

Loop Structures and Barrier Elements from *D. melanogaster* 87A7 Heat Shock Locus

Mikhail V. Glazkov*, Anna N. Shabarina

Koltzov Institute of Developmental Biology, Russian Academy of Sciences, Moscow, Russia

Email: *mvglazkov@yandex.ru

How to cite this paper: Glazkov, M.V. and Shabarina, A.N. (2016) Loop Structures and Barrier Elements from *D. melanogaster* 87A7 Heat Shock Locus. *Computational Molecular Bioscience*, 6, 53-65.

<http://dx.doi.org/10.4236/cmb.2016.64005>

Received: October 13, 2016

Accepted: December 2, 2016

Published: December 5, 2016

Copyright © 2016 by authors and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

The three-dimensional organization of the genome is closely related to its functioning. Interactions between parts of the genome located at large distances from each other have been detected within the chromosomes of different organisms, which led to the discovery of topologically associated domains (TADs). Methods that reveal such interactions between chromosomal loci imply detection of both protein-protein and protein-DNA interactions. We investigated the possibility of involvement of the direct DNA-DNA interactions in the structural and functional organization of *Drosophila melanogaster* chromosomal 87A7 locus, containing genes *hsp70Aa* and *hsp70Ab*, with the sequence analysis method. Our results indicate that the functional organization of 87A7 locus may involve different elements: chromosomal DNA fragments that attach chromosomes to the nuclear envelope, short polypurine/polypyrimidine tracts, insulators and their proteins. The combination of interactions of these elements may cause different functional states of 87A7 locus.

Keywords

hsp70 Genes, DNA-DNA Interactions, Insulators, Nuclear Envelope, TADs

1. Introduction

In recent years, great interest is attracted to interactions of different genomic regions of interphase chromosomes, which are located at large distances from each other (based on the linear dimensions of the chromosomal DNA) [1] [2] [3] [4] [5]. Physically separated chromosomal loci more frequently communicate with each other within certain domains, which are called “topological domains” [2] or “topologically associated domains”—TADs [6] with chromatin looping out of DNA region located in between. These domains persist in interphase chromosomes of different cell types; they are evolutionarily conserved on syntenic chromosomal regions of different organisms. The av-

erage size of topological domains was originally designated as 800 kb in the chromosomes of mouse and human embryonic stem cells [2]. Later [4] with the help of the Hi-C method with higher resolution (1 kb), the average size of these domains in human chromosomes was reduced to 185 kb. Similar domains with an average length of 60 kb are discovered in *D. melanogaster* chromosomes [3].

TADs were identified using various modifications of the 3C-method (Chromosome Conformation Capture). Formaldehyde fixation of protein-protein and protein-DNA interactions is one of the first steps of the method [1] [2] [3] [4] [5] as well as salt and EDTA treatment of cell nucleus that may distort the picture of the existing interactions. It is not surprising therefore that substantial part of the identified interactions refers to the enhancer-promoter interactions (related to the pattern of expressed genes), and interactions with the participation of insulator proteins.

However, one cannot exclude the involvement of direct DNA-DNA interactions in the structural and functional organization of chromosomes [7] [8]. Short polypurine/polypyrimidine tracts (9 - 10 bp), capable of forming the triple-helix DNA structure (H-form DNA) are possible participants of this process. Triplex DNA structures are revealed by immunocytochemistry in different types of chromosomes (interphase, metaphase, meiotic, polytene), they are usually detected in the intergenic regions and introns of different genes. Polypurine/polypyrimidine 9 - 10 bp long but no longer tracts are of very weak recombinogenicity (for details see [8]). When complementary polypurine and polypyrimidine tracts are located at a distance from each other, their interaction (formation of the H-form DNA) leads to looping out of nucleotide chain located in between. This was shown in *in vitro* experiments using both synthetic oligonucleotides [9] [10] [11], and the human alpha-1-antitrypsin gene [12], and while bioinformatic essays [7] [8]. It should be noted that the protein-coding and non-coding sequences (satellite DNA and repeats of LINE and SINE type) can form loop structures with fundamentally similar mechanism [8].

The question that still needs to be elucidated is how the “borders” between the adjacent TADs (including the nucleotide sequences of the chromosome regions) are organized [5]. However, based on the definition of TADs (see above) the “border” is the one to ensure cooperation between genomic regions inside the domain and to prevent (reduce the probability of) cross-domain interactions. This assumption applies to the function of insulators as well.

Earlier [13] [14] using transgenic system of *D. melanogaster yellow* and *white* genes, we have shown that the nuclear envelope DNA attachment sites (neDNA) are capable of reporter genes’ protection from the Position Effect Variegation (PEV), and the ability is improved in the presence of Wari insulator. If only one of the transgenes (*white* gene + Wari) is flanked by neDNA, it appears to be better protected against PEV. In other words, the insulator “works” well between the two neDNA fragments, while it has slight effect through the neDNA fragment. Interestingly, neDNA is evolutionarily conserved [15] [16]. These point nuclear envelope chromosomal attachment sites may also form looped chromatin structures [17] [18]. The neDNA fragment (EnvM4), used in

our transgenic experiments, possess extremely conserved motif $(AAAGA)_n$. DNA motifs (GAGA), responsible for the attachment to the nuclear envelope, were revealed within the LADs (Lamina Associated Domains) [19]. However, it should be noted that represented consensus contains AAAGA motifs as well. The most common feature of these two types of DNA motifs is the presence of polypurine tandem copies.

TADs are not detected both in mitotic chromosomes [20] and in the inactivated X-chromosome [6]. Perhaps one reason for identifying no TADs in mitotic chromosomes is the absence of the one of the components involved in the formation of TADs' borders at metaphase stage. That component turns to be the nuclear envelope/nuclear lamina, the one to attach interphase chromosomes at various sites. Concerning inactivated (compactised) X-chromosome, identifying no TADs may be of the same cause as the compaction of chromosomes is accompanied by a change in the pattern of its interaction with the nuclear envelope/nuclear lamina.

Despite the identification of TADs in many organisms and many studies in this field, much remains unclear. The data on TADs give mostly a general idea about the interactions between the different sections of DNA. A detailed study of the nucleotide sequences, transcription, epigenetic status as well as higher levels of chromatin organization can help to create a complete picture of the processes taking place in the genome.

This study is devoted to *D. melanogaster* heat shock genes (*hsp70*) locus. These genes encode proteins of *hsp70* family performing the protective functions of the cell, preventing denaturation and aggregation of proteins. It is known that heat shock genes are expressed regardless of the surrounding chromatin [21]. We explored the question of a possible mechanism for the establishment of an independent expression domain of heat shock genes at the level of the loop chromatin organization with the involvement of: neDNA (nucleotide motifs—nuclear envelope attachment “markers” $(AAAGA)_n$ and $(GAGA)_n$), short “complementary” polypurine/polypyrimidine tracts as well as insulators of the locus (*scs/scs'*-elements) and their proteins (BEAF32 and Zw5)). This paper presents an analysis of the nucleotide sequences of the locus and the attempt to compare it with the available experimental data.

2. Materials and Methods

D. melanogaster chromosomal locus 87A7 contains *hsp70Aa* and *hsp70Ab* genes (3R: 11,904,163 .. 12,006,544), length 102,382 bp (FlyBase, release 6,0).

Search for nucleotide sequences able to form three-chain DNA structures was performed according to the following criteria: polypurine/polypyrimidine tracts should be potentially able to form T(A.T), A(AT), C(GC), and G(GC) triplexes not less than 9 bp long. To simplify the simulation the search for complementary polypurine/polypyrimidine tracts was led on the same DNA strand [7] [8]. We used NCBI resources, flybase.org, including Drosophila Sequence Coordinates Converter (http://flybase.org/static_pages/downloads/COORD.html) and Vector NTI software. Juicebox software was also used to analyse Hi-C data.

3. Results

3.1. General Characteristics of the Region

The chromosomal segment of *D. melanogaster* chromosome 3 (3R: 11,904,163 .. 12,006,544) is 102,382 bp in length with 17 genes located within (from CG14731 gene to Ect3 gene) (Figure 1). The research revealed 152 polypyrimidine (Y) tracts and 186 polypurine (R) tracts no less than 9 bp long in the region (maximum size of identified tracts is 27 bp). Among them 120 polypyrimidine tracts and 135 polypurine tracts are potentially able to interact with each other with looping out of DNA located in between. Most of the tracts can possibly interact with a numerous complementary tracts located in various sites of analyzed region on chromosome 3.

Scale and representation of genes is taken from *flybase*. Arrows show the direction of transcription. Localization of the functional areas identified in *hsp70* genes locus is showed with figures: black squares—poly-A/T tracts; triangles—insulators class II (Su(Hw)); circles—insulators class I (CP190, BEAF32, CTCF); asterisks—(AAAGA)₂ tracts; rhombi—scs/scs'-elements. The border between physical domains is marked with vertical dashed lines. Diagonal hatching represents the areas of frequent intra-domain interactions. Diagonal cells denote regions that are found within the loop for *hsp70* genes/one loop for the whole locus and black rectangles denote the bases of these loop structures respectively.

The insulators of *hsp70* genes locus are localized in the area: the region of scs'-element overlaps with the promoter of CG31211 gene (3R: 11,948,272 .. 11,950,063) and includes Y82-Y84 and R108-R110 tracts, which have 31 complementary tract in the test region. Another insulator, scs'-element, is located in the area (3R: 11,962,738 .. 11,963,832) overlapping with the promoters of CG3281 and *aurora* genes. This region includes tracts R123, R124, and there is only one polypyrimidine tract complementary to R124 (data not shown). Flybase.org resource shows the localization of two types of recognition sites for the insulator proteins at the locus. The first are for the insulators class II.mE01 (contain recognition sites for Su(Hw) protein): insulator_II_2339 and insulator_II_2340. The second type of sites are for insulators class I.mE01 (contain recognition sites for two of three insulator proteins—CP190, BEAF32 and CTCF): insulator_I_3071, insulator_I_3072, insulator_I_3073, insulator_I_3074 (Table 1, Figure 1).

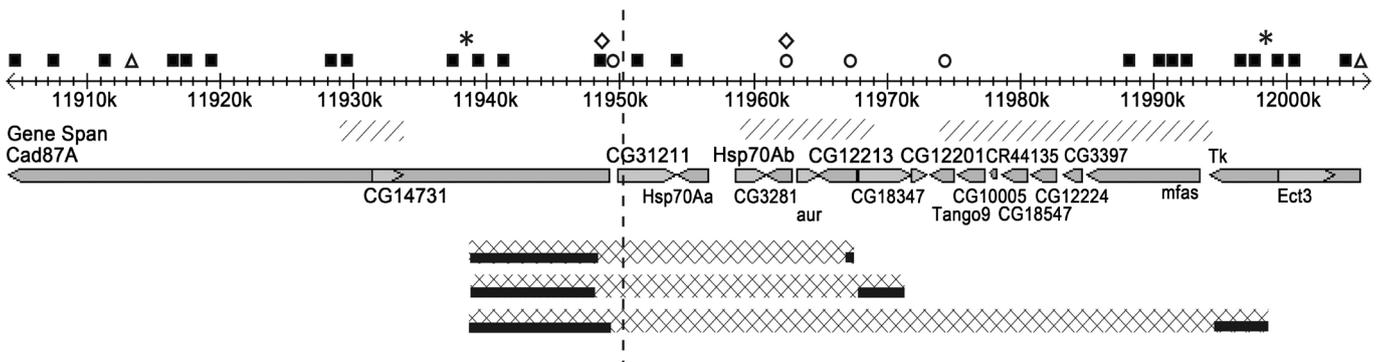


Figure 1. Schematic representation of *D. melanogaster hsp70* genes locus.

Table 1. Sequence analysis results for the *D. melanogaster hsp70* genes locus (3R: 11,904,163 .. 12,006,544; 102382 bp).

	Insulator class_II.mE01 (flybase.org)	FBsf0000154391 (11,913,012 .. 11,913,022)		
		FBsf0000154392 (12,005,260 .. 12,005,270)		
		FBsf0000158034 (11,949,333 .. 11,949,343)		
Insulators	Insulator class_I.mE01 (flybase.org)	FBsf0000158035 (11,962,949 .. 11,962,959)		
		FBsf0000158036 (11,967,340 .. 11,967,350)		
		FBsf0000158037 (11,974,832 .. 11,974,842)		
	scs	11,948,272 .. 11,950,063		
	scs'	11,962,738 .. 11,963,832		
Conservative tract of EnvM4 fragment (AAAGA) ₂		1) 11,938,828 .. 11,938,838		
		2) 11,998,731 .. 11,998,742		
		1) 11,920,264 .. 11,920,272		
		2) 11,923,379 .. 11,923,387		
		3) 11,937,609 .. 11,937,616		
		4) 11,937,706 .. 11,937,714		
	(GAGA) ₂ tracts		5) 11,947,759 .. 11,947,773	
			6) 11,947,966 .. 11,947,974	
			7) 11,971,004 .. 11,971,014	
			8) 11,973,193 .. 11,973,200	
			9) 11,980,525 .. 11,980,533	
			10) 11,985,662 .. 11,985,671	
		Poly (A/T) tracts		1) 11,904,451 .. 11,904,462
				2) 11,907,560 .. 11,907,569
				3) 11,919,007 .. 11,919,020
			4) 11,929,000 .. 11,929,017	
	5) 11,941,394 .. 11,941,420			
	6) 11,948,318 .. 11,948,326			
	7) 11,954,049 .. 11,954,058			
	8) 11,988,599 .. 11,988,615			
	9) 11,990,094 .. 11,990,098			
	10) 11,990,436 .. 11,990,454			
	11) 11,992,431 .. 11,992,443			
	12) 11,996,508 .. 11,996,516			
	13) 11,997,711 .. 11,997,719			
	14) 11,999,479 .. 11,999,489			
	15) 12,004,298 .. 12,004,306			
	16) 12,004,671 .. 12,004,680			
	17) 11,911,373 .. 11,911,382			
	18) 11,911,858 .. 11,911,869			
	19) 11,916,814 .. 11,916,826			
	20) 11,917,018 .. 11,917,027			
	21) 11,919,252 .. 11,919,261			
	22) 11,928,709 .. 11,928,721			
	23) 11,937,563 .. 11,937,572			
	24) 11,939,459 .. 11,939,473			
	25) 11,939,860 .. 11,939,871			
	26) 11,941,989 .. 11,941,999			
	27) 11,948,092 .. 11,948,104			
	28) 11,951,658 .. 11,951,667			
	29) 11,988,284 .. 11,988,294			
	30) 11,991,175 .. 11,991,183			
	31) 12,000,454 .. 12,000,469			

Continued

Sites with the potential to form single looped structure (with the help of triple-stranded DNA)	(11,938,828 .. 11,949,416) .. (11,994,767 .. 11,998,654)
Sites with the potential to form genes <i>hsp70</i> loop (with the help of triple-stranded DNA)	(11,938,916 .. 11,948,717) .. (11,967,091 .. 11,967,605) (11,938,828 .. 11,948,218) .. (11,968,206 .. 11,971,622)
TADs localization [3]	
Localization of physical domains (modelling)	Domain ID 748 11,821,528 .. 11,950,027 (Null) Domain ID 749 11,950,028 .. 12,219,527 (Null)
Areas of intra-domain interactions (± 5 kb) (experimental data)	11,929,278 .. 11,934,278 11,959,278 .. 11,969,278 11,974,278 .. 11,994,278

The sequence analysis of the region revealed a large number of single GAGA and AAAGA tracts, but not dimers that are only 10 and 2, respectively (Table 1, Figure 1). The size of the chromosomal region flanked by the dimers of conservative AAAGA tract is about 60 kb (Table 1, Figure 1). It should be noted that the area of the locus bounded with (AAAGA)₂ tracts on both sides lies within an area bounded on both sides with the insulator protein Su (Hw) binding sites. In turn, the binding sites for insulator proteins CP190, BEAF32 and CTCF are located within the area bounded by the tracks (AAAGA)₂. Most of the (GAGA)₂ tracts are also located in the same region.

The formation of the DNA loop structures (bending of the DNA molecule) is facilitated by poly-A and poly-T tracts [22]. Furthermore, it is shown that poly-A and poly-T rich DNA regions selectively interact with nuclear envelopes *in vitro* [23], enrich genomic regions that interact with nuclear lamina [24]. Sequence analysis of the locus has revealed 31 poly-A/T tracts in the region (Table 1, Figure 1). Interestingly, (AAAGA)₂ sites are enriched with poly-A/T tracts (Figure 1).

Accordingly, the detailed analysis of the possibilities for loop structures forming was carried out in the region flanked by (AAAGA)₂ tracts on both sides. A number of 55 polypyrimidine tracts capable of interacting with 189 polypurine tracts were found in the region.

3.2. The Whole Area between (AAAGA)₂: Tracts Can Form a Single Loop Structure

The first question that we were interested in is whether the entire chromosomal region, located between the two conservative dimers AAAGA, form a loop by means of three-stranded DNA structures? It turned out that it is potentially possible, with 11 options of such loops. Four polypyrimidine and 14 polypurine tracts in the region of about 10.5 kb long between genes CG14731 and CG31211 (3R: 11,938,828 .. 11,949,416) have complementary polypurine (4) and polypyrimidine (7) tracts in the region of about 4 kb long between genes *mfas* and *Ect 3* (3R: 11,994,767 .. 11,998,654). The size of these potential loop structures is about 60 kb. Interestingly, the sites with the potential to form such looped structures are flanked by conservative (AAAGA)₂ as well as they are rich in

poly-A and poly-T (**Table 1, Figure 1**).

3.3. *hsp70* Genes Can Be Localized in The Loop Structures of Small Size

Genes *hsp70Aa* and *hsp70Ab* can be organized in a loop of a smaller size (about 30 kb). The two looped structures can be formed potentially. The first loop composition: 6 tracts localized between genes CG14731 and CG31211 (11,938,916 .. 11,948,717) are complementary to 4 tracts in the CG12213 gene (11,967,091 .. 11,967,605). In this case, genes CG31211, *hsp70Aa*, *hsp70Ab*, CG3281, and *aurora* are found within the loop. The second loop composition: 11 tracts localized between genes CG14731 and CG31211 (11,938,828 .. 11,948,218) are complementary to 5 tracts in the CG18347 gene (11,968,206 .. 11,971,622). Thereby the genes CG31211, *hsp70Aa*, *hsp70Ab*, CG3281, *aurora*, and CG12213 enter the loop. Since the tracts between the CG14731 and CG31211 genes largely overlap, either one or the other loop can be realized potentially (**Figure 1**).

Interestingly, polypurine/polypyrimidine tracts located in the 5'-region of the locus also overlap when forming a single loop structure and smaller loops of the analyzed chromosome segment (**Figure 1**). It means that the formation of a smaller loop (isolation of genes *hsp70Aa* and *hsp70Ab*) is accompanied by the destruction of a larger loop, and vice versa, the formation of a "big" loop may be accompanied by the destruction of the "smaller" loop. Perhaps the destruction of the large loop appears as puffing on the cytological level. Simultaneously *scs*- and *scs'*-elements are physically placed inside the area of decondensed chromatin. This consequence of our simulation corresponds with some experimental data. First, while heat shock *scs*- and *scs'*-elements are identified within the puff, but not at the borders of condensed and decondensed chromatin [25]. Second, conservative AAAGA tracts were found at the band/interband border on *D. melanogaster* polytene chromosomes using FISH-analysis (unpublished data).

3.4. A Possible Mechanism for *hsp70* Gene Independent Expression Domain Formation

It can be assumed that the formation of a domain of active *hsp70* genes is implemented in 2 steps. At the first stage the loop is formed with the participation of polypurine/polypyrimidine tracts (the formation of two types of loops is possible). BEAF32 and Zw5 proteins of *scs*/*scs'*-elements participate at the second stage (**Figure 2**).

Step 1. While formation of the first loop (see above), CG31211, *hsp70Aa*, *hsp70Ab*, CG3281, and *aurora* genes are found within the loop. When forming the second loop (see above), CG31211, *hsp70Aa*, *hsp70Ab*, CG3281, *aurora*, and CG12213 genes enter the loop. Since the tracts between the CG14731 and CG31211 genes largely overlap, either one or the other loop can be realized potentially.

Step 2: When forming the hinge structure at step 1, *scs*- and *scs'*-elements become sufficiently closer to each other (small distance apart in the nuclear volume). Perhaps it improves the conditions for interaction of BEAF32 and Zw5 proteins, which is shown

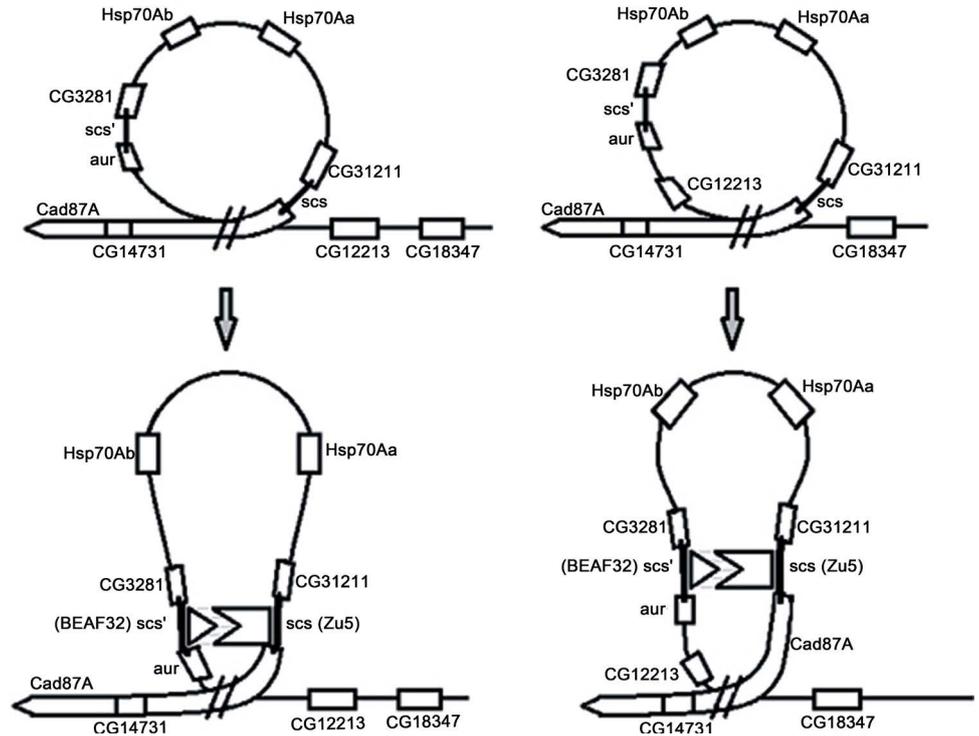


Figure 2. Schematic representation of the two variants of loop structures for *hsp70* genes that can be formed with the participation of triple-stranded DNA.

in vitro and *vivo* [26]. A loop within the loop is formed (Figure 2). This results in CG31211, CG3281, and *aurora* genes' inactivation because BEAF32 protein binding site, which specifically binds to *scs'*-element, overlaps with DREF transcription factor binding site in promoters of CG3281 and *aurora* genes [27]. Similar reason may probably underlie CG31211 gene inactivation, as *scs*-element specifically binded by Zw5 protein overlaps with the promoter of the gene. Thus, in the first loop formed with polypurine/polypyrimidine tracts only *hsp70Aa* and *hsp70Ab* genes can be in the active state. In the second loop formed with polypurine/polypyrimidine tracts inactivation of CG31211, CG3281, and *aurora* genes can be reached by the same means. Isolation of gene CG12213 from *hsp70Aa* and *hsp70Ab* genes may be due to the fact that it is situated between the bases of the loops formed by a) triplex DNA structures, b) BEAF32 and Zw5 proteins (Figure 2).

3.5. Loops and TADs of 87A7 Locus

To analyze *Drosophila* TADs we used data by Sexton *et al.* [3]. It turned out that 87A7 locus comprises two physical domains. The boundary between them is located upstream to genes *hsp70Aa* and *hsp70Ab* and falls into both “small” chromatin loops and “big” one formed by DNA-DNA interactions as well. Hi-C (TADs) data shows that the areas of frequent intra-domain interactions are localized within *hsp70Aa* and *hsp70Ab* genes loop (regardless of the loop formation variant), at insulators class I.mE01 and *scs'* region. Another area, characterized by increased frequency of contacts, is localized in

the big loop, the formation of which involves conservative tracts $(AAAGA)_n$.

4. Discussion

4.1. Chromosomal Domain of 87A7 Locus

The results of our study show that the short polypurine/polypyrimidine tracts may be involved in the loop organization of *D. melanogaster* chromosomal locus for *hsp70* genes. Tracts $(AAAGA)_2$ appeared to be the most significant nucleotide sequences to determine the domain at 87A7 locus. These tandem tracts are of extremely high evolutionary conservation, are localized mainly in the intergenic regions [16] of the genomes of various organisms (including *D. melanogaster*) and at the borders of the band/interband on polytene chromosomes of *Drosophila* (unpublished data).

The entire locus region located between two $(AAAGA)_2$ tracts can be arranged into loops with the help of complementary polypurine/polypyrimidine tracts (Figure 1, Figure 2). Moreover, this area has the potential to form a large number of smaller loops with the participation of complementary polypurine/polypyrimidine tracts. For its size (approximately 60 kb) and a large number of potential intra-domain interactions (135 of 186 polypurine tracts have complementary polypyrimidine tracts (72.6%), in turn, 120 of 152 polypyrimidine tracts have complementary polypurine tracts (78.9%)), these domains may correspond to TADs of *Drosophila* chromosomes. We assume that TAD may correspond to the area of interphase chromosome located between the two sites that anchor interphase chromosomes to the nuclear envelope (in this particular case—between two evolutionarily conserved dimers AAAGA).

The comparison of our data (Table 1 and Figure 1) with the data of Sexton *et al.* [3] shows that there are two physical domains in the investigated locus, the boundary between which is localized within the region of “small” loops (“functional” domain of *hsp70* genes). Both physical domains, comprising locus 87A7, have the “inactive domain” status. Explanation of this is the fact that *Drosophila* embryonic cells used for the experiment [3] have inactive heat shock genes, *i.e.* in the compact state, which may limit their interaction with the neighboring regions. In addition, the compact state of this region can inhibit interactions of flanking areas. It should be noted that chromatin architecture of a specific locus as well as its functional state, can be tissue- and stage-specific, therefore, this specificity should be taken into account when interpreting results.

According to Hi-C [3] two of three areas of frequent intra-domain interactions (TADs) at the 87A7 locus are located within the region bounded by a conservative tracts $(AAAGA)_n$ on both sides. More frequent interactions occur between sites that may be involved in the formation of both “large” and “small” loops of chromatin in the locus, with the help of direct DNA-DNA interactions. That is, the increased frequency of contacts is observed in the areas of chromatin, organized in loops. These data are also consistent with the fact that TAD may correspond to the interphase chromosome region located between the two sites of its “anchoring” to the nuclear envelope.

4.2. Functional Domain for *hsp70* Genes

A large number of differentially expressed genes is localized between conservative (AAAGA)₂ tracts of 87A7 locus (Figure 1). In the formation of the functional domain of *hsp70* genes about 30 kb long the short polypurine/polypyrimidine tracts as well as insulators (scs/scs'-elements), and their proteins (BEAF32, Zw5), may be involved. Perhaps, the “big” loop (60 kb) and “small” loop (30 kb) are the alternative states of chromatin in 87A7 locus, because polypurine/polypyrimidine tracts involved in the formation of both loops overlap with each other in the 5'-region of the locus.

It is not excluded that the insulators most fully perform the functional domain formation within these chromosomal domains and their “work” outside these domains may be impeded. This assumption is confirmed by our experimental data with double-gene transgenic system (*Drosophila yellow* and *white* genes). We have shown that neDNA fragments (EnvM4) when flanking the two reporter genes are able to protect them from the PEV in the host chromosomes of *D. melanogaster*. The maximum protective effect was observed in the presence of insulator Wari located in the 3'-region of *white* gene. When neDNA flank only one of the reporter genes (*white* + Wari), this very gene is protected against PEV to a greater extent than other gene (*yellow*) not flanked by neDNA fragments [13] [28].

Results of analysis for 87A7 locus loop chromatin organization can logically explain some experimental facts. Firstly, scs/scs'-elements are not able to protect the transgene from PEV when integrated into heterochromatin [29]. This may be due to the fact that transgenic systems are artificial, and may not contain the required set of chromosomal elements in the structure, ensuring the formation of a transgene independent expression domain. In addition, sites of integration are not native for the transgene and cannot contain, for example, complementary polypurine/polypyrimidine tracts necessary for the formation of the loop structures. Thus, according to the simulation for loop organization of chromatin locus 87A7 results, chromosomal binding sites with the nuclear envelope, short complementary polypurine/polypyrimidine tracts as well as scs/scs'-insulators and their proteins (Zw5, BEAF32)—the total of 3 elements may be involved in the formation of the independent expression domain of *hsp70* genes. In the case of experimental transgenic system (*yellow* and *white* genes), only 2 elements—chromosome to nuclear envelope attachment sites and insulator Wari—are involved in the formation of the independent transgenes' expression domain [13] [14] [28].

Secondly, scs/scs'-elements are not localized at the puff boundaries but inside it upon heat shock [25]. The results of our analysis indicate that the “small” loop (*hsp70* genes) is formed within the “big” expanded loop (puff??) upon activation of heat shock genes. This assumption is confirmed by the fact that (AAAGA)_n conservative tract was detected by FISH data at the band/interband borders on polytene chromosomes of *D. melanogaster* (unpublished data). The barrier function of scs/scs'-elements is also questionable because scs/scs'-elements overlap with promoters of neighboring genes: scs-element includes CG31211 gene promoter, scs'-element—promoters of CG3281 and *aurora* genes (FlyBase database). In particular, BEAF32 (insulator protein) and DREF

(transcriptional activator) proteins have a lot of interaction sites on the chromosome of *Drosophila*, and approximately in 50% of cases they overlap [30] [31]. Their competitive relationships for the DNA binding sites have been shown as well [27]. Perhaps, the overlapping of insulator protein and transcriptional activator protein sites of the adjacent genes is one of the functional mechanisms of gene *hsp70* isolation. The mechanism might be based on the relationship between the insulator protein and the transcription activator protein in the same way as “repressor-activator” attitude.

5. Conclusion

On the model of *D. melanogaster* chromosomal 87A7 locus, it was demonstrated that various elements could be involved in the structural and functional organization of chromosomes. They are short polypurine/polypyrimidine tracts, insulators and their proteins, regions that attach chromosomes to the nuclear envelope, *i.e.* not only DNA-protein but DNA-DNA interactions as well. Combinatorics of the interaction of these elements can determine alternative states of chromosomal locus and genes belonging to this locus. Moreover, the same structural and functional status of chromosomal locus can be accomplished by several embodiments of DNA-DNA interactions that may underlie self-regulation of the locus and in a wider sense—“stability” of the genetic system.

Support

This work was supported by a grant from Subprogramme “Gene pools of wildlife and its preservation” of Presidium RAS Programme for Basic Research “Biodiversity of natural systems”.

References

- [1] Dekker, J., Rippe, K., Dekker, M. and Kleckner, N. (2002) Capturing Chromosome Conformation. *Science*, **295**, 1306-1311. <https://doi.org/10.1126/science.1067799>
- [2] Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S. and Ren, B. (2012) Topological Domains in Mammalian Genomes Identified by Analysis of Chromatin Interactions. *Nature*, **485**, 376-380. <https://doi.org/10.1038/nature11082>
- [3] Sexton, T., Yaffe, E., Kenigsberg, E., Bantignies, F., Leblanc, D., Hoichman, M., Parrinello, H., Tanay, A. and Cavalli, G. (2012) Three-Dimensional Folding and Functional Organization Principles of the *Drosophila* Genome. *Cell*, **148**, 458-472. <https://doi.org/10.1016/j.cell.2012.01.010>
- [4] Rao, S.S.P., Huntley, M.H., Durand, N.C., Stamenova, E.K., Bochkov, I.D., Robinson, J.T., Sanborn, A.L., Machol, I., Omer, A.D., Lander, E.S. and Lieberman, A.E. (2014) A 3D Map of the Resolution Reveals Principles of Chromatin Looping. *Cell*, **159**, 1665-1680. <https://doi.org/10.1016/j.cell.2014.11.021>
- [5] Dekker, J. and Heard, E. (2015) Structural and Functional Diversity of Topologically Associating Domains. *FEBS Letters*, **589**, 2877-2884. <https://doi.org/10.1016/j.febslet.2015.08.044>
- [6] Nora, E.P., Lajoie, B.R., Shulz, E.G., Giorgetti, I., Okamoto, I., Servant, N., Piolot, T., van Berkum, N.L., Meisig, J., Sedat, J., Gribnau, J., Barillot, E., Bluthgen, N., Dekker, J. and

- Heard, E. (2012) Spatial Portioning of the Regulatory Landscape of the X-Inactivation Centre. *Nature*, **485**, 381-385. <https://doi.org/10.1038/nature11049>
- [7] Glazkov, M.V. (1999) Compactization of Chromosomal Loci of Eukaryotic Genes: Triplex DNA Structures (a Model). *Molecular Biology*, **33**, 511-515.
- [8] Glazkov, M.V. (2011) Loop Organization of Eukaryotic Chromosomes and Triple-Stranded DNA Structures. *Molecular Biology*, **45**, 263-274. <https://doi.org/10.1134/S0026893310061020>
- [9] Hampel, K.J., Ashley, C. and Lee, J.S. (1994) Kilobase-Range Communication between Polypurine-polypyrimidine Tracts in Linear Plasmids Mediated by Triplex Formation: A Braided Knot between Two Linear Duplexes. *Biochemistry*, **32**, 5674-5681. <https://doi.org/10.1021/bi00185a002>
- [10] Lee, J.S., Latimer, L.J.P., Haug, B.L., Pulleyblank, D.E., Skinner, D.M. and Burkholder, G.D. (1989) Triplex DNA in Plasmids and Chromosomes. *Gene*, **82**, 191-199. [https://doi.org/10.1016/0378-1119\(89\)90044-9](https://doi.org/10.1016/0378-1119(89)90044-9)
- [11] Lee, J.S., Ashley, C., Hampel, K.J., Bradley, R. and Scraba, D.G. (1995) A Stable Interaction between Separated Pyrimidine Purine Tracts in Circular DNA. *Journal of Molecular Biology*, **252**, 283-288.
- [12] Chudinov, O.S. and Glazkov, M.V. (2002) *In Vitro* Formation of Triple-Stranded DNA Structures in the Human $\alpha 1$ -Antitrypsin Gene. *Molecular Biology*, **36**, 100-104. <https://doi.org/10.1023/A:1014262926878>
- [13] Shabarina, A.N., Glazkov, M.V. and Shostak, N.G. (2010) The Role of Chromosomal Regions Anchored to the Nuclear Envelope in the Functional Organization of Chromosomes. *Russian Journal of Genetics*, **46**, 1042-1044. <https://doi.org/10.1134/S1022795410090061>
- [14] Shabarina, A.N. and Glazkov, M.V. (2012) Nuclear Envelope Attachment Sites of Interphase Chromosomes: Barrier Elements but Not Insulators. *Russian Journal of Genetics*, **48**, 864-867. <https://doi.org/10.1134/S1022795412080066>
- [15] Glazkov, M.V., Drozd, S.F. and Poltarau, A.B. (1999) Comparative Analysis of Nucleotide Sequences of Chromosomal DNA Isolated from Nuclear Envelopes and Cores of Rosette-Like Structures of Murine Interphase Chromosomes. *Russian Journal of Genetics*, **35**, 208-213.
- [16] Shabarina, A.N., Prilepa, E.I. and Glazkov, M.V. (2006) Unusual Nucleotide Sequence of a DNA Fragment Isolated from Nuclear Envelopes of Mouse Hepatocytes. *Russian Journal of Genetics*, **42**, 715-722. <https://doi.org/10.1134/S1022795406070027>
- [17] Onishchenko, G.E and Chentsov, Iu.S. (1974) Granular Layer of the Peripheral Chromatin in the Interphase Nucleus. I. Ultrastructure. *Tsitologiya*, **16**, 675-678.
- [18] Tsanev, R. and Tsaneva, I. (1986) Molecular Organization of Chromatin as Revealed by Electron Microscopy. *Methods and Achievements in Experimental Pathology*, **12**, 63-104.
- [19] Zullo, J.M., Demarco, I.A., Piqué-Regi, R., Gaffney, D.J., Epstein, C.B., Spooner, C.J., Luperchio, T.R., Bernstein, B.E., Pritchard, J.K., Reddy, K.L. and Singh, H. (2012) DNA Sequence-Dependent Compartmentalization and Silencing of Chromatin at the Nuclear Lamina. *Cell*, **149**, 1474-1487. <https://doi.org/10.1016/j.cell.2012.04.035>
- [20] Naumova, N., Imakaev, M., Fudenberg, G., Zhan, Y., Laioie, B.R., Mirny, L.A. and Dekker, J. (2013) Organization of the Mitotic Chromosome. *Science*, **342**, 948-953. <https://doi.org/10.1126/science.1236083>
- [21] Udvardy, A., Maine, F. and Schedl, P. (1985) The 87A7 Chromomere. Identification of Novel Chromatin Structures Flanking the Heat Shock Locus That May Define the Boundaries of Higher Order Domains. *Journal of Molecular Biology*, **185**, 341-358.

- [22] Moreau, J., Matyash-Smirniaguina, L. and Scherrer, K. (1981) Systematic Punctuation of Eukaryotic DNA by AT-Rich Sequences. *Proceedings of the National Academy of Sciences of the United States of America*, **78**, 1341-1345. <https://doi.org/10.1073/pnas.78.3.1341>
- [23] Comings, D.E. and Wallack, A.S. (1978) DNA-Binding Properties of Nuclear Matrix Proteins. *Journal of Cell Science*, **38**, 233-246.
- [24] Meuleman, W., Peric-Hupkes, D., Kind, J., Beaudry, J.-B., Pagie, L., Kellis, M., Reinders, M., Wessels, L. and van Steensel, B. (2013) Constitutive Nuclear Lamina-Genome Interactions are Highly Conserved and Associated with A/T-Rich Sequence. *Genome Research*, **23**, 270-280. <http://www.genome.org/cgi/doi/10.1101/gr.141028.112>
<https://doi.org/10.1101/gr.141028.112>
- [25] Kuhn, E.J., Hart, C.M. and Geyer, P.K. (2004) Studies of the Role of Drosophila Scs and Scs' Insulators in Defining Boundaries of a Chromosome Puff. *Molecular and Cellular Biology*, **24**, 1470-1480.
- [26] Blanton, J., Gaszner, M. and Schedl, P. (2003) Protein:Protein Interactions and the Pairing of Boundary Elements *in Vivo*. *Genes & Development*, **17**, 664-675.
- [27] Hart, C.M., Cuvier, O. and Laemmli, U.K. (1999) Evidence for an Antagonistic Relationship between the Boundary Element-Associated Factor BEAF and the Transcription Factor DREF. *Chromosoma*, **108**, 375-383. <https://doi.org/10.1007/s004120050389>
- [28] Shabarina, A.N. and Glazkov, M.V. (2013) Barrier Elements of Chromatin Domains and Nuclear Envelope. *Russian Journal of Genetics*, **49**, 23-28. <https://doi.org/10.1134/S1022795413010122>
- [29] Kellum, R. and Schedl, P. (1991) A Position-Effect Assay for Boundaries of Higher Order Chromosomal Domains. *Cell*, **64**, 941-950. [https://doi.org/10.1016/0092-8674\(91\)90318-S](https://doi.org/10.1016/0092-8674(91)90318-S)
- [30] Jiang, N., Emberly, E., Cuvier, O. and Hart, C.M. (2009) Genome-Wide Mapping of Boundary Element-Associated Factor (BEAF) Binding Sites in Drosophila melanogaster Links BEAF to Transcription. *Molecular and Cellular Biology*, **29**, 3556-3568.
- [31] Negre, N., Brown, C.D., Shah, P.K., Kheradpour, P., Morrison, C.A., Henikoff, J.G., Feng, X., Ahmad, K., Russell, S., White, R.A.H., Stein, L., Henikoff, S., Kellis, M. and White, K.P. (2010) Comprehensive Map of Insulator Elements for the Drosophila Genome. *PLoS Genetics*, **6**, e1000814. <https://doi.org/10.1371/journal.pgen.1000814>

List of Abbreviations

TADs—Topologically Associated Domains, neDNA—chromosomal DNA fragments that attach chromosomes to the nuclear envelope, LINE—Long Interspersed Repeat Sequences, SINE—Shot Interspersed Repeat Sequences, LADs—Lamina Associated Domains, scs/scs'—specialized chromatin structures, FISH—Fluorescence in situ Hybridization.



Submit or recommend next manuscript to SCIRP and we will provide best service for you:

Accepting pre-submission inquiries through Email, Facebook, LinkedIn, Twitter, etc.

A wide selection of journals (inclusive of 9 subjects, more than 200 journals)

Providing 24-hour high-quality service

User-friendly online submission system

Fair and swift peer-review system

Efficient typesetting and proofreading procedure

Display of the result of downloads and visits, as well as the number of cited articles

Maximum dissemination of your research work

Submit your manuscript at: <http://papersubmission.scirp.org/>

Or contact cmb@scirp.org