmann statistics. Then, we used the parameters of this potential in the conformational space of the polymer, in connection with a realistic (although arbitrary) contact function that defines the spatial dependence of the potential.

Let's assume that the number μ_{ij} of extruder molecules binding sites i and j of the chromosome can be written as an equilibrium state of an effective potential

$$P(\mu) = \frac{1}{Z_\mu} e^{-\sum_{ij} \epsilon_{ij} \mu_{ij}} , \qquad (24)$$

where ϵ_{ij} is a site-dependent effective energy.

Due to the rigid nature of the extruder, the conditional probability associated with a

10

conformation $\{r_i\}$ of the system for any given state $\{\mu_{ij}\}$ of the extruder along the chain is

$$P(r|\mu) = \frac{1}{Z} \prod_{i<j} \delta(\Delta_{ij}(r) - \mu_{ij}) e^{-\beta U_0(r)}, \tag{25}$$

where $U_0(r)$ is an underlying potential describing excluded volume and other general features of the polymer, $\beta \equiv 1/kT$ and $\Delta(r_{ij})$ is some function that defines the approaching in space of two monomers. The functional form of $\Delta(r_{ij})$ is arbitrary and cannot be determined within the present theory; due to the physical features of the system, one can envisage some kind of short-range potential.

The marginal probability of a conformation, averaged over the extruder conformations is then

$$P(r) = \int d^{N^2}\mu \; P(r|\mu) \cdot P(\mu) \tag{26}$$

Writing the delta function as $\delta(x) = \lim_{\kappa \to 0} \exp[x^2/2\kappa^2]$, the conformational probability is proportional to

$$\prod_{i<j} \int d\mu_{ij} \; \exp\left[-\epsilon_{ij}\mu_{ij} - \frac{(\Delta_{ij}(r) - \mu_{ij})^2}{2\kappa^2} - \beta U_0(r)\right]$$

that is a Gaussian integral, that gives

$$P(r_i) = \exp\left[-\beta U_0(r) - \sum_{i<j} \epsilon_{i,j} \Delta_{i,j}(r)\right], \tag{27}$$

where the term proportional to $\kappa \to 0$ has been dropped.

The effective potential has thus the form

$$U(r) = U_0(r) + k_B T \sum_{i<j} \epsilon_{i,j} \Delta_{i,j}(r) \tag{28}$$

whose parameters can be found from Eq. (24) as

$$\epsilon_{i,j} = -k_B T \log P(\mu_{ij}) - E_0 \tag{29}$$

and if the number of extruders is small (and thus μ_{ij} is essentially binary), the probabilities $p(\mu_{ij})$ can be regarded as proportional to the results $p_{i,j}$ of the rate equations calculated in Eqs. (5), (9), (10), (17), (21) and (23), according to the position and the orientation of CTCF. The arbitrary additive constant E_0 ($= -k_B T \log Z_\mu$) in the energies ϵ_{ij} must be set independently of the theory developed above (see below). Positive values of ϵ_{ij} are filtered to zero, because it is not realistic that the extruder induces a repulsive interaction in the polymer.

IV. SIMULATIONS WITH THE EFFECTIVE MODEL

To test the performance of the effective model, we performed molecular–dynamics simulations of chromosomal segments described by a chain of beads connected by springs and interacting with a potential $U = U_0 + U_{\text{eff}}$ (cf. Eq. 28) given by a polymeric term

$$U_0 = \frac{k_s}{2}(\Delta r_{ij} - a)^2 + \epsilon_0 \sum_{i<j} \left(\frac{2a^6}{\Delta r_{ij}^6} - \frac{a^{12}}{\Delta r_{ij}^{12}} \right), \tag{30}$$

where $\Delta r_{ij} \equiv |r_i - r_j|$, and the effective potential

$$U_{\text{eff}} = \sum_{i<j} \epsilon_{i,j} \left(\frac{2a^6}{\Delta r_{ij}^6} - \frac{a^{12}}{\Delta r_{ij}^{12}} \right) \tag{31}$$

representing the effect of the extruder. The rest distance a of the harmonic spring sets the elementary scale of the system, that is $a = 67 \ nm$ corresponding to a resolution of $5 \cdot 10^3 \ bp$ (cf. appendix A). The interaction range of the Lennard–Jones potentials is set to a as well. All simulations are performed at room temperature, $k_B T = 2.5 \ kJ/mol$. The harmonic constant is set to allow 10% fluctuations of the spring length, that is $k_s = 10^2 k_B T/a^2 = 250 \ kJ/mol/a^2$. The value of ϵ_0 is set to $-1.5 \ kJ/mol$ so that simulations in absence of extruders ($U_{\text{eff}} = 0$) display the polymer fragment at the θ–point. This is the simplest choice assuming that there are no other active mechanisms besides cohesin extrusion. In this case, the chromosome fragment we simulated mimics a segment of a much larger polymeric system and thus should obey ideal–chain statistics, in accordance with Flory theorem [20]. To be noted that experiments depleting cohesin display a contact probability that scales with the linear distance as a power law with exponent 1.2 [21], which is not that of an ideal chain (1.5), suggesting that some other out-of-equilibrium mechanism is at work. Choosing to simulate the polymer at the θ–point we neglect these other mechanisms.

Simulations are performed solving Langevin equations with Euler's integrator. For time steps Δt much larger than $m/\gamma \approx 10^{-4} ps$ (see Appendix A) one can use the first–order overdamped version of Langevin equations. The fastest degree of freedom is that associated with the harmonic springs defining the chain, so we expect the time step to be smaller than the associated time scale, that is $\Delta t < \gamma/k_s \sim 10^{10} \ ps$. We evaluated the quality of the simulation quantifying to which extent it satisfies the principle of detailed balance. For this purpose, we calculated the quantity

$$\tilde{H} \equiv \frac{\Delta t}{4\gamma} \left(\frac{\partial U}{\partial r} \right)^2 - \frac{r}{2} \frac{\partial U}{\partial r} + U(r) \tag{32}$$

12

that is a parameter (in energy units) which is strictly conserved if detailed balance is satisfied [22], in the present case by the Euler integrator. In simulations performed with $\Delta t \leq 10^7\ ps$ $\tilde{H}$ is conserved within an error of $k_B T$ (see Fig. 3(a)). For larger values of Δt this is no longer the case, and thus the simulations are no longer correct. We used $\Delta t = 10^7\ ps$ in the rest of the simulations.

We first applied the effective model to a small region of mouse embryonic stem cells (of coordinates ChrX:100378307-100702306, the so–called Tsix TAD). The position and orientation of CTCF are taken from ref. [23]. The map of the interaction energies ϵ_{ij} is dislayed in Fig. 3(f); the main patterns displayed by the Hi–C map (cf. Fig. 3(b)) are already apparent here. The average contact map $\overline{\Delta}_{ij}$ obtained from simulations of 16 minutes each is displayed in Fig. 3(c) and is compared with the experimental map obtained from Hi–C experiments (Fig. 3(b), [24]). After this time span, contact maps are at convergence and one can expect that the system is equilibrated.

In this procedure there are two free parameters that have to be set, namely the additive constant E_0 associated with the interaction energies (cf. Eq. (29)) and the cut–off distance R_{cont} for two beads to be defined as in contact, distance that can hardly be obtained from a molecular insight of the crosslinking process at the basis of Hi–C maps,. These parameters are then obtained maximizing the Pearson's correlation function r between all pairs of contacts in the simulated and in the experimental contact map. The correlation coefficient r as a function of E_0 is displayed in Fig. 3(d) and display a maximum at $E_0 = 14$. The dependence on R_{cont} is displayed in Fig. 3(e) and is optimal at $R_{cont} = 1.7a$. The optimal correlation obtained with these values is $r = 0.89$. The simulated contact map displays the main features of the experimental Hi–C map, including two regions with high contact probability (cf panels b and c in Fig. 3). Moreover, the simulated map displays strong contacts in the initial part of the polymer which are not present in the experimental map; this is a region lacking of any CTCF molecule.

To better understand the effective model, we simulated a toy model made of a 30-bead string with two convergent CTCF sites, as displayed in Fig. 4(a). The effective interactions ϵ_{ij} display a square of strongly interacting elements within the two CTCF sites, see Fig. 4(b); the borders of this square are even more interacting, as well as the corner where the extruder accumulates. In addition, there is a sort of 'border effect' due to extruders that bind to the CTCF–free ends of the chain. The simulation of the effective model produces a
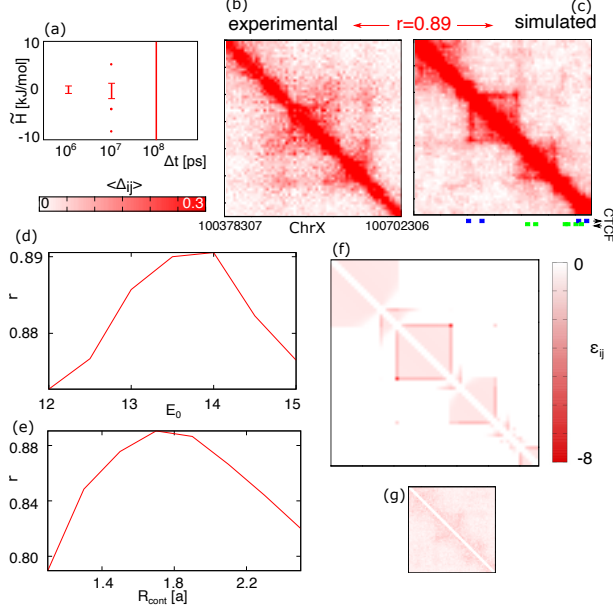
13

FIG. 3. (a) The width of fluctuations of the effective energy $\tilde{H}$ as a function of the time step Δt of the simulation, in ps. (b) The experimental Hi–C map of the Tsix region. (c) The contact map simulated with the effective model. The positions of CTCF in both orientations are indicated below the map. The correlation coefficient r between simulated and experimental contact map as a function of the shift E_0 of the energy parameters (d) and of the distance R_{cont} that defines contacts (e). (f) The interaction energy ϵ_{ij} between the beads. (g) The standard deviation between 10 simulations of 16 minutes each, plotted in the same color scale of the contact maps.

contact map which reflects essentially the interaction potential (cf. 4(c)).

V.   COMPARISON WITH EXPLICIT EXTRUDER MODEL

A relevant question we want to answer is how the the effective model performs with respect to a model in which the extruder is described explicitly [3]. In fact, we do not expect that the effective model can reproduce all details of Hi–C maps, because loop extrusion is not the only mechanism at work. For example, it is known that the formation of compartments on the scale of the whole chromosome is not driven by loop extrusion but interacts with it at a smaller scale [21]. Our main goal is then to show that the effective model can reproduce Hi–C map with the same accuracy as the explicit–extruder model.

In the explicit model, we assumed that the extruder is well–mixed around the polymer
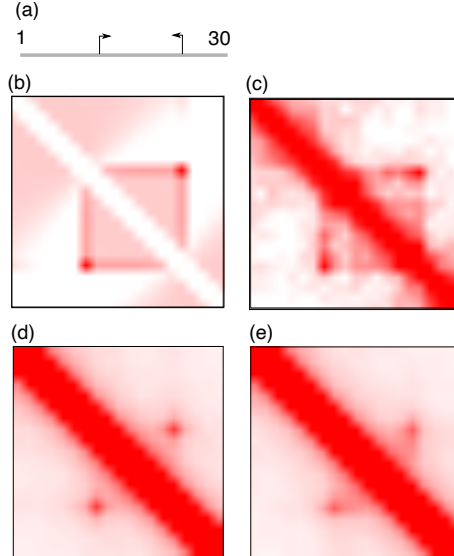
FIG. 4. (a) A toy model with two convergent CTCF molecules. (b) The effective contact energy ϵ_{ij}. (c) The result of a simulation with the effective model of 16 minutes with $E_0 = 13$. (d) The contact map resulting from the average of 30 explicit-extruder simulations of 16 minutes each in which extruders can overcome each other freely. (e) The result of simulations in which extruders cannot move into occupied sites.

and it is always available for binding. It can bind to a pair of adjacent sites with rate k_{on}, each side of the extruder can walk with rate k and it can detach with rate k_{off}. The monomers linked by an extruder experience a harmonic force characterized by a harmonic constant k_s and a rest distance a, that is the same force that guarantees the integrity of the polymer. We assumed that different bound extruders cannot overcome each other and they cannot overcome CTCF sites. The numerical parameters are given in Appendix A.

The average contact map obtained from 30 simulations, calculated in the same way as those obtained with the effective model (cf. Sect. IV), is displayed in Figs. 5(a) and (b) for the Tsix domain. Contact maps seem to be at convergence. The correlation coefficient with the experimental map is $r = 0.89$, which is identical to that of the effective model.

The main difference between the explicit and the effective model is in the fluctuations around the average. In Fig. 5(c) we showed the result of three individual simulations and in Fig. 5(d) the standard deviation associated with the simulations. It is apparent that in explicit–extruder simulations the average map is given by the contribution of maps which are quite different from each other. In fact, the standard deviation is comparable with the
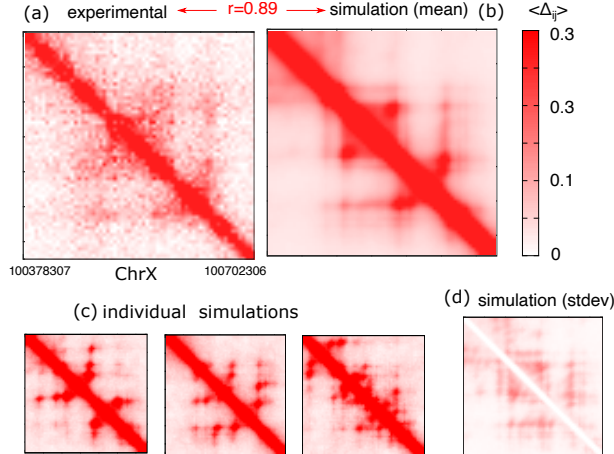
15

FIG. 5. Results of the simulation of the Tsix TAD with explicit extruders. (a) The experimental map. (b) The mean contact function $\langle\Delta_{ij}\rangle$, averaged over 30 simulations (c) Examples of individual simulations that contribute to the average. (d) Their standard deviation.

average. This result is different than that of the effective model, in which the different simulations generate maps which are much more homogeneous (cf. Fig. 3(g)).

In the case of the toy model of Fig. 4, explicit-extruder simulations produce contact maps in which corner peaks are more evident, and the overall domain is less clear. We also compared the results of simulations in which extuders are freely allowed to overcome each other (Fig. 4(d)) with simulations in which an extruder cannot occupy a site which is already occupied (Fig. 4(e)). The two maps are essentially identical, suggesting that the hypothesis done in connection with Eq. (4) is not critical.

We repeated similar calculations for other two regions of the chromosome X of mouse embryonic stem cells, of 1300 and 2600 kbp, respectively. The energy maps ϵ_{ij} are displayed in Fig. 6. As in the case of Tsix, the energy maps contain most of the features displayed by the experimental maps. The comparison between the results of the explicit–extruder model, those of the effective model and the experimental map are displayed in Fig. 7. Also in these cases, the effective model ($r = 0.82$ and $r = 0.76$ for the two regions, respectively) performs similarly, if not better, than the explicit–extruder model ($r = 0.78$ and $r = 0.71$, respectively). The main features of the Hi–C maps are captured by both the explicit and the effective models. While small TADs are captured well, large TADs display in the models a finer structure which is not apparent in the experiment. Interestingly, the explicit and the effective models produce very similar maps (r=0.89 and 0.80, respectively), even in the
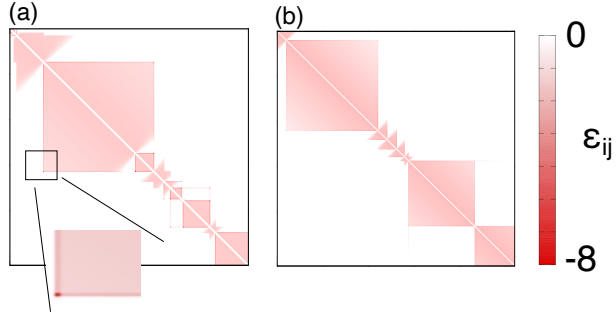
16

FIG. 6. The interaction energies ϵ_{ij} between beads for (a) region chrX:102278307-103570000 and (b) region chrX:103578307-106170000. The inset is the zoom of the squared region.

patterns that are not in the experimental map.

The larger correlation of the effective model with the experimental data is presumably due to simulation statistics: while the explicit-extruder model requires to average the motion of the extruders over multiple simulations (30 in the case shown above), the effective model already contains implicitly this average. The most apparent difference between the two models is that while the explict extruder generates maps whose elements are spatially correlated with their neighbors, the maps obtained with the effective model display abrupt changes between neighboring elements. This is not unexpected, since the effective model, assuming a stationary distribution of the extruder along the chain, neglects correlations between consecutive sites associated with the motion of the extruder on short time scales.

A popular way of summarizing the information contained in contact maps of chromosomes is studying the average contact probability between sites as a function of their distance $|i-j|$ along the chain, that is usually a power law [25].

The three sets of experimental data we studied display scaling coefficients $\beta = 0.71$, $\beta = 0.87$ and 0.77, respectively. The simulations display a central region (around 10 beads, corresponding to 50 kbp) in which the contact probability is a power law with good exponents both in the case of the explicit–extruder (giving 0.74, 0.86 and 0.75, respectively) and the effective model (0.73, 0.86 and 0.78, respectively). In addition, all models display a bend at $|i-j| < 5$, likely due to the coarse–graining of the model and an exponential cut-off due to finite–size effects.

This power–law dependence of the contact probability on the linear distance is not surprising in the light of the effective model, if this describes realistically the effective interactions
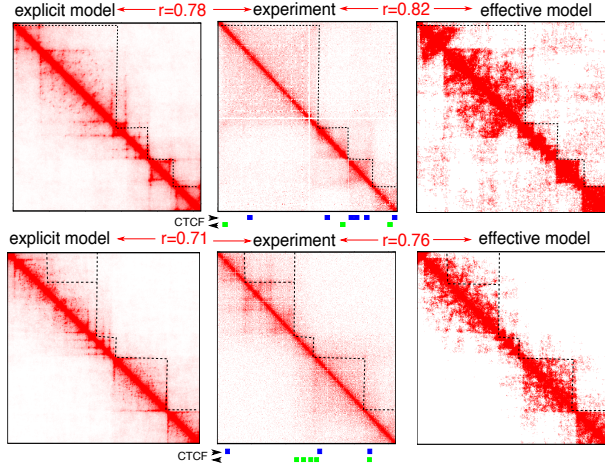
FIG. 7. Comparison between the experimental data (central panels), the results of the explicit–extruder model (left panels) and those of the effective model (right panels). The upper panels are region chrX:102278307-103570000, the lower panel chrX:103578307-106170000. The dashed lines are a guide to the eye.

between the beads of the polymer. In fact, both the interaction energy (cf. Eq. 29) and the polymer looping entropy display such a power–law dependence.

VI. CONCLUSIONS

We developed an effective model for the dynamics of chromosomes based on the assumption that the interactions that stabilize TADs are mediated by extruders running along the polymer consuming energy. The effective model is built in such a way that its equilibrium conformations approximate the conformations visited in the long run by the out–of–equilibrium extrusion mechanism.

We showed that simulations performed with the effective model produce average contact maps that are as similar to the experimental Hi–C map as those from an explicit extrusion mechanism. Even if they allow to detect TADs, the agreement with the experimental data for both kinds of model is still not perfect. This is not surprising because they are based on a minimal amount of information, that is the position of CTCF along the chain. There are indeed models [26–28] that produce contact maps closer to the experimental ones, but at the price of a larger amount of input information, being thus less predictive. Importantly, the present model can be improved by adding information from the experimental maps, such as
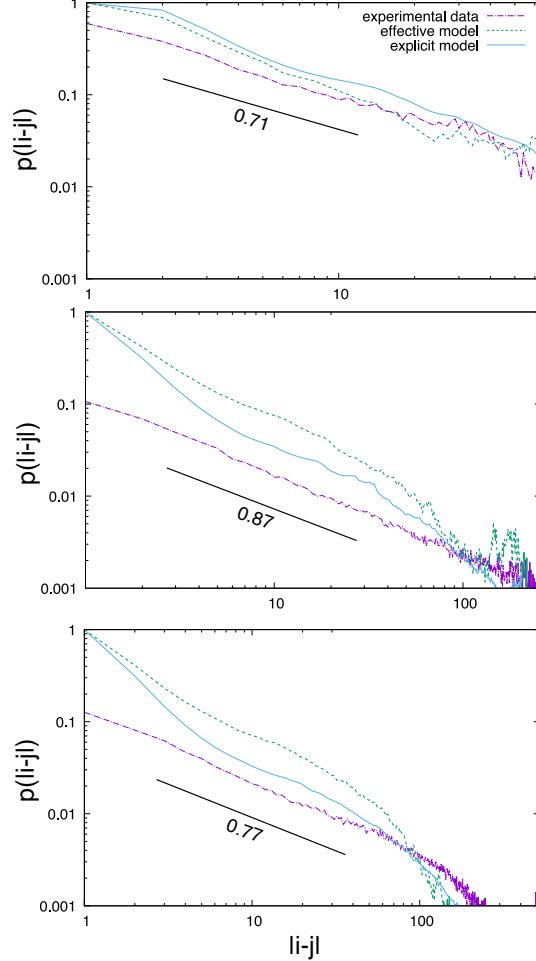
FIG. 8. The contact probability between sites as a function of their distance along the chain, for the Tsix domain (upper panel), for region chrX:102278307-103570000 (middle panel) and region chrX:103578307-106170000 (lower panel).

the presence of compartments on a length scale larger than that of TADs [16].

The effective model in controlled by a free parameter (E_0) because interaction energies are defined but for an additive constant. This parameter cannot be determined by the position of CTCF but has to be tuned manually. Although the details of the simulated maps depend on this choice, the overall partitioning of the chromosome into domains seems quite robust with respect to it.

The effective model is based on averaging out the position of the extruder along the chain, so it is a mean–field approximation. Although this appears to be good enough to reproduce average maps, by definition it cannot account for fluctuations. Thus, the price to be payed to reduce the complexity in the description of the system is the loss of information about

cell–to–cell variability.

Nonetheless, the maps that summarize the interaction energies ϵ_{ij} display the main patterns present in the experimental Hi–C maps, suggesting that polymeric entropy plays a limited role in shaping the architecture of chromosomes, at least at the scale of Mbp.

Another strong approximation that we implemented is that the extruder cannot overcome CTCF sites. This approximation allowed us to obtain an analytical expression for the distribution of the extruder. However, it is known that CTCF is bound to its binding sites only for $\approx 50\%$ of time [29], resulting in an effective permeability of CTCF sites. Explicit simulations of the motion of phantom and of sterically-interacting extruders give similar results, at least in a simple toy system.

An approach analogous to ours was followed in ref. [17], in which a Fokker–Planck equation for the binding probability of the extruder is solved in case of constant velocity, of pure diffusion and of diffusion in an effective potential that reflects the entropic cost of polymer looping. Only in the last case the binding probability displays a power-law scaling, reflecting the dependence of the entropy cost on the linear distance of the loop. However, recent experiments [9, 10] indicate that cohesin uses ATP not only to bind/unbind but also to run on the chromosome. Since ATP hydrolysis rate in cohesin is approximately 2 s^{-1} [9], assuming that hydrolysis provides ≈ 30 kJ/mol, the provided power is approximately 60 kJ/mol per second. On the other hand, cohesin runs at a rate of 2 kbp/s, making a loop of 4 kbp every second. Assuming a persistence length of the order of a kbp, the energy loss associated with the formation of the loop is at most $T \log 4 \approx 3$ kJ/mol, which is one twentieth of the provided energy. Consequently, one does not expect entropy loss to be the main determinant of the distribution of cohesin on the fiber.

The computational gain offered by the effective model is quite consistent, not only because one has not to solve the equations for the extruder, but also because in a single simulation one describes in an effective way multiple trajectories. As an example, the simulation of the Tsix domain with the effective model takes of the order of 0.2 h per simulated minute per cpu core, to be compared with 6 h with the explicit model.

The effective model is useful not only to make simulations more efficient, but also to clarify the physics of chromatin. For example, in the light of the form of the potential developed for the effective model, it is not surprising that the contact probability scales with the linear distance as a power law. In fact, both the looping energy and the associated

20

entropy loss scale as power laws.

Te code to perform the simulations can be downloaded at https://github.com/martanit/LEAD

Appendix A: Numerical parameters of the system

Experiments of fluorescence recovery after photobleaching indicate that the mean residence time of cohesin on the chromatin fiber is 13 minutes, corresponding to a detachment rate $k_{off} = 1.3 \cdot 10^{-3} s^{-1}$ [30]. Total internal reflection microscopy of reconstructed cohesin–chromatin in a flow cell indicate that the stepping rate of cohesin is $k = 10^3$ bp/s [9]. Fluorescence correlation spectroscopy experiments show that approximately $c = 250,000$ copies of cohesin are present in human cells in G1 phase and that 64% of them are bound to chromatin [30]. The same order of magnitude but a smaller number ($c = 109,000$) is obtained for mouse embryonic stem cells [31]. The binding rate k'_{on} of cohesin on chromatin per base can be estimated from k_{off} and from the fraction of bound molecules, that is

$$k'_{on} = k_{off} \frac{V_n}{N} \frac{c_b}{c - c_b}, . \tag{A1}$$

where V_n is the nuclear volume, N is the total number of base pairs, c_b is the number of bound cohesin molecules and c is the total number of cohesin molecules. Using $V_n = 500$ μm^3, $N = 3 \cdot 10^9$, $c = 2.5 \cdot 10^5$ and $c_b = 1.6 \cdot 10^5$ [30] one obtains $k'_{on} = 3.9 \cdot 10^{-13}$ $nm^2/(ps \cdot bp)$.

The (effective) diffusion coefficient of the chromatin fiber measured by live–cell imaging is $D_{ch} = 3 \cdot 10^{-3}$ $\mu m^2/s$ [32]. The diffusion coefficient of cohesin can be estimated by Stoke's law, using a hydrodynamic radius of $R = 8.5$ nm [33] and a viscosity for the nucleoplasm of $\eta = 1.5$ cP [34]. One obtains for cohesin $D_{co} = 18$ $\mu m^2/s$, which is five orders of magnitude larger than that of chromatin, justifying the well–mixed hypothesis.

The time scale τ_{ext} associated with extrusion on the TAD length scale (i.e., $L \sim 10^6$ bp extruded by $n_{ext} \sim 30$ cohesin molecules) is $\tau_{ext} \sim L/(k\, n_{ext}) \sim 30$ s. The time scale associated with the motion of the chain is $\tau \sim L^2/D_{ch} \sim 300$s (using for typical TADs $L \sim 100$ nm [35]).

The Hi–C maps we used as reference have a resolution of $5 \cdot 10^3$ bp, an thus we used this as elementary unit of the model. From the density obtained from ref. [26], this corresponds to $a = 67$ nm. We used this quantity as elementary length scale for the model. The friction constant of the polymer can be obtained from Einstein's equation $\gamma = D_{ch}/k_B T$ and, in

21

terms of the length scale a, at room temperature is $\gamma = 4 \cdot 10^{12} \ kJ\,ps/mol\,a^2$. The stepping rate of cohesin is $k = 2 \cdot 10^{-13} \ a/ps$. The loading rate used for simulations of N monomers is $k_{\text{on}} = Nk'_{\text{on}}$.

Assuming the mass density typical of biomolecules, $1 \ g/cm^3$, the mass of a monomer is of the order of $10^{-22} \ Kg$.

Appendix B: Solution of the recursive equation

Let's define $b_{n,m} = a_{n+1,m+1}$. Eq. (B1) can be written as

$$b_{n,m} = b_{n-1,m} + b_{n,m-1}, \tag{B1}$$

that can be solved iteratively starting from $b_{n,0} = 2^{n+1} - 1$ and $b_{0,m} = 1$.

Let's define the bivariate generating function

$$f(x,y) = \sum_{n,m=0}^{\infty} b_{n,m} x^n y^m \tag{B2}$$

Separating the terms $m, n = 0$ one obtains

$$
\begin{aligned}
f(x,y) &= \sum_{n=0}(2^{n+1} - 1)x^n + \sum_{m>0} y^n + \sum_{n,m>0} b_{n,m} x^n y^m \\
&= \frac{1}{1-x} + \frac{y}{1-y} + \sum_{n,m>0} b_{n,m} x^n y^m.
\end{aligned} \tag{B3}
$$

Substituting Eq. (B1) and renaming the indexes,

$$
\begin{aligned}
f(x,y) &= \frac{1}{(1-x)(1-2x)} + \\
&\quad + \frac{y}{1-y} + \sum_{n,m>0} [b_{n-1,m} + b_{n,m-1}] x^n y^m = \\
&= \frac{1}{(1-x)(1-2x)} + \frac{y}{1-y} + x \sum_{n=0,m>0} b_{n,m} x^n y^m - \\
&\quad + y \sum_{n>0,m=0} b_{n,m} x^n y^m = \\
&= \frac{1}{(1-x)(1-2x)} + \frac{y}{1-y} + xf(x,y) - \\
&\quad - \frac{x}{(1-2x)(1-x)} + yf(x,y) - \frac{y}{1-y} = \\
&= \frac{1}{(1-2x)} + xf(x,y) + yf(x,y)
\end{aligned} \tag{B4}
$$

22

Thus,

$$f(x,y) = \frac{1}{(1-x-y)(1-2x)},$$ (B5)

whose series expansion is

$$b_{n,m} = 2^{n+m+1} - \frac{\Gamma(n+m+2)_2F_1(1,n+m+2,n+2,1/2)}{2\Gamma(n+2)\Gamma(m+1)}$$ (B6)

where $_2F_1$ is the Gaussian hypergeometric function. This expression can be simplified to

$$b_{n,m} = \binom{n+m+1}{m+1} {}_2F_1(1,-n.2+m,-1)$$ (B7)

and thus

$$a_{n,m} = \binom{n+m-1}{m} {}_2F_1(1,1-n.1+m,-1)$$ (B8)

[1] J. R. Dixon, S. Selvaraj, F. Yue, A. Kim, Y. Li, Y. Shen, M. Hu, J. S. Liu, and B. Ren, Nature **485**, 376 (2012).

[2] Y. Zhan, L. Mariani, I. Barozzi, E. G. Schulz, N. Blüthgen, M. Stadler, G. Tiana, and L. Giorgetti, Genome Res. **27**, 479 (2017).

[3] G. Fudenberg, M. Imakaev, C. Lu, A. Goloborodko, N. Abdennur, and L. A. Mirny, Cell Rep. **15**, 2038 (2016).

[4] G. Fudenberg, N. Abdennur, M. Imakaev, A. Goloborodko, and L. A. Mirny, Cold Spring Harb. Symp. Quant. Biol. **82**, 45 (2017).

[5] S. K. Ghosh and D. Jost, Brief. Funct. Genomics **19**, 119 (2020).

[6] R. J. Spencer, B. C. del Rosario, S. F. Pinter, D. Lessing, R. I. Sadreyev, and J. T. Lee, Genetics **189**, 441 (2011).

[7] G. Wutz, C. Várnai, K. Nagasaka, D. A. Cisneros, R. R. Stocsits, W. 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, EMBO J. **36**, 3573 (2017).

[8] S. S. P. Rao, S.-C. Huang, B. Glenn St Hilaire, J. M. Engreitz, E. M. Perez, K.-R. Kieffer-Kwon, A. L. Sanborn, S. E. Johnstone, G. D. Bascom, I. D. Bochkov, X. Huang, M. S. Shamim, J. Shin, D. Turner, Z. Ye, A. D. Omer, J. T. Robinson, T. Schlick, B. E. Bernstein, R. Casellas, E. S. Lander, and E. L. Aiden, Cell **171**, 305 (2017).

[9] I. F. Davidson, B. Bauer, D. Goetz, W. Tang, G. Wutz, and J.-M. Peters, Science. **366**, 1338 (2019).

[10] Y. Kim, Z. Shi, H. Zhang, I. J. Finkelstein, and H. Yu, Science. **366**, 1345 (2019).

[11] A. L. Sanborn, S. S. P. Rao, S.-C. Huang, N. C. Durand, M. H. Huntley, A. I. Jewett, I. D. Bochkov, D. Chinnappan, A. Cutkosky, J. Li, K. P. Geeting, A. Gnirke, A. Melnikov, D. McKenna, E. K. Stamenova, E. S. Lander, and E. L. Aiden, Proc. Natl. Acad. Sci. U. S. A. **112**, E6456 (2015).

[12] M. Yin, J. Wang, M. Wang, X. Li, M. Zhang, Q. Wu, and Y. Wang, Cell Res. **27**, 1365 (2017).

[13] Y. Li, J. H. I. Haarhuis, Á. Sedeño Cacciatore, R. Oldenkamp, M. S. van Ruiten, L. Willems, H. Teunissen, K. W. Muir, E. de Wit, B. D. Rowland, and D. Panne, Nature **578**, 472 (2020).

[14] S. S. P. Rao, M. H. Huntley, N. C. Durand, E. K. Stamenova, I. D. Bochkov, J. T. Robinson, A. L. Sanborn, I. Machol, A. D. Omer, E. S. Lander, and E. L. Aiden, Cell **159**, 1665 (2014).

[15] G. Tiana and L. Giorgetti, Curr. Opin. Struct. Biol. **49**, 11 (2018).

[16] J. Nuebler, G. Fudenberg, M. Imakaev, N. Abdennur, and L. A. Mirny, Proc. Natl. Acad. Sci. U. S. A. **115**, E6697 (2018).

[17] C. A. Brackley, J. Johnson, D. Michieletto, A. N. Morozov, M. Nicodemi, P. R. Cook, and D. Marenduzzo, Phys. Rev. Lett. **119**, 138101 (2017).

[18] C. A. Brackley, J. Johnson, D. Michieletto, A. N. Morozov, M. Nicodemi, P. R. Cook, and D. Marenduzzo, Nucleus **9**, 95 (2018).

[19] J. F. Marko, P. De Los Rios, A. Barducci, and S. Gruber, Nucleic Acids Res. **79**, 51905 (2019).

[20] A. Y. Grosberg and A. Khokhlov, *Statistical Physics of Macromolecules* (AIP Press, 1994).

[21] W. Schwarzer, N. Abdennur, A. Goloborodko, A. Pekowska, G. Fudenberg, Y. Loe-Mie, N. A. Fonseca, W. Huber, C. Haering, L. Mirny, and F. Spitz, Nature **152**, 1270 (2017).

[22] G. Bussi and M. Parrinello, Phys. Rev. E, Stat. nonlinear, soft matter Phys. **75**, 56707 (2007).

[23] E. P. Nora, A. Goloborodko, A. L. Valton, J. H. Gibcus, A. Uebersohn, N. Abdennur, J. Dekker, L. A. Mirny, and B. G. Bruneau, Cell **169**, 930 (2017).

[24] J. Redolfi, Y. Zhan, C. Valdes-Quezada, M. Kryzhanovska, I. Guerreiro, V. Iesmantavicius, T. Pollex, R. S. Grand, E. Mulugeta, J. Kind, G. Tiana, S. A. Smallwood, W. de Laat, and L. Giorgetti, Nat. Struct. Mol. Biol. **26**, 471 (2019).

[25] E. Lieberman-Aiden, N. L. van Berkum, L. Williams, M. Imakaev, T. Ragoczy, A. Telling, I. Amit, B. R. Lajoie, P. J. Sabo, M. O. Dorschner, R. Sandstrom, B. Bernstein, M. A. Bender,

M. Groudine, A. Gnirke, J. Stamatoyannopoulos, L. A. Mirny, E. S. Lander, and J. Dekker, Science. **326**, 289 (2009).

[26] L. Giorgetti, R. Galupa, E. P. Nora, T. Piolot, F. Lam, J. Dekker, G. Tiana, and E. Heard, Cell **157**, 950 (2014).

[27] D. Marenduzzo, Genome Biol. **17**, 1 (2016).

[28] D. Jost, C. Vaillant, and P. Meister, Curr. Opin. Cell Biol. **44**, 20 (2017).

[29] A. S. Hansen, I. Pustova, C. Cattoglio, R. Tjian, and X. Darzacq, Elife **6**, 10.7554/eLife.25776 (2017).

[30] J. Holzmann, A. Z. Politi, K. Nagasaka, M. Hantsche-Grininger, N. Walther, B. Koch, J. Fuchs, G. Dürnberger, W. Tang, R. Ladurner, R. R. Stocsits, G. A. Busslinger, B. Novák, K. Mechtler, I. F. Davidson, J. Ellenberg, and J.-M. Peters, Elife **8**, 10.7554/eLife.46269 (2019).

[31] C. Cattoglio, I. Pustova, N. Walther, J. J. Ho, M. Hantsche-Grininger, C. J. Inouye, M. J. Hossain, G. M. Dailey, J. Ellenberg, X. Darzacq, R. Tjian, and A. S. Hansen, Elife **8**, e40164 (2019).

[32] G. Tiana, A. Amitai, T. Pollex, T. Piolot, D. Holcman, E. Heard, and L. Giorgetti, Biophys. J. **110**, 1234 (2016).

[33] S. Weitzer, C. Lehane, and F. Uhlmann, Curr. Biol. **13**, 1930 (2003).

[34] L. Liang, X. Wang, D. Xing, T. Chen, and W. R. Chen, J. Biomed. Opt. **14**, 024013 (2009).

[35] A. N. Boettiger, B. Bintu, J. R. Moffitt, S. Wang, B. J. Beliveau, G. Fudenberg, M. Imakaev, L. A. Mirny, C.-t. Wu, and X. Zhuang, Nature **529**, 1 (2016).