Dynamic Random Loops Can Explain the Appearance of Topologically Associating Domain in Chromosome Capture Experiments

1 Introduction

The spatio-temporal organization of the chromatin plays an essential role in the regulation of sub-cellular activity such as gene expression[1].

3d structure of the Igh locus was suggested by [4] after tagging several position along the loci.

2 Experimental data and Methods

2.1 The experimental data

We used the experimental 5C data generated by Nora et al.[5] for the chromosome contact frequencies of the X chromosome in a 4.5 Gb region encompassing the X inactivation center. In our work we focused on a subset of the data, including a 94,082 bp region, termed TAD D and E (see [5]).

The maps generated by the 5C experiments describes the contact between genomic loci of variable sizes over millions of nuclei. We followed data coarse-graining procedure, similarly to the one described in [3], to map segment encounter frequencies to that of evenly spaced, equal size beads. A bead size of 3000 bp was chosen according to the mean size of restriction segments resulted by HindIII enzyme digestion, used in the process of the 5C experiments (see [3] Supplementary Material). This choice of bead size resulted in coarse-grained polymer of 307 beads.

The coarse-grained pair-wise bead encounter frequencies includes 14,509 data points and was used to calculate the bead encounter probability as a function of the distance in beads units. Encounter frequencies for equidistant beads were averaged to give the 'one sided' bead encounter frequency. The bead encounter probability was derived by dividing the bead encounter frequency by the total number of encounters of that bead.

We fitted the experimental bead encounter probability with a function of the form

$$p(d) = \alpha d^{-\beta} \tag{1}$$

with p the encounter probability, $\alpha = \frac{1}{\sum_{j=1}^{k} j^{-\beta}}$, d is the distance in bead units, and β is a parameter to be determined by the fitting procedure.

2.2 The polymer model

To explore the different polymer conformations that can explain the appearance of the TADs, we chose to use the Rouse chain. The Rouse chain describes the dynamics of a linear polymer as a collection of massless beads connected by harmonic springs and driven by the thermal forces of diffusion. The system of stochastic differential equations describing the time progression of a chain of N beads is given in the 3-dimensional case by

$$\frac{dR}{dt} = -\frac{3D}{b^2} KR + \sqrt{2D}\frac{dW}{dt} \tag{2}$$

where, $R(t) = [R_1(t), R_2(t), .., R_N(t)]^T$ describes the 3D coordinates of N beads at time t, D is the diffusion constant, b is the standard-deviation of the distance between adjacent beads of the chain, W is an independent $N \times 3$ Brownian motion with mean 0 and variance 1 in each component, and K is the Kirchhoff bead connectivity matrix, which reflects different chain connectivities.

We have constructed our polymer model to have L loops of random sizes. To form each loops, we have randomly chosen 2 non-neighboring beads and altered the connectivity in the Kirchhoff matrix, with the condition that no bead can participate in the formation of more than one loop.

2.3 Simulations

Throughout simulations, for each fixed number of loops, the chosen beads to connect varied randomly. Such a choice was made to refer to the heterogene-

ity if the spatial organization inside TAD between cells, even in the same cell phase [5].

Simulations were always carried out until the chain's relaxation time, in which point any two beads were determined to have encountered if their distance at the end of the simulation satisfied $|R_j - R_k| < \epsilon < b, \quad (j \neq k)$. The chain's relaxation time is given by the slowest mode of the linear chain

$$\tau = \frac{b^2}{12D \sin(\frac{\pi}{2N})}$$

for which the number of simulation steps performed is $\frac{\tau}{\Delta t}$. The time step, Δt was set so to prevent simulation 'blow-ups' by demanding that the quotient of the norms of beads position at two subsequent time steps would be smaller that unity, which resulted in $\Delta t < \frac{b^2}{12D}$.

For each tested polymer connectivity we constructed the bead encounter frequencies histogram and derived the bead encounter probability from it. The bead encounter probability was then fitted similarly to the fitting in eq. 1.

For a linear Rouse chain with nearest neighbor interactions, the expected value of β is 1.5 [2]. We interpret $\beta < 1.5$ as long range interaction resulting from non nearest-neighbor bead connections. Since adding non-neighboring connections to the linear chain can only increase the long range encounter probability, we have focused on interpreting the fitted values in the range $\beta < 1.5$.

3 Results

3.1 Analysis of the experimental data

To evaluate the mean encounter probability in the experimental data, we have calculated the β value for the 3 cases of TAD D, TAD E and TAD D+E for each of the beads in those genomic region. This provides us with a basis for comparison of the results of simulations with the experimental data and to the inference on the spatial organization of the chromosome.

The calculation of β for each bead in the case of TAD D + E resulted in a pattern which was correlated with the significant long range interactions (Figure 1 lower panel), represented by the peaks the encounter probability graph (Figure 1, upper panel)

The mean β value for the case of TAD D and E was 0.729 (Figure 1, upper panel), well below the expected 1.5 of a linear Rouse chain, which suggested the presence of long range specific bead interactions.

4 Discussion

5 Figure

References

[1] Thomas Cremer and Christoph Cremer. Chromosome territories, nuclear architecture and gene regulation in mammalian cells. *Nature reviews genetics*, 2(4):292–301, 2001.

[2] M. Doi and S.F. Edwards. *The theory of polymer dynamics*. International series of monographs on physics. Clarendon Press, 1986.

[3] Luca Giorgetti, Rafael Galupa, Elphège P. Nora, Tristan Piolot, France Lam, Job Dekker, Guido Tiana, and Edith Heard. Predictive polymer modeling reveals coupled fluctuations in chromosome conformation and transcription. *Cell*, 157(4):950–963, May 2014.

[4] Suchit Jhunjhunwala, Menno C van Zelm, Mandy M Peak, Steve Cutchin, Roy Riblet, Jacques JM van Dongen, Frank G Grosveld, Tobias A Knoch, and Cornelis Murre. The 3d structure of the immunoglobulin heavy-chain locus: implications for long-range genomic interactions. *Cell*, 133(2):265–279, 2008.

[5] Elphège P Nora, Bryan R Lajoie, Edda G Schulz, Luca Giorgetti, Ikuhiro Okamoto, Nicolas Servant, Tristan Piolot, Nynke L van Berkum, Johannes Meisig, John Sedat, Joost Gribnau, Emmanuel Barillot, Nils Blüthgen, Job Dekker, and Edith Heard. Spatial partitioning of the regulatory landscape of the x-inactivation centre. *Nature*, 485:381–5, 2012 May 17 2012.

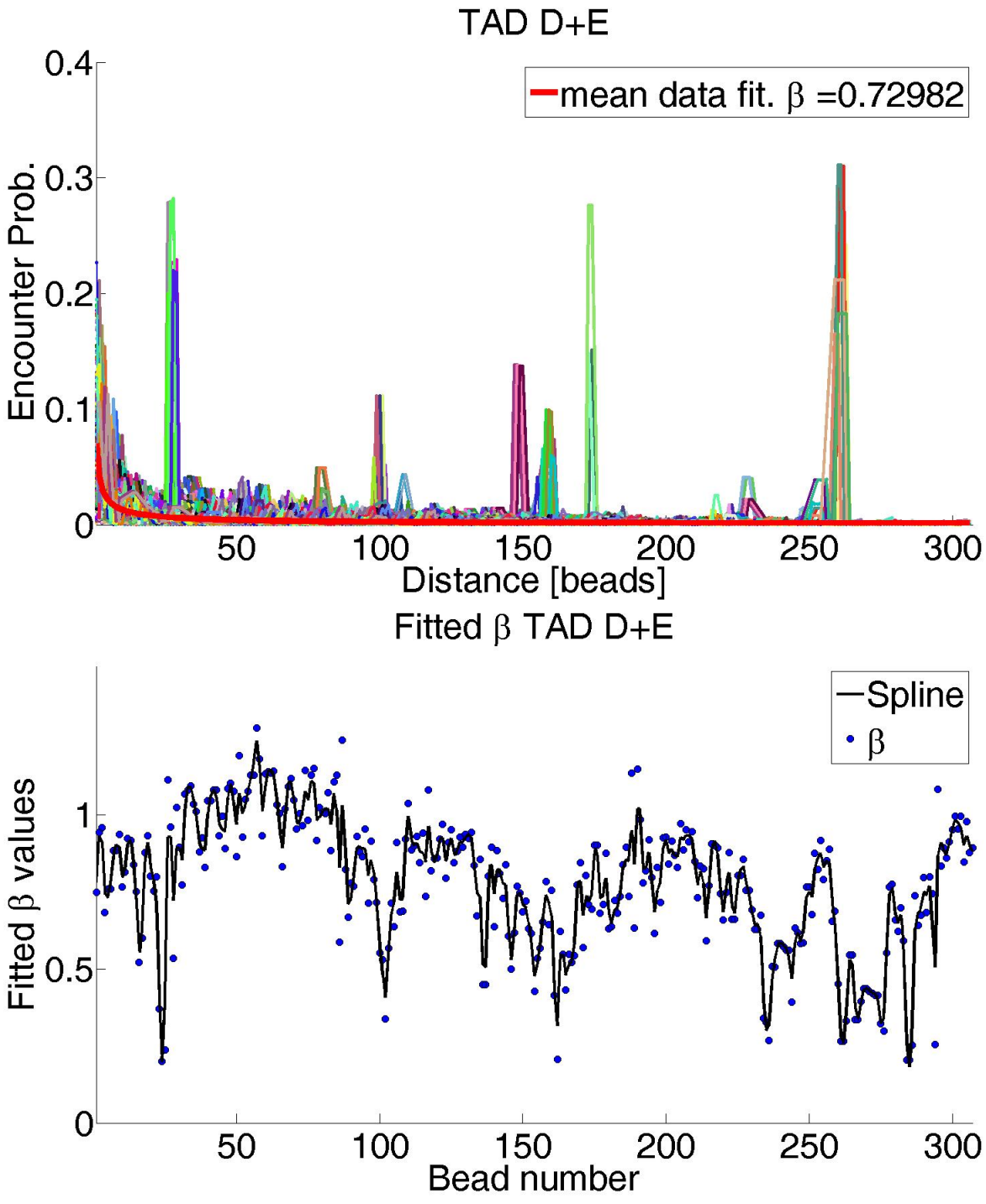
TAD D+E
mean data fit. β =0.72982
Encounter Prob.
0.4
0.3
0.2
0.1
0
50
100
150
200
250
300
Distance [beads]
Fitted β TAD D+E
Spline
β
Fitted β values
1
0.5
0
50
100
150
200
250
300
Bead number

Fig. 1: