

Comprehensive Mapping of Long-Range Interactions Reveals Folding Principles of the Human Genome

Authors: Erez Lieberman-Aiden, Nynke L. van Berkum, Louise Williams, Maxim Imakaev, Tobias Ragozy, Agnes Telling, Ido Amit, Bryan R. Lajoie, Peter J. Sabo, Michael O. Dorschner, Richard Sandstrom, Bradley Bernstein, M. A. Bender, Mark Groudine, Andreas Gnirke, John Stamatoyannopoulos, Leonid A. Mirny, Eric S. Lander and Job Dekker

Science **326**:289-293, 2009.

Recommended with a commentary by Yitzhak Rabin, Bar-Ilan University.

It is only on rare occasion that theoretical predictions based on simple physical considerations with limited experimental input are recognized as important (or even relevant) by the cell biology community, Watson and Crick's double helix model of DNA being perhaps the most famous counterexample. It therefore came as a surprise when a group of biologists, mathematicians and physicists reported the results of a comprehensive study of the folding principles of entire human genome (46 DNA molecules) in the nucleus of a cell[1] and concluded that they are consistent with the so called fractal globule concept, introduced many years ago under the name "crumpled globule"[2, 3].

By cross-linking neighboring (in space, not along the sequence) DNA strands, cutting them from the rest of the polymer, ligating them together and analyzing their sequences, Lieberman-Aiden et al were able to partition the chromatin into 1 Mb (megabase) segments and construct a contact matrix m_{ij} whose entries represent the frequency of contacts between segments i and j . Consistent with previous observations of chromosome territories[4], they found that different DNA molecules are segregated in the nucleus and that certain chromosome pairs tend to be near one another in space. Analysis of the contact matrix showed that each chromosome can be decomposed into two sets of segments (compartments) A and B such that contacts within each compartment occur with high probability and contacts between different compartments are suppressed. For a given genomic distance (i.e., separation along the contour of DNA), contacts between segments in set B occur with higher probability than in set A, consistent with the observation that while set A is strongly correlated with the presence of transcribed genes and therefore with expanded (accessible) chromatin, DNA sequences in set B appear to be correlated with the more densely packed, non-transcribing heterochromatin.

In order to obtain (statistical) information about the structure of chromatin folds within each compartment the authors measured the contact probability $I(s)$ as a function of genomic distance s . They found that in the range $500\text{ kb} < s < 7\text{ Mb}$ which corresponds to the typical range of sizes of expanded and condensed chromatin domains, this probability scales as $1/s$. This simple result allows one to rule out the possibility that the chromatin is folded into

an equilibrium globule state familiar from polymer physics, which forms (upon long equilibration) when a dilute polymer solution is brought to poor solvent conditions (e.g., by changing temperature). Notice that a polymer in a good solvent is a self-similar object with fractal dimension $d_f = 5/3$ since the mean spatial distance $r(s)$ between 2 points separated by contour distance s scales as $s^{3/5}$, for any s smaller or equal to the chain length L [5]. Similarly, an ideal polymer (a Gaussian random walk) is a fractal object with $d_f = 2$. Such self-similar (scale-independent) behavior is not observed for the equilibrium globule. While on small scales ($s < L^{2/3}$) the equilibrium globule behaves as an ideal solution of disconnected polymers with $r \sim s^{1/2}$, on larger scales r becomes independent of s and approaches the radius of the globule $R \sim L^{1/3}$ [6]. It was suggested that the equilibrium globule is not the only possible collapsed state of polymers and a new class of fractal collapsed objects for which $r \sim s^{1/3}$ on all length scales (and thus $d_f = 3$) was proposed[2, 3]. Such objects were named "crumpled globules" because, unlike the heavily self-knotted equilibrium globule, the crumpled globule represents a set of folded polymer conformations that do not contain entanglements and it is therefore metastable with respect to the entropically-favored, knotted equilibrium state whose formation involves time scales that are not accessible to DNA within a typical cell cycle, even if the machinery of topological enzymes is taken into account[7]. They argued that the absence of entanglements makes the crumpled globule a particularly attractive candidate for the dense packaging of chromatin in the nucleus of a cell[3], since massive entanglements would adversely interfere with the processes of transcription and replication. The crumpled globule was also implicated in simulations of concentrated solutions of uncatenated and unknotted rings, based on the idea that as different crumples of the same chain remain segregated so are also the different rings – and for the same topological reason[8].

Coming back to the results of Lieberman-Aiden et al., one can now check whether the observed contact probability obeys the expected scaling of the equilibrium or of the fractal globule. Note that since in both cases this probability is inversely proportional to the volume $r^3(s)$, it should scale as $1/s^{3/2}$ for the equilibrium globule but only as $1/s$ for the fractal globule, and the observation of the latter scaling provides strong support for the crumpled globule, self-similar model of chromatin. Using Monte-Carlo simulations and Alexander polynomial analysis of the degree of knotting, Lieberman-Aiden et al demonstrated the lack of entanglements in fractal globules. They also showed (this point was missed in ref. [3]) that unlike other crumpled and unknotted curves (e.g., Peano curve), a fractal globule consists of heavily interpenetrating "crumples" such that the number of contacts between two neighboring crumples of contour length s is proportional to their volume (s) rather than to their surface area ($s^{2/3}$). This surprising result suggests that even though different segments of the crumpled globule are not entangled (in the sense of knots), they are strongly interdigitated.

References

- [1] E. Lieberman-Aiden et al., *Science* **326**, 289 (2009).
- [2] A. Yu. Grosberg, S. K. Nechaev and E. I. Shakhnovich, *J. Phys. France* **49**, 2095 (1988).
- [3] A. Yu. Grosberg, Y. Rabin, S. Havlin and A. Neer, *Europhys. Lett.* **23**, 373 (1993).
- [4] T. Cremer and C. Cremer, *Nat. Rev. Genet.* **2**, 292 (2001).
- [5] P. -G. deGennes, *Scaling Concepts in Polymer Physics* (Cornell Univ. Press, Ithaca, NY, 1979).
- [6] I. M. Lifshitz, A. Yu. Grosberg and A. R. Khokhlov, *Rev. Mod. Phys.* **50**, 683 (1978).
- [7] J.-L. Sikorav and G. Jannink, *Biophys. J.* **66**, 827 (1994).
- [8] T. Vettorel, A. Yu. Grosberg and K. Kremer, *Phys. Biol.* **6**, 025013 (2009).