Unraveling the mechanisms of chromatin fibril packaging

Alexey A. Gavrilov, Yuri Y. Shevelyov, Sergey V. Ulianova, Ekaterina E. Khrameeva, Pavel Kos, Alexander Chertovich, and Sergey V. Razin

Department of Molecular Genetics of Cell, Institute of Molecular Genetics RAS, Moscow, Russia; Department of Molecular Biology, Lomonosov Moscow State University, Moscow, Russia; Institute for Information Transmission Problems (The Kharkevich Institute) RAS, Moscow, Russia; Physics Department, Lomonosov Moscow State University, Moscow, Russia

ABSTRACT
Recent data indicate that eukaryotic chromosomes are organized into Topologically Associating Domains (TADs); however, the mechanisms underlying TAD formation remain obscure. Based on the results of Hi-C analysis performed on 4 Drosophila melanogaster cell lines, we have proposed that specific properties of nucleosomes in active and repressed chromatin play a key role in the formation of TADs. Our computer simulations showed that the ability of “inactive” nucleosomes to stick to each other and the lack of such ability in “active” nucleosomes is sufficient for spatial segregation of these types of chromatin, which is revealed in the Hi-C analysis as TAD/inter-TAD partitioning. However, some Drosophila and mammalian TADs contain both active and inactive chromatin, a fact that does not fit this model. Herein, we present additional arguments for the model by postulating that transcriptionally active chromatin is extruded on the surface of a TAD, and discuss the possible impact of this organization on the enhancer-promoter communication and on the segregation of TADs.

In eukaryotic cells, a DNA chain is folded thousands of times to fit the small volume of a cell nucleus. Although in textbooks this folding is commonly represented as a continuity of hierarchical structures (10-nm nucleosomal chain, 30-nm chromatin fiber, loops of the 30-nm chromatin fiber, etc.) the current results suggest that only irregularly folded 10-nm fibers are present in the nuclei of living cells (for a review, see refs.1,2). Recent advances in the study of genome topology using the high resolution chromosome conformation capture procedure (Hi-C) have made it possible to demonstrate that chromosomes of eukaryotes are partitioned into so-called Topologically Associating Domains (TADs). These domains are defined by the high contact probabilities of regions located within them but with decreased contact probabilities of regions located in different TADs. In mammals, TADs appear to be composed of smaller looped and ordinary domains. TADs are separated by short border regions or longer inter-TADs. Partitioning of chromosomes into TADs is typical for different organisms, from unicellular ones to humans. However, with an increase in genome complexity, the tendency for chromosomes to partition into TADs becomes more pronounced. For example, TADs are not observed in the gene dense and smaller genomes of Plasmodium and Arabidopsis, whereas they are present in the larger genomes of mammals. The profiles of genome partitioning into TADs have been reported to be conserved between different types of cells and even between cells of related species, such as human and mouse. The borders of mammalian TADs and inter-TAD regions were found to be enriched in CTCF insulator protein binding and the promoters of housekeeping genes. TADs were also identified in Drosophila possessing a genome of a similar length but less gene dense than that of Arabidopsis. The borders of Drosophila TADs were also found to contain active genes and binding sites for insulator and architectural proteins such as CTCF and cohesin, suggesting the role of these proteins in the demarcation of TADs. However, the mechanisms underlying TAD formation remain obscure.
Recently, we performed Hi-C analysis on *Drosophila* cell lines of different origins. Contrary to the previous studies, we did not observe a strong enrichment of TAD borders/inter-TADs with CTCF deposition sites, and instead another insulator protein Su(Hw) was preferentially present within the TADs. However, *Drosophila* inter-TADs harbor active chromatin and constitutively transcribed (house-keeping) genes. The tissue-specific genes reside preferentially in TADs and their transcription appears to correlate with a partial de-compaction of the TADs. Interestingly, we found that the inter-TADs possess all the features previously attributed to decompacted interbands of polytene chromosomes and their positions mostly correspond to the previously reported positions of the interbands. This observation is further supported by the recently published Hi-C analysis of *Drosophila* polytene chromosomes.

We propose that TADs are self-organized condensed chromatin domains depleted in active chromatin marks that are separated by regions of active chromatin. The latter is poorly folded into compact structures because of the high level of histone acetylation. According to our model, the interaction of nucleosomes plays a key role in the chromatin partitioning into TADs. Nucleosomes as a whole are negatively charged repulsing units. However, histones are positively charged, and positively charged histone tails present on a nucleosome may interact with negatively charged sites on another nucleosome (such as an H2A/H2B acidic patch or DNA chain), thus bridging 2 nucleosomes together and forming a compact tertiary structure of a nucleosomal array. Acetylation of the histones present in active chromatin regions diminishes the positive charge of their tails, thus preventing the nucleosomes from clumping together with each other. A similar effect may be caused by histone ubiquitylation. Moreover, the presence of both acetylation and ubiquitylation gives a synergistic effect on chromatin fiber decompaction. Noteworthily, we observed an enrichment of both these epigenetic marks inherent to active chromatin at TAD borders and in inter-TADs (Fig. 1 and ref. 11). It may be also relevant that TAD borders and inter-TADs are depleted in linker histone H1, which is known to participate in chromatin compaction, and in the core histones. This may reflect the increased histone turnover in the actively transcribed regions and the presence of nucleosome depleted regions in the genome regulatory elements. A disturbance in the regularity in nucleosome positioning and a decreased compensation for the negative charge of DNA caused by partial dissociation of the histones could contribute to the decompaction of chromatin within inter-TADs.

Taking into consideration the ability of nucleosomes of repressed chromatin to establish short-lived interactions with each other and the absence/reduction of this ability in the nucleosomes of active chromatin, we performed a computer simulation of the folding of a chromatin fiber, represented by beads on the string, partitioned in blocks of “inactive” and “active” nucleosomes. The polymer collapsed in a series of condensed blobs (analogous to TADs), consisting mostly of “inactive” nucleosomes separated by relatively stretched spacers (analogous to inter-TADs) of “active” nucleosomes. In this simulation, each “inactive” nucleosome was allowed to establish only one interaction with another nucleosome at a time (saturated interactions). Collectively, the results of our study strongly suggest that the ability of “inactive” nucleosomal arrays for self-aggregation and the lack of such ability in “active” arrays are sufficient for the 3D-spatial segregation of these types of chromatin.

In principle, “active” arrays tend to be extruded from compact structures formed by longer “inactive” arrays. The final configuration of the polymer will strongly depend on the distribution of “active” and “inactive” arrays along a linear polymer chain. In our original modeling, the blocks of 500 “inactive” nucleosomes were interspersed by blocks of 50 “active” nucleosomes. Simulation of this polymer folding resulted in a characteristic pattern of compact TADs separated by unfolded inter-TADs. However, the insertion of short arrays of “active” nucleosomes into a long region composed of “inactive” (self-interacting) monomers results in the formation of a similar compact TAD covered by extruded “active” arrays (Fig. 2A, middle globule). This may give a key to understanding the mechanism of the formation of TADs assembled from “inactive” nucleosomes with various insertions of active chromatin. Here, small active regions are looped out from the TAD globule, whereas long active regions form linkers between different TADs. The location of active genes at the TAD surface may facilitate transcription and RNA release to the interchromatin channels. This model is in agreement with the microscopy studies suggesting that interphase chromosomes are organized in a series of ~1 Mb globular domains immersed in the interchromatin...
compartment,\textsuperscript{18} and that transcription occurs in the perichromatin region at the surface of the condensed chromatin domains.\textsuperscript{19,20} FISH staining of individual TADs supports the supposition that TADs correspond to globular chromatin domains.\textsuperscript{4,21} A layer of “active” (non-interacting) nucleosomes on the surface of an “inactive” TAD core would insulate this core from occasional interaction/intermingling with other TADs.

The proposed model suggests that active enhancers that are transcribed, bear active chromatin marks and colocalize with nucleosome depleted regions,\textsuperscript{22-25} should also be extruded to the TAD surface. This may facilitate enhancer-promoter communication as the search for a target promoter would occur in the chromatin shell surrounding TAD, where the mobility should be much faster.

One of the important observations of our polymer modeling is that the structure of TADs in individual cells can differ significantly. In individual simulations, we were able to detect the fusion of TADs, their assembling into large aggregates (mega-TADs), the destruction and splitting of TADs into smaller blobs (sub-TADs), as well as the formation of long-range contacts.\textsuperscript{11} Strikingly, averaging the contact maps for individual simulations facilitates enhancer-promoter communication as the search for a target promoter would occur in the chromatin shell surrounding TAD, where the mobility should be much faster.

One of the important observations of our polymer modeling is that the structure of TADs in individual cells can differ significantly. In individual simulations, we were able to detect the fusion of TADs, their assembling into large aggregates (mega-TADs), the destruction and splitting of TADs into smaller blobs (sub-TADs), as well as the formation of long-range contacts.\textsuperscript{11} Strikingly, averaging the contact maps for individual simulations facilitates enhancer-promoter communication as the search for a target promoter would occur in the chromatin shell surrounding TAD, where the mobility should be much faster.

One of the important observations of our polymer modeling is that the structure of TADs in individual cells can differ significantly. In individual simulations, we were able to detect the fusion of TADs, their assembling into large aggregates (mega-TADs), the destruction and splitting of TADs into smaller blobs (sub-TADs), as well as the formation of long-range contacts.\textsuperscript{11} Strikingly, averaging the contact maps for individual simulations facilitates enhancer-promoter communication as the search for a target promoter would occur in the chromatin shell surrounding TAD, where the mobility should be much faster.

**Figure 1.** Distribution of the ubi-H2B chromatin mark around TAD borders in 3 Drosophila cell lines. Box-plots show the density of ubiquitylated histone H2B in 20-kb bins located at the same position relative to a TAD boundary, averaged over all TADs. The TAD boundaries were aligned so that the inter-TAD bins were placed to the left of the boundary bin (“0 kb” bin) and TAD bins were placed to the right of the boundary bin. The data on the ubi-H2B mark were downloaded from the modENCODE database and processed as described in ref.\textsuperscript{11}.

**Figure 2.** Computer simulation of the folding of a linear polymer composed of blocks of “inactive” and “active” nucleosomes. (A) One of the predicted spatial configurations of the polymer arranged as follows: 1000A-100B-(100A-10B)\times 10-100B-1000A, where A is the “inactive” nucleosomes (green) and B is the “active” nucleosomes (black). The simulation was done using the dissipative particle dynamics (DPD) algorithm, as described in ref.\textsuperscript{11}, with the only difference being that the conformation was evaluated during the time of $1 \times 10^6$ DPD steps. (B) Spatial proximity map (distance heat map) of the polymer conformation presented in (A) (note, the scale is in beads). A scheme of the model polymer is presented below the map; blocks of “inactive” and “active” nucleosomes are shown in green and black, respectively.
erased the information about this diversity and produced a regular pattern of TADs, each represented by a single block of “inactive” nucleosomes. These results are in line with the published single-cell Hi-C data, which demonstrated cell-to-cell variability in the structure of individual chromosomes. However, resolution of the single-cell Hi-C analysis performed so far is not sufficient to track the individual TADs.

The loop extrusion model was recently proposed to direct chromatin compaction in mammals. The borders of extruded loops are thought to be established by CTCF/cohesin binding. The convergent CTCF-binding motifs were indeed found at many mammalian TAD borders. Moreover, the causative role of CTCF-binding sites in the establishment of TAD borders in mammals was demonstrated using CRISPR/Cas9-based deletions and inversions of these sites. We note that in all these cases, the CTCF-mediated loops at TADs borders gave bright contact spots in the tops of triangles, reflecting self-interacting domains in the heat maps. Such bright spots were not detected in the tops of TADs in Drosophila cells. Therefore, currently there is no evidence supporting the existence of such a mechanism in Drosophila.

In conclusion, we suggest that the basic principle of chromatin self-partitioning into TADs is based on the different properties of “active” and “inactive” types of nucleosomal arrays. This mechanism appears to work almost in “pure form” in Drosophila. In mammals, the mechanism of TAD formation is probably more complex. Loop extrusion and demarcation of the loop borders by CTCF/cohesin binding play an essential role in the generation of TAD patterns. One of the current models proposes that TADs are generated due to the presence of TAD-specific sets of communicator proteins which bridge their target sites within a TAD. However, currently there is no evidence for the existence of TAD-specific sets of communicator proteins. In our opinion, communicator proteins may rather ensure functionally-dependent shaping of TADs generated by aggregation of nucleosomal arrays. The structure of TADs can also be modulated by DNA supercoiling. In a recent study, it was shown that transcriptionally inactive chromatin is topologically overwound and forms compact large scale chromatin structures, whereas transcriptionally active chromatin is underwound and characterized by a more decompacted structure. Including supercoiling into models of topological domains allows a much better recapitulation of the experimentally determined Hi-C data. Enzymes mediating DNA supercoiling (RNA polymerase, DNA topoisomerases) can work in cooperation with “loop extrusion machines” to create a characteristic TAD structure. Nevertheless, it is likely that interactions of the nucleosomes underlie all the proposed mechanisms of TADs formation. Indeed, the compaction of extruded chromatin loops will still be directed by simple physical laws and will proceed via interaction of the nucleosomal particles.

**Abbreviations**
- DPD dissatisfying particle dynamics
- TAD topologically associating domain

**Disclosure of potential conflicts of interest**
No potential conflicts of interest were disclosed.

**Funding**
This work was supported by the Russian Science Foundation (project #14-14-01088).

**References**

1. Maeshima K, Ide S, Hibino sK, Sasai M. Liquid-like behavior of chromatin. Curr Opin Genet Dev 2016; 37:36-45; PMID:26826680; http://dx.doi.org/10.1016/j. cged.2015.11.006
2. Razin SV, Gavrilov AA. Chromatin without the 30-nm fiber: constrained disorder instead of hierarchical folding. Epigenetics 2014; 9:653-7; PMID:24561903; http://dx.doi.org/10.4161/epi.28297
3. Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, Amit I, Lajoie BR, Sabo PJ, Dorschner MO, et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science 2009; 326:289-93; PMID:19815776; http://dx.doi.org/10.1126/science.1181369
4. Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, Hu M, Liu JS, Ren B. Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature 2012; 485:376-80; PMID:22495300; http://dx.doi.org/10.1038/nature11082
5. Nora EP, Lajoie BR, Schulz EG, Gioretti L, Okamoto I, Servant N, Pirolt T, van Berkum NL, Meisig J, Sedat J, et al. Spatial partitioning of the regulatory landscape of the X-inactivation centre. Nature 2012; 485:381-5; PMID:22495304; http://dx.doi.org/10.1038/nature11049
6. Hou C, Li L, Qin ZS, Corces VG. Gene density, transcription, and insulators contribute to the partition of the Drosophila genome into physical domains. Mol Cell 2012; 48:471-84; PMID:23041285; http://dx.doi.org/10.1016/j.molcel.2012.08.031
[7] Sexton T, Yaffe E, Kenigsberg E, Bantignies F, Leblanc B, Hoichman M, Barral N, Tanay A, Cavalli G. Three-dimensional folding and functional organization principles of the Drosophila genome. Cell 2012; 148:458-72; PMID:22265598; http://dx.doi.org/10.1016/j.cell.2012.01.010

[8] Rao SS, Huntly MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, Sanborn AL, Machol I, Omer AD, Lander ES, et al. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. Cell 2014; 159:1665-80; PMID:25497547; http://dx.doi.org/10.1016/j.cell.2014.11.021

[9] Ay F, Bunnik EM, Varoquaux N, Bol SM, Prudhomme J, Vert JP, Noble WS, Le Roch KG. Three-dimensional modeling of the P. falciparum genome during the erythrocytic cycle reveals a strong connection between genome architecture and gene expression. Genome Res 2014; 24:974-88; PMID:24671853; http://dx.doi.org/10.1101/gr.169417.113

[10] Feng S, Cokus SJ, Schubert V, Zhai J, Pellegrini M, Jacobsen SE. Genome-wide Hi-C analyses in wild-type and mutants reveal high-resolution chromatin interactions in Arabidopsis. Mol Cell 2014; 55:694-707; PMID:25132175; http://dx.doi.org/10.1016/j.molcel.2014.10.026

[11] Uljanov SV, Khrameeva EE, Gavrillov AA, Flyamer IM, Kos P, Mikhailova EA, Penin AA, Logacheva MD, Imaeak MV, Chertovich A, et al. Active chromatin and transcription play a key role in chromosome partitioning into topologically associating domains. Genome Res 2016; 26:70-84; PMID:26518482; http://dx.doi.org/10.1101/gr.196006.115

[12] Zhimulev IF, Zykova TY, Goncharov FP, Khoroshko VA, Demakova OV, Semeshin VF, Pokholkova GV, Boldyrev LV, Demidova DS, Babenko VN, et al. Genetic organization of interphase chromosome bands and interbands in Drosophila melanogaster. PLoS One 2014; 9:e101631; PMID:25072930; http://dx.doi.org/10.1371/journal.pone.0101631

[13] Eagen KP, Hartl TA, Kornberg RD. Stable chromosome condensation revealed by chromosome conformation capture. Cell 2015; 163:934-46; PMID:26544940; http://dx.doi.org/10.1016/j.cell.2015.10.021

[14] Barbieri M, Chotalia M, Fraser J, Lavitas LM, Dostie J, Lander ES, et al. Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. Proc Natl Acad Sci U S A 2015; 112:E6456-7; PMID:26504220; http://dx.doi.org/10.1073/pnas.1517972112

[15] Thurman RE, Rynes E, Humbert R, Vierstra J, Maurano MT, Haugen E, Sheffield NC, Stergachis AB, Wang H, Vernot B, et al. The accessible chromatin landscape of the human genome. Nature 2012; 489:75-82; PMID:22955617; http://dx.doi.org/10.1038/nature11232

[16] Melamed P, Yosefzon Y, Rudnizky S, Pnueli L. Transcriptional enhancers: Transcription, function and flexibility. Transcription 2016; 7:26-31; PMID:26934309; http://dx.doi.org/10.1080/1541264.2015.1128517

[17] Li W, Lam MT, Notani D. Enhancer RNAs. Cell Cycle 2014; 13:3151-2; PMID:25485487; http://dx.doi.org/10.4161/cellcycle.27950

[18] Cremer T, Cremer M. Chromosome territories. Cold Spring Harbor Perspect Biol 2010; 2:a003889; PMID:20300217; http://dx.doi.org/10.1101/cshperspect.a003889

[19] Niedojadlo J, Perret-Vivancos C, Kalland KH, Cmarko D, Cremer T, van Driel R, Fakan S. Transcribed DNA is preferentially located in the perichromatin region of mammalian cell nuclei. Exp Cell Res 2011; 317:433-44; PMID:21056558; http://dx.doi.org/10.1016/j.yexcr.2010.10.026

[20] Fabbre PJ, Benke A, Joye E, Nguyen Huynh TH, Manley S, Douboe D. Nanoscale spatial organization of the HoxD gene cluster in distinct transcriptional states. Proc Natl Acad Sci U S A 2015; 112:13964-9; PMID:26504220; http://dx.doi.org/10.1073/pnas.1517972112

[21] Fabre PJ, Benke A, Joye E, Nguyen Huynh TH, Manley S, Douboe D. Nanoscale spatial organization of the HoxD gene cluster in distinct transcriptional states. Proc Natl Acad Sci U S A 2015; 112:13964-9; PMID:26504220; http://dx.doi.org/10.1073/pnas.1517972112

[22] Mousavi K, Zare H, Koulmis M, Sartorelli V. The emerging roles of eRNAs in transcriptional regulatory networks. RNA Biol 2014; 11:106-10; PMID:24525859; http://dx.doi.org/10.4161/15384101.2014.962860

[23] Navarro-Torres F, Pinto L, Luo M, Maro R, Passer PM, Barry PS, et al. Single-cell Hi-C reveals cell-to-cell variability in chromosome structure. Nature 2013; 502:59-64; PMID:22955617; http://dx.doi.org/10.1038/nature11232

[24] Pnueli L. Transcriptional enhancers: Transcription, function and flexibility. Transcription 2016; 7:26-31; PMID:26934309; http://dx.doi.org/10.1080/1541264.2015.1128517

[25] Li W, Lam MT, Notani D. Enhancer RNAs. Cell Cycle 2014; 13:3151-2; PMID:25485487; http://dx.doi.org/10.4161/cellcycle.27950

[26] Fudenberg G, Imakaev M, Lu C, Golobordoko A, Abdennur N, Mirny LA. Formation of chromosomal domains by loop extrusion. Cell Rep 2016; 15:2038-49; PMID:27210764; http://dx.doi.org/10.1016/j.celrep.2016.04.085

[27] Sanborn AL, Roberts SC, Huang SC, Durand NC, Huntley MH, Jewett AI, Bochkov ID, Chinnappan D, Cutkosky A, Li J, et al. Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. Proc Natl Acad Sci U S A 2015; 112:E6456-65; PMID:26499245; http://dx.doi.org/10.1073/pnas.1518552112

[28] Barbieri M, Chotalia M, Fraser J, Lavitas LM, Dostie J, Pombo A, Nicodemi M. Complexity of chromatin folding is captured by the strings and binders switch model. Proc Natl Acad Sci U S A 2012; 109:1673-8; PMID:22988072; http://dx.doi.org/10.1073/pnas.1204799109
[30] Fraser J, Ferrai C, Chiariello AM, Schueler M, Rito T, Laudanno G, Barbieri M, Moore BL, Kraemer DC, Aitken S, et al. Hierarchical folding and reorganization of chromosomes are linked to transcriptional changes in cellular differentiation. Mol Syst Biol 2015; 11:852; PMID:26700852; http://dx.doi.org/10.15252/msb.20156492

[31] Naughton C, Avlonitis N, Corless S, Prendergast JG, Mati IK, Eijk PP, Cockroft SL, Bradley M, Ylstra B, Gilbert N. Transcription forms and remodels supercoiling domains unfolding large-scale chromatin structures. Nat Struct Mol Biol 2013; 20:387-95; PMID:23416946; http://dx.doi.org/10.1038/nsmb.2509

[32] Benedetti F, Dorier J, Burnier Y, Stasiak A. Models that include supercoiling of topological domains reproduce several known features of interphase chromosomes. Nucleic Acids Res 2014; 42:2848-55; PMID:24366878; http://dx.doi.org/10.1093/nar/gkt1353