

Implications of the polymer brush theory for the budding yeast nucleus.

Anton Goloborodko, Geoffrey Fudenberg.

May 17, 2012

Abstract

The budding yeast nucleus is a model system for studies of chromosomal organization. The current biological understanding of yeast chromosomal organization is that chromosomes are tethered at their centromeres to a relatively small patch on the inner surface of the almost spherical nuclear envelope. In this work, we study the implications of such organization from the point of view of the polymer theory. We model the nucleus with the modified DiMarzio & Rubin lattice model and compare the results with the existing experimental data.

Introduction.

The conformation of DNA in a nucleus controls the spacial proximity of genes and has implications on its function. However, until recently there were very experimental methods to study the folding of chromosomes. Recently, a method named Hi-C was developed to assess the relative contact frequency between each pair of genomic regions, or loci. The method quickly gained popularity and several organisms has already been studied. However, the theory of contact frequencies in polymeric systems has not been developed thoroughly, probably, because of the absence of the demand from the experimentalists.

The budding yeast nucleus represents a convenient biological system for the studies of chromatin organization [1]. It has long been known that its interphase chromosomes assume so-called “Rabl conformations” [2], i.e. they are tethered at their centromeres to a comparatively small region on the nuclear envelope. Such type of chromosomal organization is not limited to yeast, but also occurs in plants and worms. The implications of such DNA conformation on the contact probability have never been studied.

In this paper we present a simple lattice model of chromatin in the budding yeast nucleus. We model the nucleus as a rectangular prism with the chromosomes grafted on the inner surface of one of the faces. This model falls under the definition of a polymer brush, which theory was developed in the 90s [3].

The standard approach to brush modelling is to use the self-consistent field approximation (SCF), where the interactions between otherwise phantom chains are accounted for by a mean field potential of mutual exclusion.

The key assumption behind the model is that in an ensemble of nuclei, the chromosomes are grafted at random points. Under this assumption, the average spatial distribution of each individual locus is independent of its coordinates along the grafting plane and depends only on the distance to it. Therefore, the dimensionality of a problem gets reduced from three to one. Moreover, the average relative frequency of contacts between two genomic loci in the ensemble will depend only on the correlation of their one-point spacial probability distributions. It is important to note that this assumption may not be applied to calculate the contact frequency of two loci located on the same chromosome, since it does not account for the two-point correlations.

Lattice model.

The nucleus is modeled as a rectangular prism with a base area $A=1.69\mu m^2$ and the height $H=1.2\mu m$. The chromosomes are split at their centromeres into a pair of chromosomal arms and each arm was modeled as a freely-jointed phantom chain with the Kuhn length of $b=120$ nm per 3000 base pairs and the width $w = 20$ nm. Totally, 32 polymer chains of the lengths varying from 33 to 360 b are considered in the model. The simplifying assumption made in the model is that the spatial distributions of individual monomers vary only with a distance z to the grafting surface, i.e. the distribution of monomers within (x,y) plane is considered homogeneous. In this assumption, the system is described by a one-dimensional lattice with $H/b = 10$ nodes, where one node represented a plane parallel to the grafting surface. Therefore, the set of vectors $P_j^{n,(i)}$ containing the probability of observing the i -th monomer of the n -th chain in the j -th plane fully characterize the spatial distributions of genomic loci. The first monomer of each chromosomal arm is always localized at the wall.

The probability distribution of monomers among the nodes of a lattice is calculated via the modified DiMarzio & Rubin lattice model [4, 5, 6]. In this model, the probability distribution of $(i+1)$ -th monomer is derived from the probability distribution of i -th monomer using the transfer matrix:

$$\mathbf{P}^{n,(i+1)} = \mathbf{T}^{n,(i)} \mathbf{P}^{n,(i)}$$

The transfer matrix \mathbf{T}_{ij} contains the probability of a step from j -th to the i -th level and is defined as:

$$\mathbf{T}^{n,(i)} = \left\| \begin{array}{cccc} \frac{4Z(L_n-i,1)}{4Z(L_n-i,1)+Z(L_n-i,2)} & \frac{Z(L_n-i,2)}{Z(L_n-i,1)+4Z(L_n-i,2)+Z(L_n-i,3)} & 0 & \dots \\ \frac{Z(L_n-i,2)}{4Z(L_n-i,1)+Z(L_n-i,2)} & \frac{4Z(L_n-i,2)}{4Z(L_n-i,2)+Z(L_n-i,3)} & \frac{Z(L_n-i,2)}{Z(L_n-i,2)+4Z(L_n-i,3)+Z(L_n-i,4)} & \dots \\ 0 & \frac{Z(L_n-i,2)+Z(L_n-i,3)}{Z(L_n-i,1)+4Z(L_n-i,2)+Z(L_n-i,3)} & \frac{4Z(L_n-i,3)}{Z(L_n-i,2)+4Z(L_n-i,3)+Z(L_n-i,4)} & \dots \\ 0 & 0 & \frac{Z(L_n-i,4)}{Z(L_n-i,2)+4Z(L_n-i,3)+Z(L_n-i,4)} & \dots \\ \dots & \dots & \dots & \ddots \end{array} \right\|$$

where L_n is the length of n-th chain and $Z(L, j)$ if the partition function of a chain of length L with one end fixed in j-th layer of the lattice.

The tethering of the polymers is reflected in the spacial distribution of their first monomers:

$$\mathbf{P}^{n,(1)} = \left\| \begin{array}{c} 1 \\ 0 \\ \dots \\ 0 \end{array} \right\|$$

The partition function $Z(L_n - i, j)$ used in the probability transfer matrix \mathbf{T} is calculated via another modification of the DiMarzio & Rubin approach:

$$\mathbf{Z}^{(i+1)} = \mathbf{W} \mathbf{Z}^{(i)}$$

$$\mathbf{W} = \left\| \begin{array}{cccc} 4 \exp(-\phi(1)) & \exp(-\phi(1)) & 0 & \dots \\ \exp(-\phi(2)) & 4 \exp(-\phi(2)) & \exp(-\phi(2)) & \dots \\ 0 & \exp(-\phi(3)) & 4 \exp(-\phi(3)) & \dots \\ 0 & 0 & \exp(-\phi(4)) & \dots \\ \dots & \dots & \dots & \ddots \end{array} \right\|$$

$$Z(N, j) = \mathbf{1}^T \mathbf{W}^{L-1} \mathbf{Z}^{(1)}(\mathbf{j})$$

$$\mathbf{Z}^{(1)}(\mathbf{j}) = \left\| \begin{array}{c} 0 \\ \dots \\ \exp(-\phi(j)) \\ \dots \\ 0 \end{array} \right\|$$

, where the only non-zero element is at j-th row.

The self-consistent field is defined as:

$$\phi(j) = \frac{a}{A} \sum_{n=1}^N \sum_{i=1}^{L_n} P_j^{n,(i)}$$

, where $a = w^2$ is the cross sectional area of the polymer, A is the area of a lattice plane and the first sum goes over different chains.

The distribution vectors are calculated iteratively by substituting the resulting $P_j^{n,(i)}$ to the equation for the self-consistent field.

Finally, the relative probability of a contact between two monomers in a chain is defined as a dot product of their probability density profiles:

$$P_c(n, i, n', i') = \sum_j P_j^{n,(i)} P_j^{n',(i')}$$

Results.

The relative occupancy of the lattice layers is shown on Fig.1. The comparison with the calculations for non-interacting chains suggests that the resulting distribution of monomer density is affected by a combination of two effects. The dominant one is the effective repulsion of polymers from the walls. In the limit of infinitely long chains, the probability density of a phantom chain in a slit of height H has the form of $\sin^2(\frac{\pi z}{H})$ [7]. In our case, this distribution is skewed towards the grafting plane, since most of the chains are not long enough to reach the opposite wall. As seen in the figure, the inter-chain repulsion counterbalances the confinement and leads to a more uniform distribution of monomer density in the bulk of the nucleus.

Fig.2 shows the difference between spacial distribution of monomers of the longest polymer of the system, the right arm of chromosome 4 and the relatively short left arm of chromosome 1. Again, the repulsion from the walls and grafting seem to be singularly responsible for the distribution of the monomers that are relatively close to the centromere, both for the short and long chromosomes. However, the loci that are further away from the centromere get repelled to the far region of the nucleus. This effect may play role in the formation of the nucleolus by the repetitive rDNA regions of the chromosome 12.

The availability of the Hi-C data allows us to compare the predicted contact frequency with the experiment (Fig.3). The calculated contact maps qualitatively agree with the measurements. The experiment confirms the prediction of low contact frequency between the centromeric loci with the rest of the chromosomes due to wall repulsion. The chain-chain repulsion is predicted to further increase the isolation of the telomeres from the centromeres which is also qualitatively confirmed by the experimental data.

These very first results of the application of our model agree with the major features of the yeast genomic contact maps. While not designed to replace the Brownian dynamics simulations, the proposed method has several complementary advantages. The lattice approach is semi-analytical and promotes the separation of effects with relatively easy interpretation of the results. The high speed of computation allows the researcher to quickly test hypotheses and explore the effect of changing the free parameters of the model. The future directions of the research include the quantitative comparison with the experimental data, as well as the modification of the method to account for the spherical geometry of a nucleus. Another major improvement would be to allow the model to predict

the interaction frequency between loci within the same chromosome.

We thank Prof. Leonid Mirny and our labmates for thoughtful discussions.

References

- [1] A. Taddei, H. Schober, and S. M. Gasser, Cold Spring Harbor perspectives in biology **2**, a000612 (2010).
- [2] C. Rabl, *Über Zelltheilung*. (W. Engelmann, Leipzig, 1887).
- [3] S. T. Milner, Science (New York, N.Y.) **251**, 905 (1991).
- [4] E. A. DiMarzio and R. Rubin, The Journal of Chemical Physics **55**, 4318 (1971).
- [5] J. M. H. M. Scheutjens and G. J. Fleer, The Journal of Physical Chemistry **83**, 1619 (1979).
- [6] J. M. H. M. Scheutjens and G. J. Fleer, The Journal of Physical Chemistry **84**, 178 (1980).
- [7] A. I. Grosberg and A. R. Khokhlov, *Statistical physics of macromolecules* (AIP Press, New York, 1994).

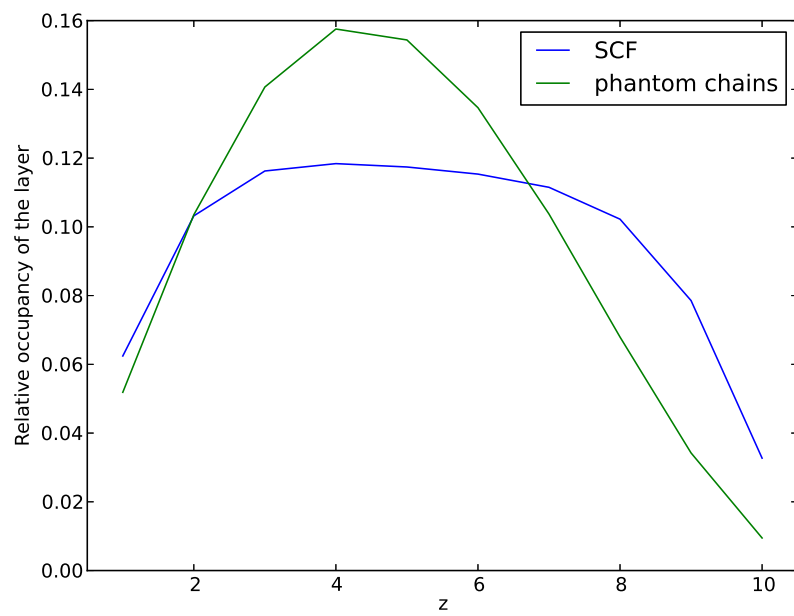


Figure 1: The monomer density profile

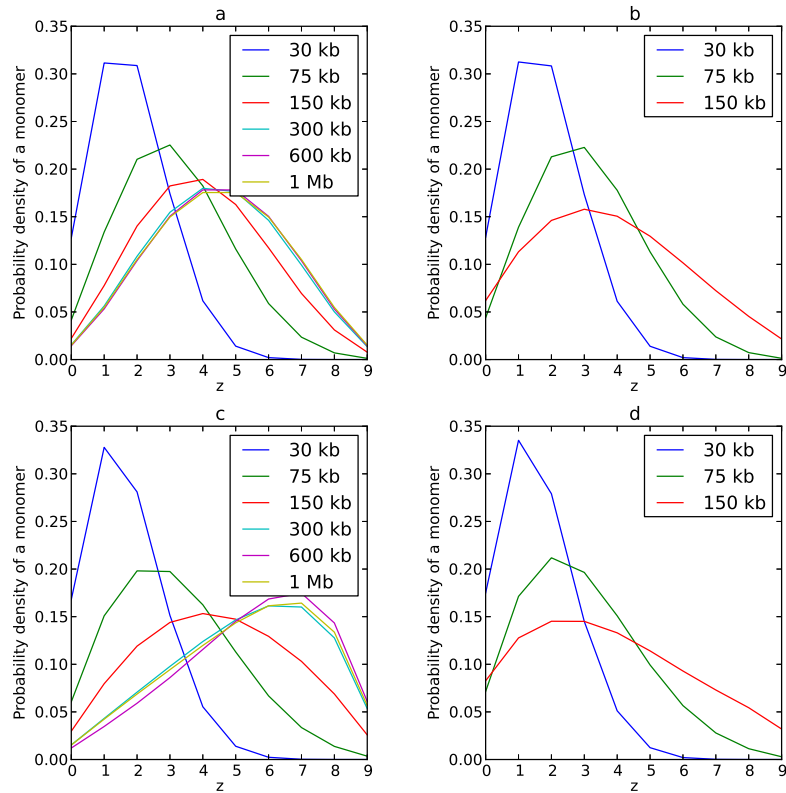


Figure 2: The spacial distributions of individual genomic loci. (a) and (c) shows the distributions calculated for the loci of the right arm of chromosome 4 in the non-interacting and SCF case, respectively. (b) and (d) shows the distributions calculated for the loci of the left arm of chromosome 1 in the non-interacting and SCF case, respectively. Different curves correspond to the spatial distributions of loci at various genomic distances from the centromere.

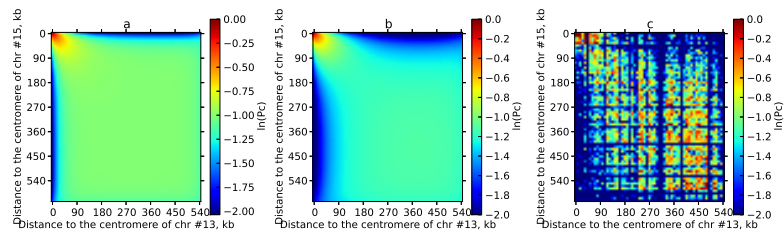


Figure 3: The relative frequency of contacts between the genomic loci of the left arms of chromosome 13 and 15. (a) - non-interactive chains calculations, (b) - self-consistent field calculations, (c) - experimental data.