# Highly heterogeneous long-range interactions explain chromosome compartmentalization

Eidelman Y.A.[1], Salnikov I.V.[1], Slanina S.V.[1], Andreev S.G.[1,2]*
[1] Institute of Biochemical Physics, Russian Academy of Sciences. Kosygin str. 4, 119334 Moscow, Russia
[2] National Research Nuclear University MEPhI, Kashirskoye shosse 31, 115409 Moscow, Russia
*andreev_sg@mail.ru

## Abstract

The explanation for the compartmentalization of the chromosome contact maps observed by Hi-C method is still an unsolved mystery. The most natural and generally accepted explanation is that blocks of frequent/rare contacts on the entire map are associated with the existence of blocks of homogeneous elements along the chromosome that interact in 3D so that elements of the same type frequently of different types rarely. We study the polymer model of the chromosome, in which there are no blocks of homogeneous elements and homogeneous interactions and all the interaction energies are different. We demonstrate that such a heteropolymer model is able to describe chromosomal maps with high accuracy. The differences from compartment polymer models of chromosomes, which actually reflect the mechanism of microphase separation in polymers, are discussed further.

## Introduction

The study of chromosome structure compartmentalization is important. According to some data, compartmentalization correlates with gene expression, while according to other it does not. It changes in the course of differentiation and malignant transformation. In [1], the compartmentalization of contact maps obtained by the Hi-C method was explained by the physically natural and simplest assumption that homogeneous contact blocks are explained by the presence of blocks of homogeneous elements on the chromosomes themselves. The number of element types proposed was two [1] or six [2]. Is there any other physical explanation?

In [3,4] we introduced a completely heteropolymer model of chromosomes, where all potentials are different. The calculations with 100 kilo base pairs (kbp) resolution showed quite good ability of the model to capture the main properties of contact maps, visible in experiments. What will change with increasing model resolution? To answer this question, in this paper we investigate the properties of the previously introduced heteropolymer model of the chromosome with a higher resolution, with 50 and 10 kbp resolution.

The distinct feature of model [3] is that it does not suggest that the genome/chromosome is partitioned into two [1] or six [2] types of *sub*-compartments equivalent to two or six types of polymer blocks of monomers. In [3] and in the present model we consider completely heterogeneous polymer model where all chromosome subunits are assumed to be different and interaction energies between any pair of elements are different. Here we further investigate the model's ability to describe chromosome contact Hi-C data and compare some results with prediction of models based on compartment idea which is, in turn, based on suggestion of microphase separation.

We obtain whole-chromosome contact maps and compare the quality of their description of experimental maps with a polymer model of six types of chromatin subunits. The conclusion we arrive at is that macroscopic compartmentalization, first observed in [1], does not require a microphase separation mechanism to explain it, as it is widely interpreted.

## Results

*Heterogeneous polymer chromosome model describes the pattern of intrachromosomal contacts better than model postulating microphase separation*

To test the descriptive capacity of the heterogeneous chromosome model, we used contact maps of chromosome 22 in GM12878 cells at higher resolution than previously used [3]. Fig.1 shows the results of comparison between simulated and experimental [2] contact maps with different resolutions.

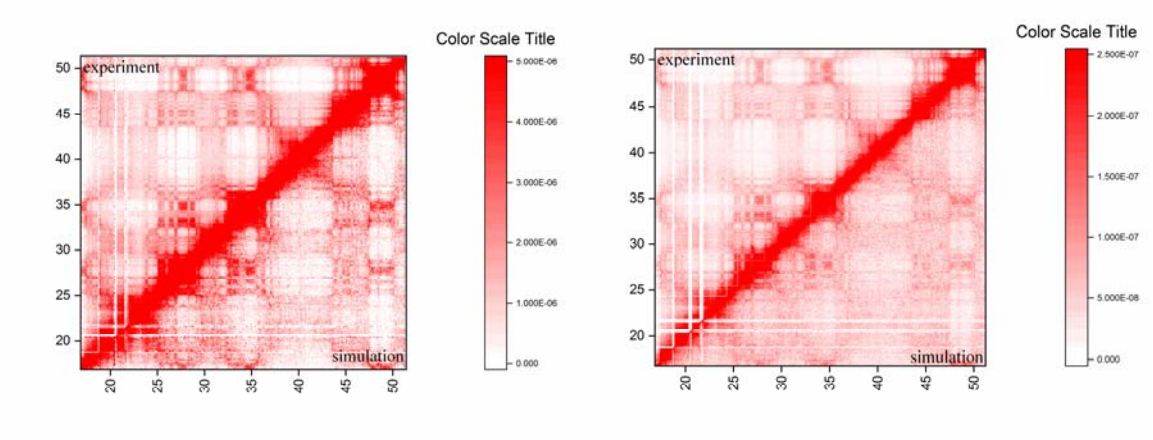


Fig.1. Heterogeneous polymer model describes Hi-C maps for chromosome 22 in GM12878 cells. Model resolutions are 50 kbp (left) and 10 kbp (right). Experiment: [2]. Pearson correlation: for 50 kbp R=0.982, for 10 kbp R=0.978.

For comparison, the results of calculations on the 6-subcompartment model with microphase separation [5] with 50 kbp resolution are presented, see Fig.2.

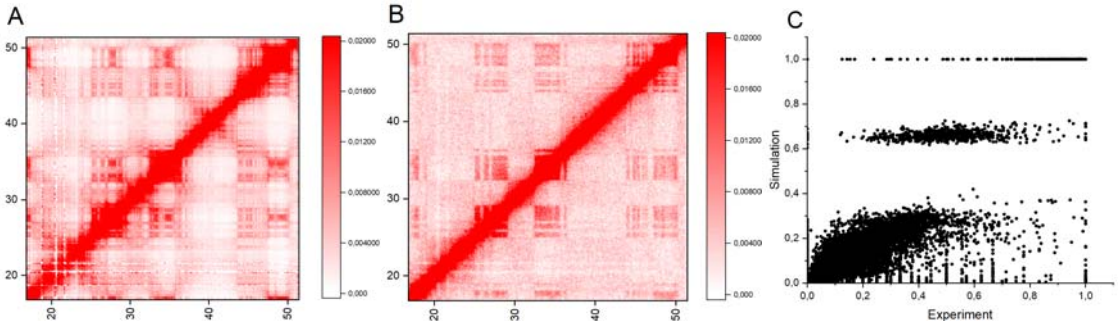


Fig.2. Simulation by 6-subcompartment model compared to Hi-C data [2]. The model and its parameters were taken from [5]. **A, B**: Hi-C maps with 50 kbp resolution. **A**: experiment. **B**: simulation. **C**: correlation graph. Unlike [5], all points are presented in the graph. Pearson correlation R=0.863.

*Checking the contact map compartments for uniformity*

The concept of compartments imposes a number of conditions which should be met within the compartments on the map: homogeneity of contacts within the compartments, homogeneous potentials, similar interactions of elements within the compartments of given type [1,2,5].

We checked these postulates by simulations using our more detailed heterogeneous model, which describes the entire map well without assuming homogeneous blocks of elements. To do this, we selected a number of blocks in the contact maps (Fig.3) and tested them for the requirements of "homogeneity". Instead of raw contact maps, we used “observed/expected” maps to offset the contact heterogeneity associated with the influence of genomic separation.

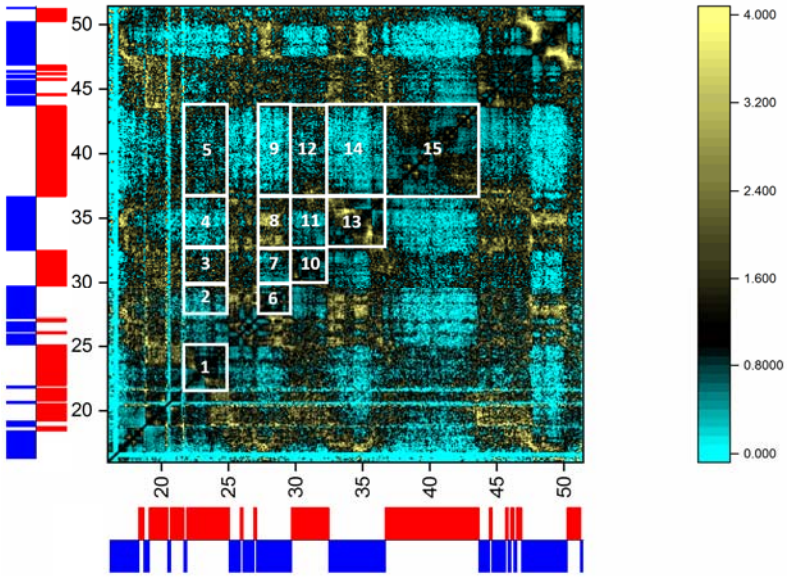


Fig.3. Simulated contact map (observed/expected) of chromosome 22 for contact heterogeneity analysis. Heterogeneous chromosome model, resolution 50 kbp, map coarsened to 100 kbp. To the left and below: ideogram of A/B compartments obtained with 100 kbp resolution. A: red, B: blue. The map was divided into blocks according to compartment boundaries. 15 blocks picked for analysis together with their numbers are shown in the map (white).

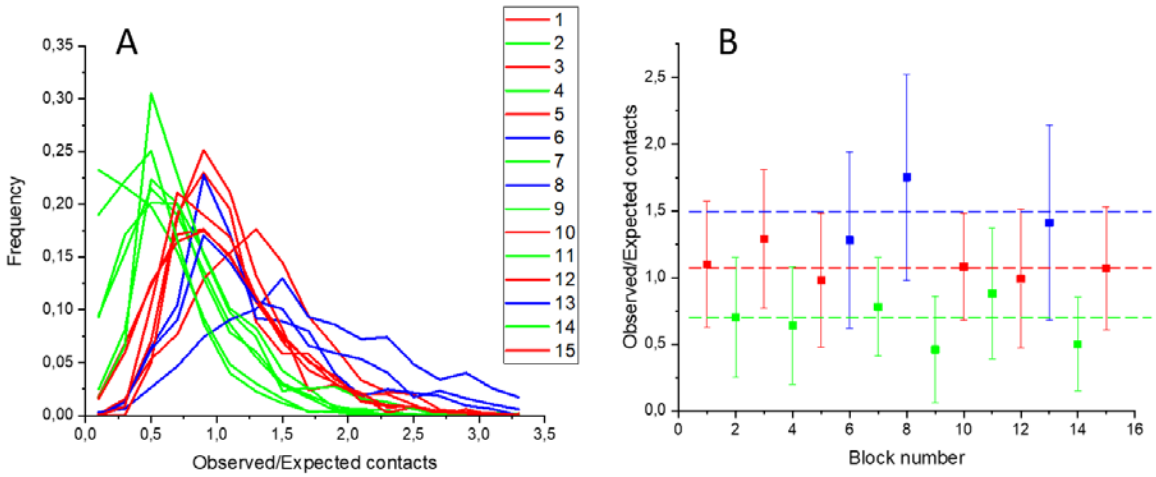


Fig.4. Analysis of contact heterogeneity in selected blocks of chromosome 22 contact map with resolution 50 kbp. Observed/Expected map was used to neutralize the bias caused by genomic separation. Block boundaries were determined according to boundaries of A/B compartments. **A**: distribution of contacts within each block. Block numeration is shown in Fig.3. Colors: red – A-A blocks (i.e. both regions of the chromosome belong to A compartment, see Fig.3); blue – B-B blocks; green – A-B blocks. **B**: Mean number of contacts and standard deviations for each block. Dashed lines denote means for all blocks of corresponding types.

One can conclude from Figs 3 and 4 that there are heterogeneous distributions of contacts within compartment blocks already in 50 kbp resolution and strong deviations from the average values of frequency of contacts. Moreover, the frequency of contacts of different types (A-B) does not differ much from the frequency of contacts of the same type (A-A or B-B). This fact directly contradicts the postulate of two compartments on the contact map of chromosome 22 with 50 kbp resolution and supports our heteropolymer model.

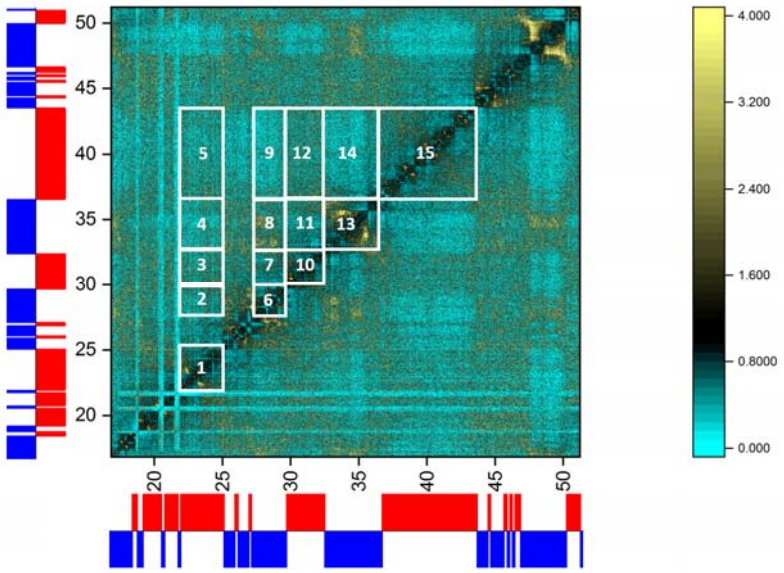

Fig.5. Simulated contact map (observed/expected) of chromosome 22 for contact heterogeneity analysis. Resolution 10 kbp. To the left and below: ideogram of A/B compartments obtained with 100 kbp resolution. A: red, B: blue. The map was divided into blocks according to compartment boundaries. The same 15 blocks as for 50 kbp model (Fig.3) were picked for analysis (white).

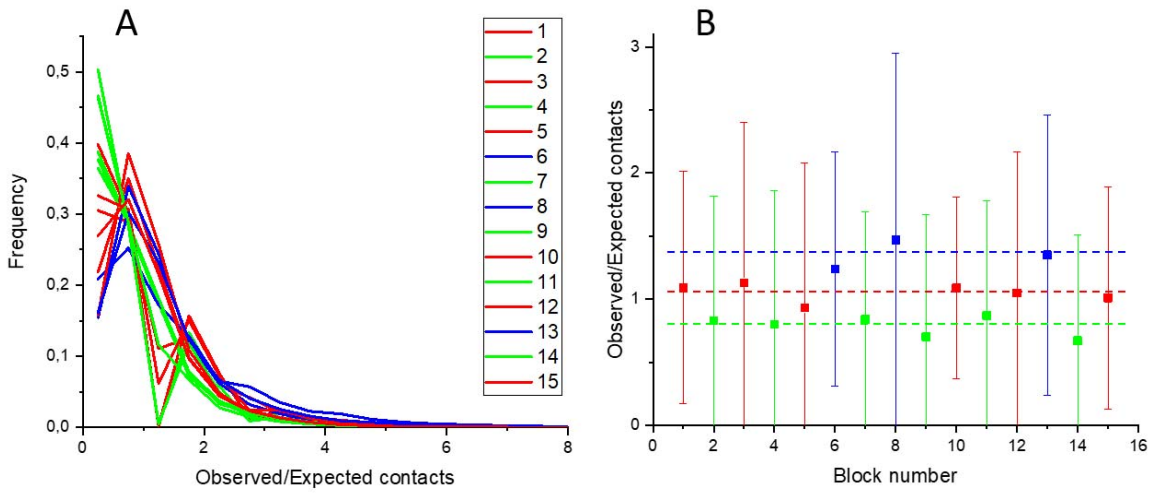


Fig.6. Analysis of contact heterogeneity in selected blocks of chromosome 22 contact map with resolution 10 kbp. Block boundaries were determined according to boundaries of A/B compartments. **A**: distribution of contacts within each block. Block numeration is shown in Fig.5. Colors: red – A-A blocks; blue – B-B blocks; green – A-B blocks. **B**: Mean number of contacts and standard deviations for each block. Dashed lines denote means for all blocks of corresponding types.

Figs 5 and 6 show results of the same analysis performed on maps with higher resolution, 10 kbp. The conclusion is the same as for 50 kbp resolution. Moreover, the heterogeneity of contact frequencies within each block is even higher and difference between same-type and different-type blocks is even lower than for 50 kbp resolution which makes the compartment conception meaningless.

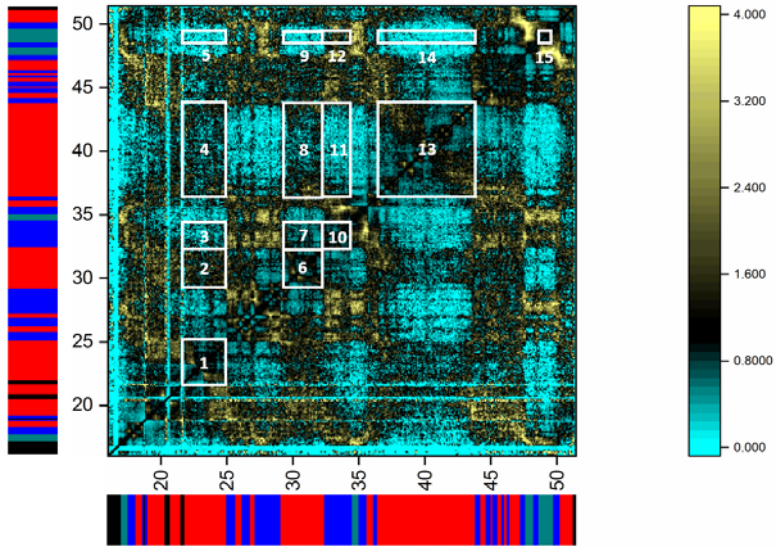


Fig.7. Simulated contact map (observed/expected) of chromosome 22 for contact heterogeneity analysis. Heterogeneous chromosome model, resolution 50 kbp, map coarsened to 100 kbp. To the left and below: ideogram of A1/B1/B2 subcompartments from [2], 100 kbp resolution. A1: red, B1: blue, B2: teal, N/A: black. A2 subcompartment is not present in chromosome 22. The map was divided into blocks according to subcompartment boundaries. 15 blocks picked for analysis together with their numbers are shown in the map (white). Their positions are partially different from those picked previously (Figs 3, 5) as subcompartment boundaries are different from A/B case considered previously.

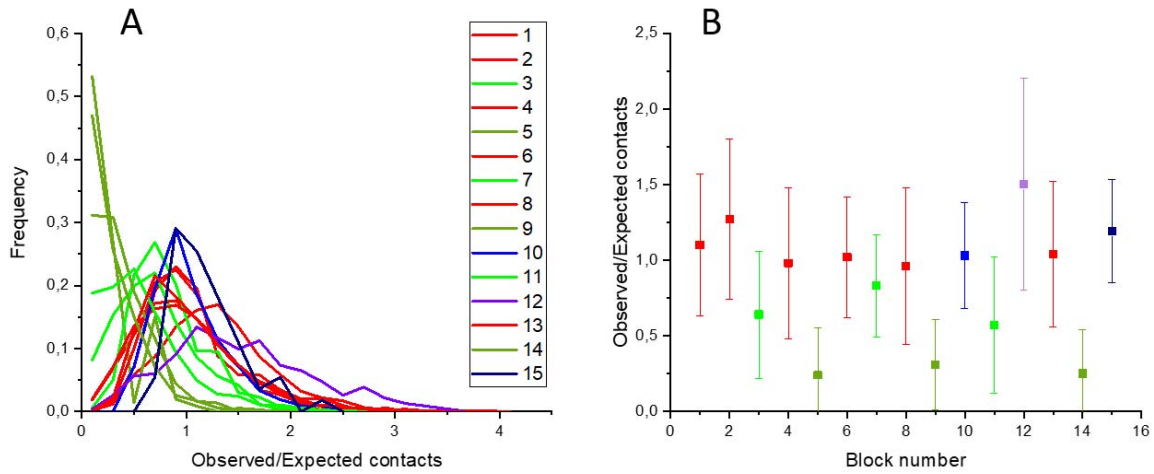


Fig.8. Analysis of contact heterogeneity in selected blocks of chromosome 22 contact map with resolution 50 kbp. Block boundaries were determined according to boundaries of A1/B1/B2 subcompartments. **A**: distribution of contacts within each block. Block numeration is shown in Fig.7. Colors: red – A1-A1 blocks; blue – B1-B1 blocks; dark blue – B2-B2 blocks; green – A1-B1 blocks; dark green – A1-B2 blocks; violet – B2-B2 blocks. **B**: Mean number of contacts and standard deviations for each block.

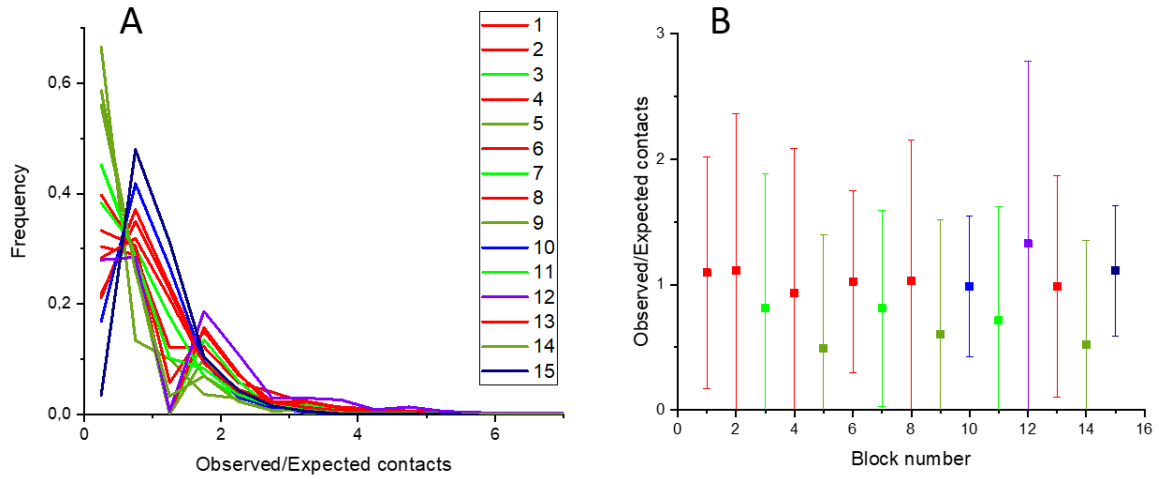


Fig.9. Analysis of contact heterogeneity in selected blocks of chromosome 22 contact map with resolution 10 kbp. Block boundaries were determined according to boundaries of A1/B1/B2 subcompartments. **A**: distribution of contacts within each block. Block positions and numeration and the same as for 50 kbp resolution (Figs 7, 8). Colors: red – A1-A1 blocks; blue – B1-B1 blocks; dark blue – B2-B2 blocks; green – A1-B1 blocks; dark green – A1-B2 blocks; violet – B2-B2 blocks. **B**: Mean number of contacts and standard deviations for each block.

The same analysis was repeated for 6-subcompartment conception [2] which was supposed to be more relevant than 2-compartment conception at resolutions as high as 50-100 kbp [2,5]. The results are presented in Figs 7-9. Only one type of distributions was visually different from others (A1-B2, resolution 50 kbp, Fig.8), but in a whole picture the previous conclusion stands: heterogeneity within each block is higher than between same-type and different-type blocks. This contradicts the postulate of several subcompartments on the contact map of chromosome 22 with both 50 and 10 kbp resolution and supports our heteropolymer model.

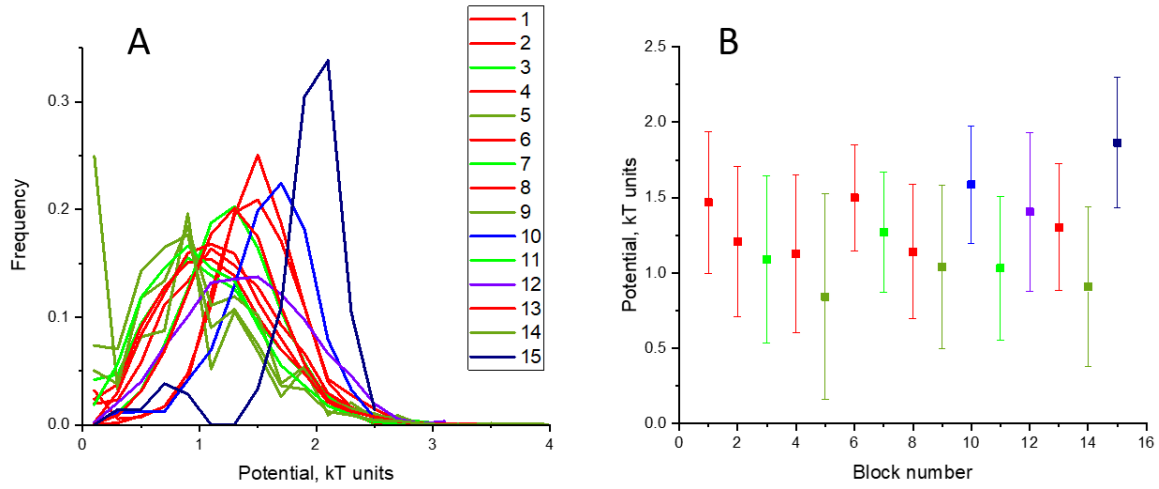


Fig.10. Analysis of potential heterogeneity in selected blocks of chromosome 22 contact map with resolution 50 kbp. Blocks are the same as in contact heterogeneity analysis (Fig.8). **A**: distribution of potentials within each block. Block numeration is shown in Fig.7. Colors: red – A1-A1 blocks; blue – B1-

B1 blocks; dark blue – B2-B2 blocks; green – A1-B1 blocks; dark green – A1-B2 blocks; violet – B2-B2 blocks. **B**: Mean potentials and standard deviations for each block.

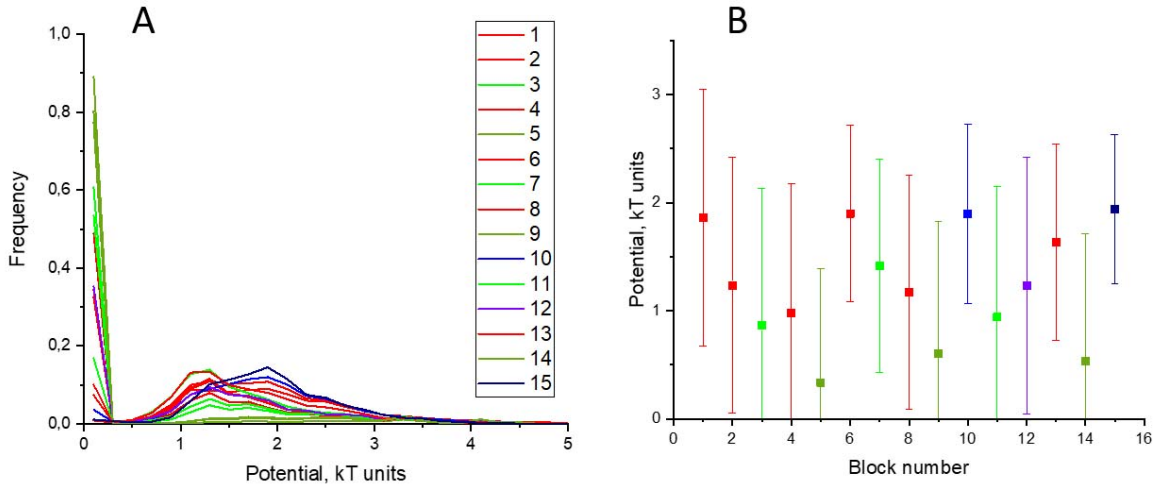


Fig.11. Analysis of potential heterogeneity in selected blocks of chromosome 22 contact map with resolution 10 kbp. Blocks are the same as in contact heterogeneity analysis (Fig.8). **A**: distribution of potentials within each block. Block numeration is shown in Fig.7. Colors: red – A1-A1 blocks; blue – B1-B1 blocks; dark blue – B2-B2 blocks; green – A1-B1 blocks; dark green – A1-B2 blocks; violet – B2-B2 blocks. **B**: Mean potentials and standard deviations for each block.

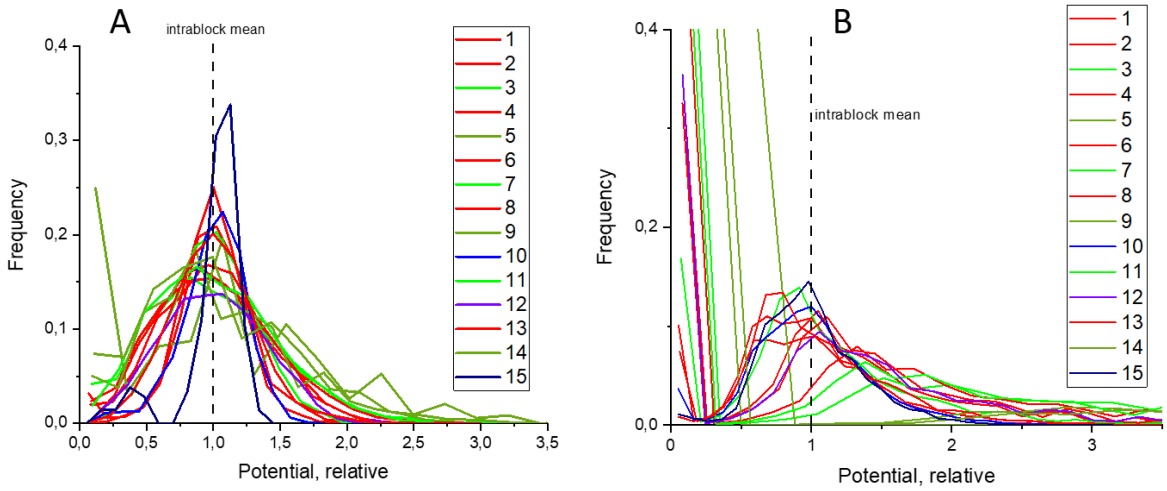


Fig.12. Potentials within each block are highly heterogeneous. **A**: distribution for relative potentials for 50 kbp simulation. **B**: for 10 kbp simulation. Relative potentials were obtained for each block by dividing potentials by the mean for this block. Thus, all means equal unity (vertical dashed lines). Block numbers and line colors are the same as in corresponding distributions for absolute potentials (Figs 10, 11).

Figures 10-12 show the presence of wide potential distributions (heterogeneity) necessary to describe the contact frequencies in different blocks with 50 and kbp resolutions according to our heterogeneous model. They are very different from the homogeneous potentials in blocks, which are postulated by a model with a microphase separation with 50 kbp resolution [5]. For 10 kbp the potential spread is even

wider than for 50 kbp, which indicates the heterogeneity of interactions and does not support the assumption of homogeneous blocks of similar elements and, accordingly, blocks of identical interaction energies between repeating elements [1,2,5]

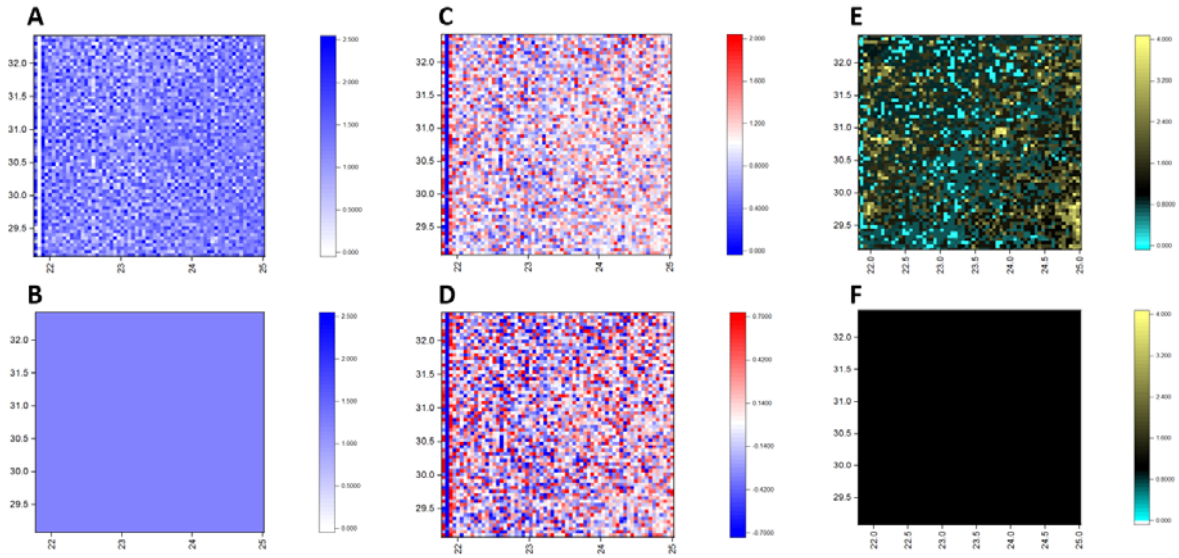


Fig.13. Interconnection between heterogeneity of potentials, i.e. absence of (sub)compartments, and heterogeneity of contacts. As an example, block 2 for 50 kbp resolution model is taken, 21.8-25 x 29.1-32.4 Mbp, A1-A1 type (see Fig.7). **A**: simulated heterogeneous map of potentials. **B**: the corresponding map for the case with all potentials equal to the mean (<u>=1.21 kT). **C** and **D** illustrate potential heterogeneity. **C**: simulated map of relative potentials $u_1(i,j)=u(i,j)/\langle u\rangle$. **D**: simulated map of potential relative deviation $u_2(i,j)=(u(i,j)-\langle u\rangle)/\langle u\rangle$. **E**: contact map (observed/expected) obtained with heterogeneous potentials in panel A. **F**: contact map assumed from homogeneous potentials (panel B).

Fig.13 demonstrates that heterogeneous contacts are more adequately described by heterogeneous potentials that homogeneous one for given block or compartment.

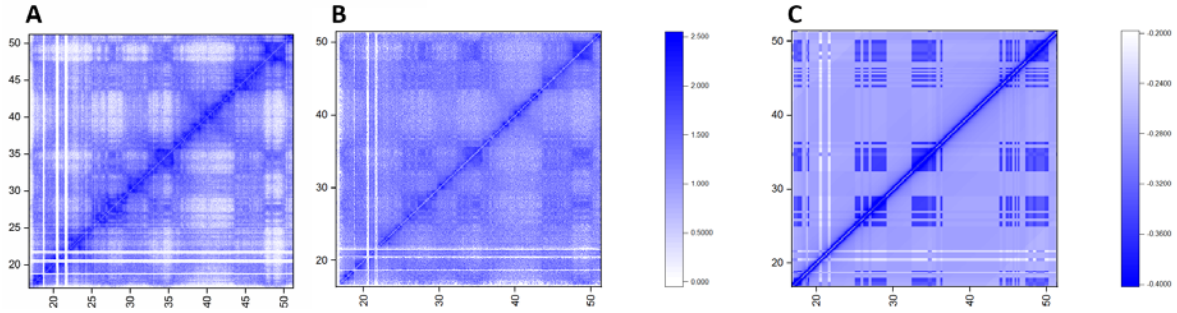


Fig.14. Maps of potentials for different models of chromosome 22. A: heterogeneous model, resolution 10 kbp. B: heterogeneous model, resolution 50 kbp. C: 5-subcompartment model [5], resolution 50 kbp. The values are in kT units.

Fig.14 shows the difference in distribution of energy interactions along the chromosome between heterogeneous chromosome model and model based on microphase separation.

## Discussion

We conclude that there are heterogeneous distributions of contact densities in different parts of chromosome which are due to heterogeneous interaction energies. Chromosome compartmentalization can be explained not by microphase separation of monomer blocks along the chromosome but by effect of heterogenic network of volumetric interaction potentials in chromosome. A radical consequence of this approach is that there are sequence-specific generalized volumetric potentials represent which on molecular level crosslinking factors, (multi)protein complexes, transcription factors, ncRNA, *etc*. They can define specificity of interaction energies on different scales. The macro-compartmentalization seen on high-resolution maps is partly the result and even the artifact of map coarsening. At high resolution there are wide distributions in all parameters (contacts, potentials) in all blocks which, when resolution is decreased and the map is coarsened, produce the effect of distinct blocks, sharp borders, compartments, *etc*. Homogeneous blocks of same-type elements at the microscopic level are not observed after analyzing the contact map with high resolution (50 and 10 kbp). Alternatively, we cannot rule out that the frequency jumps and skips in the maps and the resulting heterogeneity of interaction energies may be related to not very high-quality maps, and not to the heterogeneous nature of long-range interactions in chromosomes.

In conclusion, the main result of the introduction of a new heteropolymer model of the chromosome is as follows. Explanation of the compartmentalization of the contact map (chromosome 22, GM12878 cells) into 2 (A-B) or 6 types of chromatin subunits, is possible without the mechanism of microphase separation. Thus, contact maps can be explained without repeating blocks of same-type elements along the chromosome. I.e. our model does not require "compartments" along the chromosome, understood as blocks of the same type of elements. Although the mechanism of microphase separation is physically simplest and natural for the explanation of the checkerboard type of contact maps, it is not the only one and, as we show, not necessary.

## Methods

The interphase structure of human chromosome 22 was modeled with the inhouse software using molecular dynamics package OpenMM. A chromosome was modeled as a heteropolymer chain of subunits with either 50 kbp or 10 kbp DNA content. The interaction potential between any pair of non-neighbor subunits ($i$, $j$) consists of excluded volume potential and attraction potential. The former is the same for all pairs of subunits, while the latter depends on the position of the subunits along the chain. This is similar to the method developed by us in [2,3]. The coefficients for the attraction potential are determined to provide the best fit by the model the experimental Hi-C maps. To this end we developed the iterative algorithm described in [2,3].

**Acknowledgements**
The study was supported by MEPhI Academic Excellence Project (Contract No. 02.a03.21.0005); Russian Foundation for Basic Research grant 14-01-00825; The research is carried out using the

equipment of the shared research facilities of HPC computing resources at Lomonosov Moscow State University and Joint Supercomputer Center of the Russian Academy of Sciences.

**Author Contributions**

Y.E.: software, scientific analysis, work on manuscript. I.S.: software. S.S.: data acquisition and analysis. S.A.: conceptualization, scientific analysis, work on manuscript.

**Conflict of Interests**

The authors declare they have no conflicts of interest.

## References

1.*Lieberman-Aiden E., van Berkum N.L., Williams L. et al*. Comprehensive mapping of long-range interactions reveals folding principles of the human genome // Science. 2009. V. 326. № 5950. P. 289–293. doi: 10.1126/science.1181369

2.*Rao S.S., Huntley M.H., Durand N.C. et al*. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping // Cell. 2014. V. 159. № 7. P. 1665–1680. doi: 10.1016/j.cell.2014.11.021

3.*Eidelman Y.A., Slanina S.V., Pyatenko V.S., Andreev S.G.*
Chromosome compartmentalization replacement during stem cell differentiation
http://dx.doi.org/10.1101/284190 doi: bioRxiv preprint first posted online Mar. 18, 2018;

4.*Eidelman Y.A., Salnikov I.V., Slanina S.V., Andreev S.G.*
Spatial organization of the mouse chromosomes influences the landscape of intrachromosomal exchange aberration breakpoints
http://dx.doi.org/10.1101/852277 doi: bioRxiv preprint first posted online Nov. 23, 2019;

5.*Di Pierro M., Zhang B., Aiden E.L. et al*. Transferable model for chromosome architecture // Proc. Natl Acad. Sci. USA. 2016. V. 113. № 43. P. 12168–12173. doi: 10.1073/pnas.1613607113