# Supporting Information


## Mukhopadhyay et al. 10.1073/pnas.1103845108


### SI Text

**Further Background on Biology of Looping Models or "Whose Chromatin Is It, Anyway?".** Regulatory action at distance is a phenomenon found in many organisms, mostly metazoans. The most famed locus for investigation of long-distance enhancer-promoter interaction is the beta-globin locus for mammals (see article by Palstra, et al. in ref. 6, pages 107–142). However, such distal enhancer activity exists in invertebrates as well. Prompted by the phenomenon of position effect variegation [(PEV), see the article by Girton and Johansen in ref. 6, pages 1–44] *Drosophila melanogaster* became the organism where much of the genetic investigations of long-range regulatory influence happened. Drosophila is the organism where one of the coauthors of this article (P.D.S.) discovered insulators (see ref. 7). Vertebrate insulators (CTCF binding sites) were found later. Enhancer and insulators have also been studied in vitro, a major endeavor in the lab of another coauthor of this article (V.M.S.) (5). Without exaggerating, one might say that most biologists expect that there is a set of common mechanisms underlying long-distance enhancer action in metazoans and the same is true of the inhibitory influence of insulators. Hence, perhaps, an effort to understand theoretical aspects of the general problem makes sense in this particular context.

**Simulation Details.** The model for our simulation of chromatin polymer needed to be computationally minimal but nevertheless be able to capture the essential features of intermediate-to-long-range properties of chromatin. Over several trials, we settled on the following model, all of whose parameters could be learned from a more microscopic modeling approach adopted by many other groups (including our collaborators Wilson group).

A bead and spring model captures the polymer aspect of the chromatin. The beads are spherical and the springs connecting them are of zero energy length equal to bead diameter, bending rigidity of a few $k_BT$ and stretching rigidity approximately double the bending rigidity. We explored bending rigidities in the range of 2.5–5$k_BT$ and stretching in the range 5–10$k_BT$, and all the data presented in the paper is for 2.5$k_BT$ bending and 5$k_BT$ stretching rigidity. The stretching however, was modeled by a FENE (finite extensible nonlinear elastic) potential such that length change within 20% of the spring length was allowed. In order ensure noncrossing of the polymer, we introduced phantom beads of the same diameter as the real beads and placed in the middle of the vector connecting each neighboring real bead centers. The phantom beads only interacts with real beads that are not its immediate neighbor and all other phantom beads through a $1/r^{12}$ potential, with an exceptionally large energy scale so that the repulsion energy cost for overlap is prohibitive. We explored both 200 beads and 100 beads ring polymer configuration, all data presented in the paper is for 200 beads. All real beads also experience the same nonoverlap repulsion.

Modeling the electrostatic interaction between nucleosomes required several trial and error and input from the more microscopic molecular level modeling by Wilma Olson group and the existing literature on modeling nucleosomes (see refs. 1–5). We based our model on two observations from these studies:

1. The nucleosome interactions assuming rigid histone tails, but considering the detailed charge distribution on the histone cores and tails (modeled as cylinders) depend on their relative orientations in space. This dependence is owing to the very complex nature of the charge distributions.
2. The interaction is moderately attractive at short ranges–energy no more than a few $k_BT$ and a range of the order of nucleosome size. Coupled with the first observation, a typical nucleosome pair interacts intermittently. The flexibility of the histone tails should only strengthen this feature.

With these observations, we modeled our nucleosome to have an internal binary random variable. Only two *available* nucleosomes (i.e., both with internal variable one) can interact attractively. This internal variable is updated every hundreds of Monte Carlo (MC) steps for each bead and the energy cost for availability is controlled by Boltzmann factor with energy of few $k_BT$. We call this energy *availability*. This move is done by generating a random update time from an exponential distribution. Two *available* bead, when in close proximity to each other, interact by an attractive potential that is modeled as a Gaussian potential well (depth of few $k_BT$s) with a cutoff of a 1.5 $*$ bead size. We call this potential *attraction*. We have studied a number of attraction and availability pairs. All combinations from the two lists (availability: 1.00$k_BT$, 1.25$k_BT$, 1.50$k_BT$; attraction: 2.00$k_BT$, 2.25$k_BT$, 2.50$k_BT$) were extensively simulated (see below).

In order to save computation time, the configuration of the polymer was checked at a MC time scale slower than the update of the internal variable of nucleosomes. The simulation was done by MC methods. Details are as follows:

A set of random initial configurations were generated. This procedure involved equilibration step from a circular configuration by fifty million attempted MC moves for *each* bead. The attempted move in space of a randomly chosen bead for each MC step is 5% of its size. The success rate of a move is on the average around 60–65%. We ensured from polymer statistics considerations (bead to bead distance, pair correlation function, energy, etc.) this initial step resulted in a random equilibrium configuration.

The data is gather from at least 25 runs for each combination of parameters (availability and attraction), where each run is fifty million attempted MC moves for each bead. The average run took of the order of seven-eight hours on a computer cluster (C++ programming language). Given that we have explored nine combinations of parameters extensively, the entire dataset is a result of around sixteen hundred hours of computational time for the ring polymer set-up and the same time for the pinched polymer set-up. Typically, for each run we store the configuration every twenty thousand MC steps (for 200 beads) and check the configuration to attempt to establish or break attractive bonds. The updates of the internal variables are done asynchronously for all beads at roughly one tenth that time scale, as explained above. The dataset has been analyzed for quality check by computing the pair correlation functions, participation of beads in bonds, bead to bead distance, and fluctuations in radius of gyration etc. to ensure that the polymer is not stuck in a configuration (e.g., a fully collapsed state or a long-lived compact state). Our parameter choice and model construction was to avoid exploring collapsed polymer state–the attractive interaction is low and entropy still dominates the free energy.

For the pinched configuration, all possible four "promoter" beads at equal distances to an "enhancer" bead were explored in the data analysis that established insulation properties. However, this more careful approach meant reducing our total data to a less sizable one given by such specific picks. Over all insulation is, of course, far less noisy.

**Comparing Length Scales Between the Simulation and Relevant Biological Systems.** In our system, the persistence length is about two links, based on the peak in looping probability. Thus, our polymer

length is at least 50 times the persistence length. Making an estimate of persistence length in chromatin is tricky because it could depend upon nucleosome density, secondary structure (like 10 nm or 30 nm fiber), and other conditions. See ref. 1, for a fit to yeast data and ref. 2, for a more theoretical discussion of nucleosome density affects persistence length. In terms of genomic distance, this number could vary from a few Kb to 30 Kb. Hence 100 monomers could represent 100 Kb to 1.5 Mb, depending upon which locus one is applying this model to. On the other hand, this is the right range for distal enhancer action in mammals (beta-globin LCR 50 Kb, ZRS-Shh 1 Mb, see articles by Palstra et al and the one by Kleinjan and Lettice in ref. 6). Most known distal enhancers in Drosophila (see the review in ref. 3) operate from less than 100 Kb away. Typical CTCF demarcated insulator domains are order 100–200 Kb (4).


1. Bystricky K, Heun P, Gehlen L, Langowski J, Gasser SM (2004) Long-range compaction and flexibility of interphase chromatin in budding yeast analyzed by high-resolution imaging techniques. *Proc Natl Acad Sci USA* 101:16495–16500.
2. Diesinger PM, Heerman DW (2009) Depletion effects massively change chromatin properties and influence genome folding. *Biophys J* 97:2146–2153.
3. Dorsett D (1999) Distant liaisons: long-range enhancer-promoter interactions in Drosophila. *Curr Opin Genet Dev* 9:505–514.
4. Kim TH, et al. (2007) Analysis of the vertebrate insulator protein CTCF-binding sites in the human genome. *Cell* 128:1231–1245.
5. Polikanov YS, Studitsky VM (2009) Analysis of distant communication on defined chromatin templates in vitro. *Methods Mol Biol* 43:563–576.
6. van Heyningen V, Hill RE eds (2008) *Long-range control of gene expression* (*Academic Press*, London).
7. Kellum R, Schedl P (1991) A position-effect assay for boundaries of higher order chromosomal domains. *Cell* 64:941–950.