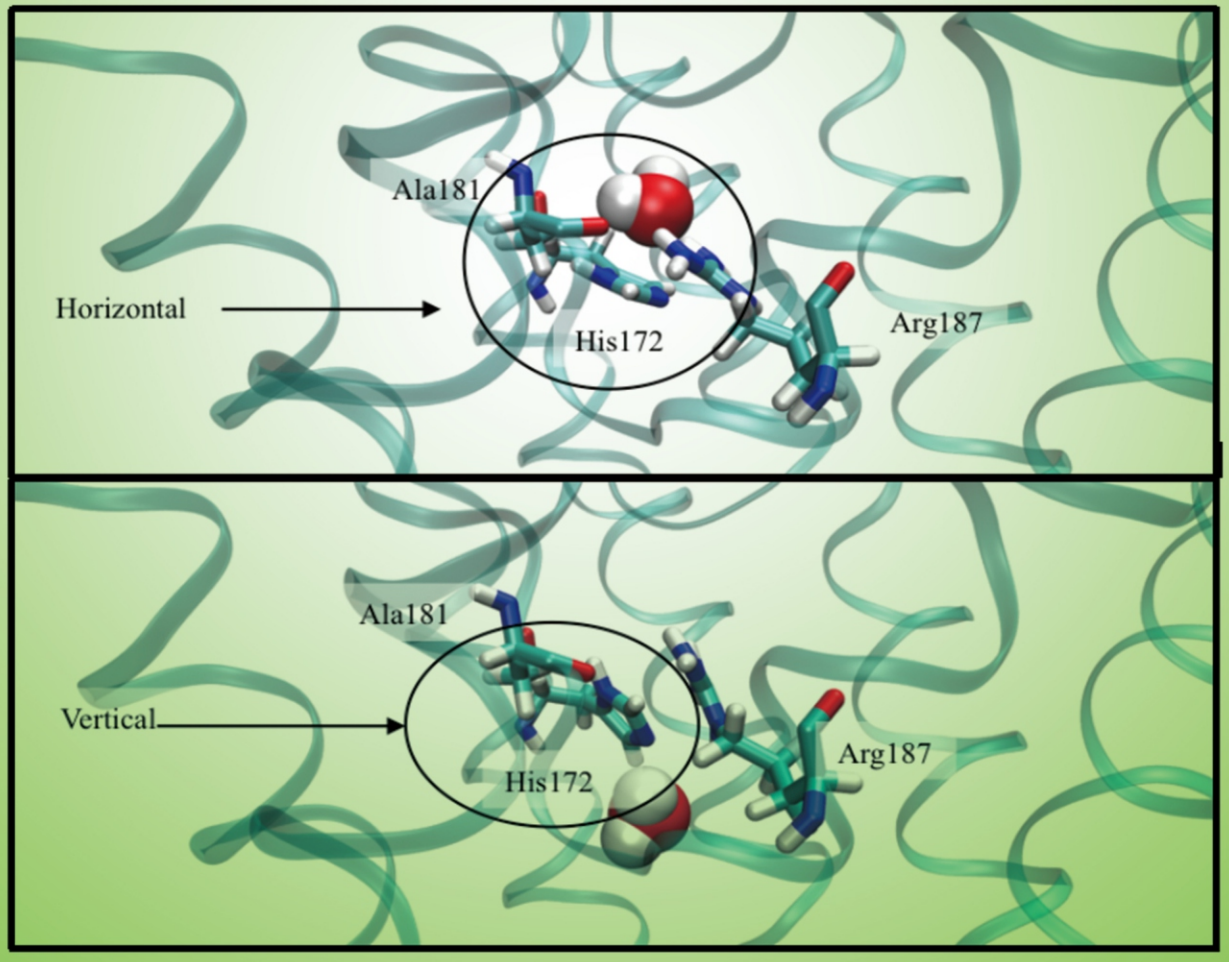
ISSN: 2165-3445 Volume 6, Number 4, December 2016

Scientific Research Publishing

Computational Molecular Bioscience



ISSN: 2165-3445

www.scirp.org/journal/cmb

Journal Editorial Board

ISSN: 2165-3445 (Print) ISSN: 2165-3453 (Online)

http://www.scirp.org/journal/cmb

Editor-in-Chief

Editorial Board

Scientific Research Publishing

Computational Molecular Bioscience, 2016, 6, 53-79
http://www.scirp.org/journal/cmb
ISSN Online: 2165-3453
ISSN Print: 2165-3445

Table of Contents

Volume 6 Number 4 December 2016

The figure on the front cover is from the article published in Computational Molecular Bioscience, 2016, Vol. 6, No. 4, pp. 66-79 by L. K. Balcziak, T. H. Bach, L. R. Montgomery, M. E. Warwick and U. Akgun.

Computational Molecular Bioscience (CMB)
Journal Information

SUBSCRIPTIONS

The *Computational Molecular Bioscience* (Online at Scientific Research Publishing, www.SciRP.org) is published quarterly by Scientific Research Publishing, Inc., USA.

Subscription rates:
Print: $79 per issue.
To subscribe, please contact Journals Subscriptions Department, E-mail: sub@scirp.org

SERVICES

Advertisements
Advertisement Sales Department, E-mail: service@scirp.org

Reprints (minimum quantity 100 copies)
Reprints Co-ordinator, Scientific Research Publishing, Inc., USA.
E-mail: sub@scirp.org

COPYRIGHT

PRODUCTION INFORMATION

For manuscripts that have been accepted for publication, please contact:
E-mail: cmb@scirp.org

Computational Molecular Bioscience, 2016, 6, 53-65
http://www.scirp.org/journal/cmb
ISSN Online: 2165-3453
ISSN Print: 2165-3445

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

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 *hsp*70*Aa* and *hsp*70*Ab*, 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

*hsp*70 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 recombinogenity (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 (compactisied) 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 (*hsp*70) locus. These genes encode proteins of *hsp*70 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 *hsp*70*Aa* and *hsp*70*Ab* 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 *hsp*70 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 *hsp*70 genes/one loop for the whole locus and black rectangles denote the bases of these loop structures respectively.

The insulators of *hsp*70 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).

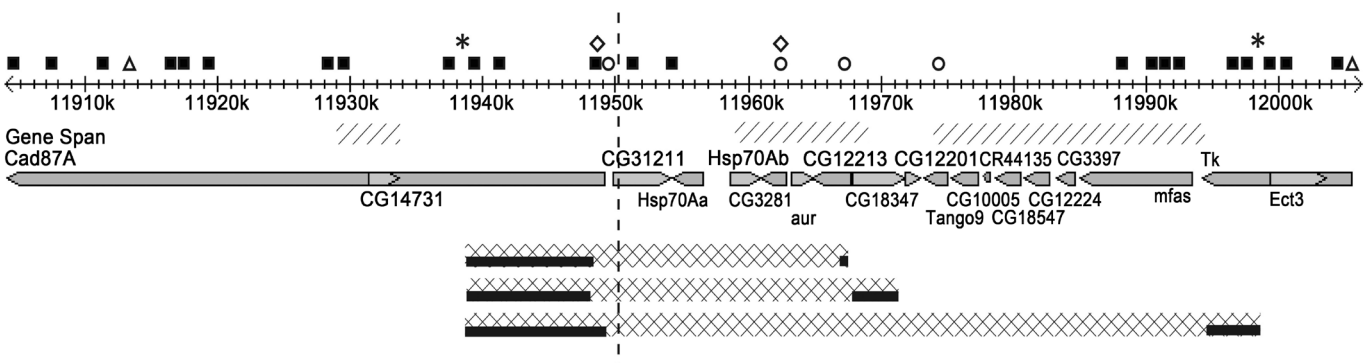


Figure 1. Schematic representation of *D. melanogaster hsp*70 genes locus.

Table 1. Sequence analysis results for the *D. melanogaster hsp*70 genes locus (3R: 11,904,163 .. 12,006,544; 102382 bp).

Insulators	Insulator class_II.mE01 (flybase.org)	FBsf0000154391 (11,913,012 .. 11,913,022)
		FBsf0000154392 (12,005,260 .. 12,005,270)
	Insulator class_I.mE01 (flybase.org)	FBsf0000158034 (11,949,333 .. 11,949,343)
		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
$(GAGA)_2$ tracts		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 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 *hsp*70 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. *hsp*70 Genes Can Be Localized in The Loop Structures of Small Size

Genes *hsp*70*Aa* and *hsp*70*Ab* 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, *hsp*70*Aa*, *hsp*70*Ab*, 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, *hsp*70*Aa*, *hsp*70*Ab*, 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 *hsp*70*Aa* and *hsp*70*Ab*) 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* polythene chromosomes using FISH-analysis (unpublished data).

3.4. A Possible Mechanism for *hsp*70 Gene Independent Expression Domain Formation

It can be assumed that the formation of a domain of active *hsp*70 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, *hsp*70*Aa*, *hsp*70*Ab*, CG3281, and *aurora* genes are found within the loop. When forming the second loop (see above), CG31211, *hsp*70*Aa*, *hsp*70*Ab*, 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

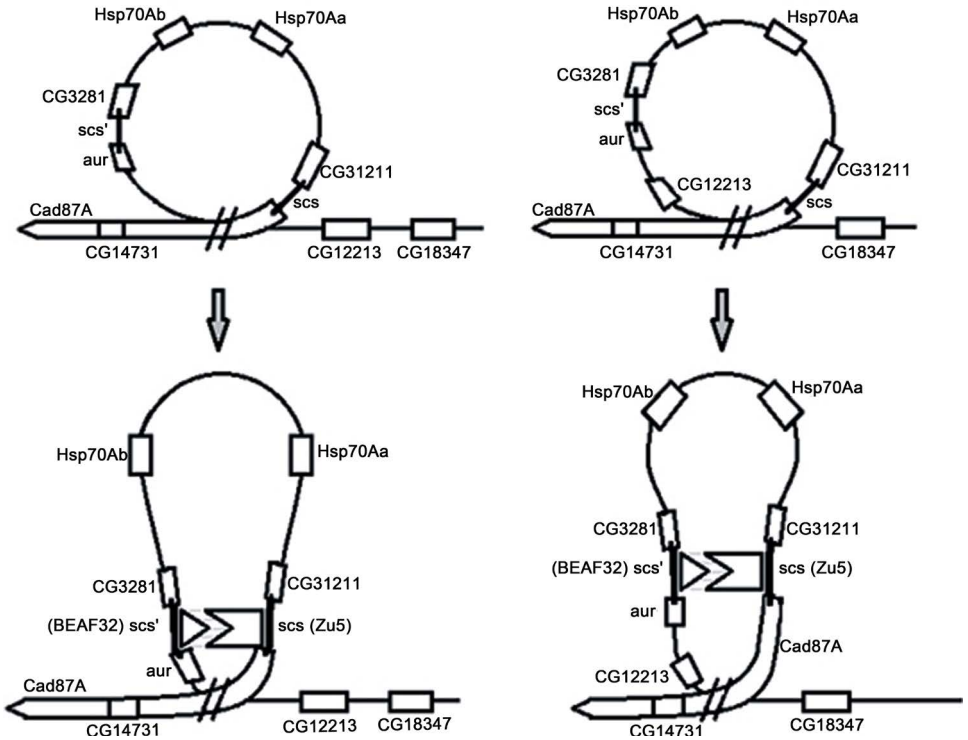


Figure 2. Schematic representation of the two variants of loop structures for *hsp*70 genes that can be formed with the participation of triple-stranded DNA.

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 *hsp*70*Aa* and *hsp*70*Ab* 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 *hsp*70*Aa* and *hsp*70*Ab* 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 *hsp*70*Aa* and *hsp*70*Ab* 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 *hsp*70*Aa* and *hsp*70*Ab* 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 *hsp*70 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 polythene 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 *hsp*70 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 *hsp*70 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 *hsp*70 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 *hsp*70 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 (*hsp*70 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 polythene 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 *hsp*70 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

[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 *a*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 Poltaraus, 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. *Tsitologiia*, **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

[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.

Scientific Research Publishing

Computational Molecular Bioscience, 2016, 6, 66-79
http://www.scirp.org/journal/cmb
ISSN Online: 2165-3453
ISSN Print: 2165-3445

Calmodulin Bound Aquaporin-0 Reveals Two Distinct Energy Profiles

L. K. Balcziak*, T. H. Bach*, L. R. Montgomery, M. E. Warwick, U. Akgun

Physics Department, Coe College, Cedar Rapids, IA, USA
Email: uakgun@coe.edu

How to cite this paper: Balcziak, L.K., Bach, T.H., Montgomery, L.R., Warwick, M.E. and Akgun, U. (2016) Calmodulin Bound Aquaporin-0 Reveals Two Distinct Energy Profiles. *Computational Molecular Bioscience*, **6**, 66-79.
http://dx.doi.org/10.4236/cmb.2016.64006

Received: October 31, 2016
Accepted: December 17, 2016
Published: December 20, 2016

Abstract

Aquaporin-0 (AQP0) contributes to the nurturing and cleaning of the eye lens of waste products. It is a tetrameric protein composed of four identical monomers, each of which has its own water pore. AQP0 water conduction is regulated by pH, Ca^{2+} concentration, and the phosphorylation of serine residues at the C-terminal. High cellular Ca^{2+} concentration enhances the binding of Calmodulin (CaM), a Ca^{2+} dependent protein, to AQP0 from cytoplasm. This study focuses on determining the differences between the AQP0-CaM and the open AQP0 systems, by using Molecular Dynamics (MD) methods. The water conduction energy profiles are measured with two separate MD simulation techniques revealed two distinct channel profiles for the AQP0-CaM combined model. While the CaM bound channels' energy barriers exceed the 6 kcal/mol, the no CaM bound AQP0 energy profile stays below 3 kcal/mol. The structural analysis of these different pores during the free equilibrations also supported this conclusion with distinct pore diameters. Unlike the previous report, this study observed Phe75 and His66 taking role in stabilizing the CSII restriction site in CaM bound AQP0.

Keywords

Aquaporin-0, Calmodulin, Water Channel, MD Simulations

1. Introduction

When Peter Agre discovered the first water-conducting channel in human red blood cells, denoted CHIP28 at the time of identification, genomic sequencing revealed that it belonged to the MIP family, which derives its name from the Major Intrinsic Protein, found in lens fiber cell [1] [2]. Once the water-channel function of CHIP28 was determined, the MIP family was renamed the aquaporin family; CHIP28 was renamed Aq-

*These authors contributed equally to this work.

uaporin-1, and MIP was renamed AQP0 [2]. Proteins homologous to these water channels are found in all domains of life, which have since been grouped into approximately 30 subfamilies through phylogenetic analysis [3]. In humans, there are 13 isoforms (AQP0-AQP12), expressed in a tissue-specific manner [1] [2] [4] [5] [6]. The properties of each aquaporin (selectivity and permeability) correspond to its molecular form.

The general structure for an aquaporin is a homotetramer. Each monomer is made up of 6 alpha-helices that form a barrel-shaped pore, 3 extracellular loops (A, C, and E), 2 intracellular loops (B and D), and cytoplasmic C and N termini [2] [4]. Two restriction sites regulate the channel. It selects against hydronium ions with highly conserved NPA motifs (Asn-Pro-Ala) located on loops B and E, which form a constriction [4] [7] [8] [9] [10]. The pore restricts larger solutes through the ar/R constriction site, also known as construction site I, composed of four residues including arginine and aromatic amino acids. The pore diameter at the ar/R constriction site creates an opening 3 Å wide in water-selective aquaporins, allowing water in, but excluding larger solutes [7] [8] [11] [12].

Out of the 13 human isoforms of aquaporin channels, 9 are expressed in the human eye [13]. In the lens fiber, there is a unique isoform that is of special interest. AQP0, which is expressed solely in the mammalian lens fiber, accounts for more than 60 percent by weight of its plasma membrane proteins. Functioning as a junction as well as an aquaporin, AQP0 is crucial for maintenance of the structural integrity of cortical fiber cells, resulting in lens transparency, allowing for light to be focused on the retina.

There are two forms of AQP0 present in a lens. The open configuration functions as a water channel, contributing to the microcirculation of water that nurtures and cleans the lens of waste products [14] [15] [16] [17]. As the fiber cell matures, AQP0 is modified into the closed configuration. The C-terminal is cleaved through post-transcriptional modification, which causes two AQP0 links to form an intercellular adhesion junction called a thin lens junction [2] [10] [18]-[22]. These modifications include phosphorylation at the calmodulin-binding domain [23]. While the absence of AQP0 from the lens does not prevent interlocking protrusions of young fiber cells, it did cause the loss of function of these protrusions to maintain the structural integrity of the fiber cells, leading to cell separation and cataract formation [13]. A number of mutations in AQP0 contribute to congenital cataracts in humans and mice by interfering with the protein's ability to maintain osmotic balance [24] [25] [26] [27]. Hence, determining the structure of AQP0 has clinical relevance to cataract formation.

The water permeability of AQP0 is 40 times lower compared to AQP1 [10] [28]. This has been attributed to two unique tyrosine residues (Tyr149 and Tyr23). The NPA site in AQP0 is also narrower than in AQP1, resulting from a substitution of Leu151 to Phe141 in loop B [10]. AQP0 channels are regulated by external pH [28] [29] [30], external and internal concentration of calcium ions (Ca^{2+}) [28] [29] [30], and phosphorylation. Protonation and deprotonation of His40 has been shown to be caused by the pH regulation of AQP0 [3]. The optimal pH for AQP0 water conduction is at 6.5: an in-

crease in pH from 6.5 to 7.2 decreases water permeability by half [28] [29] [30]. Calmodulin is a universal regulatory protein that mediates the effects of Ca^{2+} concentration on the water permeability of AQP0 through allosteric modulation [28]-[36]. Calmodulin (CaM) is a small protein (16.7 kDa) with a C-lobe and an N-lobe connected by a flexible linkage [37] [38] [39] [40]. The binding of four Ca^{2+} ions to CaM exposes hydrophobic regions that can interact with hydrophobic regions of target proteins [33] [34] [35]. Recent research on the pseudoatomic model suggests that one CaM can cooperatively bind non-canonically to two open configuration AQP0 monomers and allosterically inhibit the channel by narrowing constriction site II (CSII), a feature absent in other AQPs [32] [36]. CSII is a located on the cytoplasmic half of the pore and involves Tyr149 extending into the pore and interacting with His66 and Phe75 [10]. The pH and Ca^{2+} concentration vary inside the layers of the lens, so regulation by these properties may be physiologically significant [3]. The modifications that affect responsiveness of the fiber cell include phosphorylation at the CaM-binding domain [23]. An anchoring protein, AKAP2 positions protein kinase A, which phosphorylates AQP0 at Ser235 in the CaM binding site in order to prevent CaM modulation. This may be done in order to preserve fluid circulation in the middle of the lens [23] [41]. Another study found when either Ser235 or Ser231 was phosphorylated, as it is in posttranslational modification, binding affinity is reduced for dansyl-CaM in a bovine model [34].

Previous dynamic simulations suggested that CaM binding to an AQP0 tetramer reduced the dynamics of all AQP0 channels [36]. Permeability assays carried out on oocytes expressing wild type AQP0 showed that when cytoplasmic Ca^{2+} increased from 0 to 1.8 mM, water permeability decreases by a factor of 0.4 [28]-[36]. Another study showed that calcium and protons can regulate AQP0 via single-monomer regulation, as well as in a cooperative fashion. The non-canonical double binding of CaM suggests that the regulation of Ca^{2+} is through cooperative modulation: the CaM molecule linking the two AQP0 bound monomers to each other [42]. This leads to the question of whether the dynamic restriction of AQP0 is the only mechanism through which CaM-Ca^{2+} regulates AQP0's permeability to water. In this study, we applied various Molecular Dynamics Simulation techniques to investigate the structural and energetic variations among the AQP0 monomers with and without CaM binding.

2. Methods

Molecular visualization programs VMD 1.9.1 [43] and CHARMM36 [44] were used for modeling and visualization. MD Simulations were performed via NAMD2.0 [45], on local CPU and GPU clusters, the University of Iowa HPC facility, and the University of Texas TACC. Two separate tetrameric models were built with free-AQP0: (PDB ID: 2B6P [22]) and CaM-bound AQP0 (PDB ID: 3J41 [42]). In each model, hydrogen atoms were added to the backbone of an AQP0 tetramer by VMD *psfgen*. Then the AQP0 tetramer was embedded in 1-palmitoyl-2-oleoyl-sn-glycerol-3-phosphocholine (POPC) bilayer of 150 Å × 150 Å by VMD *membrane builder*, solvated in TIP3 water molecules using VMD *solvate* and neutralized with NaCl by VMD *autoionize*. The final CaM bound

AQP0 model had 386,322 atoms, and the open AQP0 simulation model has 284,903 atoms. Both systems were minimized for 2000 steps, and equilibrated for 30 ns before the production runs.

Steered molecular dynamic simulation (SMD) was used to study water conduction energy profile, transport dynamics, as well as determination of the critical residues in the AQP0 pathway. An extracellular water molecule approximately 5 Å away from the opening was aligned with the center z-axis of each channel. In other words, SMD was conducted with four water molecules in a free-AQP0 or a CaM-bound AQP0 model. The water molecules were pulled along the z-axis with a force constant of 1.5 kcal/mol/Å^2 and a constant velocity of 5.0 Å/ns to AQP0 pore. All of the simulations were carried out in a periodic cell of 150 × 150 × 170 Å, at 310 K, and 1 atm. Z-coordinate and force were recorded every 100 fs. The center of mass of a water molecule was defined as the center of mass of the oxygen atom. Instantaneous force was calculated using the equation: $F = z_i + vt - z$, where z_i is initial position of the center of mass, v is pulling velocity, t is time and z if position of the center of mass at the point of measurement. Work is the product of the average force between two consecutive data point and the distance between these two data point. The energy profile of each water molecule was calculated based on Jarzynsky's equality [46]: $\exp(-\beta W) = \exp(-\beta \Delta G)$, where $\beta = 1/k_B T$, k_B is the Boltzmann constant and T is temperature. In addition, RMSD was used to keep track of the dynamics of critical residues at constriction site I and constriction site II, especially Tyr149. VMD was used to study water conduction pathway and conformational changes of critical residues.

Umbrella sampling was used to investigate the energy profile of a water molecule in an AQP0 channel. The simulation was carried out for 4 channels simultaneously with a simulation window size of 0.5 Å. First, a water molecule approximately 5 Å directly above each channel was fixed using VMD *mdff*, and then the systems were minimized for 2000 steps. After that, these water molecules were constrained in z-coordinate, the coordinate along the longitude of the channel, with a force constant of 10 (kcal/mol/Å^2). The simulation was conducted at 310 K and 1 atm for 450 ps with data recorded every 500 fs, which resulted in 900 data points. The first 300 data points were used for equilibration and the next 600 data points were used for analysis. The initial position of the water molecule was defined as the fixed position of oxygen during minimization.

HOLE [47] was used to analyze the channels of AQP0 to confirm the results of the free energy profiles obtained with SMD and Umbrella sampling. The HOLE software creates a visual output and coordinates of the radius of the channel for each monomer. A HOLE profile and visualization was created for each individual monomer, one for the open conformation and another for the closed conformation after over 50 ns of MD simulations. In the end 8 outputs were created. This allowed for the comparison of the channel dynamics when CaM was bound versus when the channel was in an open conformation. Comparisons of the HOLE pathways and SMD trajectories were created by superimposing the channels with their alpha-carbons onto one another using the RMSD tool on VMD for alignment.

RMSD values of the side chains were gathered from the VMD software *RMSD Trajectory Tool*. Data acquired from equilibration of the two models was used for the analysis. Both of the models were equilibrated and had 200 usable DCD frames over the span of 30 nanoseconds per model.

3. Results and Discussion

3.1. AQP0-Cam System Has Two Distinct Water Conduction Energy Profiles

The various MD simulations applied to CaM bound and unbound AQP0 tetramers revealed three unique energy profiles for water conduction. The first representing the four channels in free-AQP0, the second representing the two channels that were not covered by CaM in CaM-bound AQP0, and the final energy profile representing the two channels that were bound by CaM. For convenience, these three energy profiles are denoted open, closed-open, and closed-closed, respectively. Each of the energy profiles has three prominent peaks.

Figure 1 highlights both models used in the simulations. Figure 1—LEFT is the

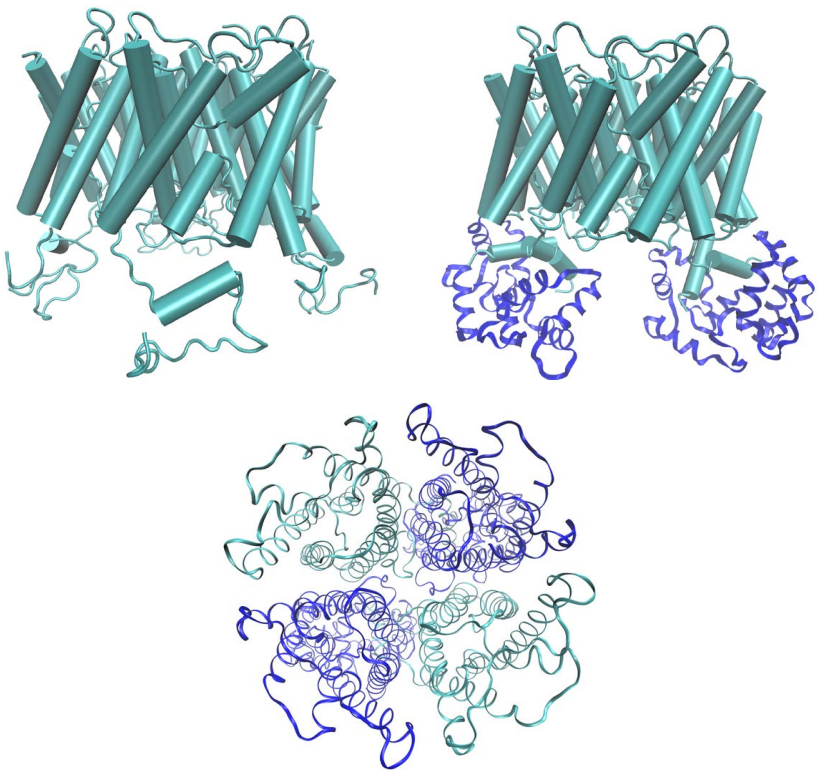


Figure 1. (Top Left) AQP0 tetramer in the open conformation with no CaM attached to the protein. (Top Right) AQP0 tetramer bound to two CaM. In the closed configuration each CaM is aligned with one of the AQP0 pores. In order to simplify the image, the lipids and water molecules are not shown. (Bottom) Top down perspective of the closed configuration with CaM. In this study, the monomers in dark blue are named as closed-closed, and those in cyan are closed-open.

AQP0 tetramer in the open conformation without CaM. Figure 1—RIGHT shows the two CaM units are bound to the bottom portion of the tetramer. CaM is highlighted in dark blue ribbon below the tetramer, facing outwards towards the intracellular region.

The radius of the open monomer is always at a value above 1 Å (see Figure 2). In contrast with the closed-open and the closed-closed monomers, the channel is very clear and permeable to water. The only slight point of constriction is along Z coordinate 3.53 where the radius dips down to a value of 1.001 angstroms for a short section. This corresponds to a group of residues that constitute the ar/R site (includes His172, Ala181, and Arg187). The degree of constriction along the initial portion of the channel depends on how close His172 is to Ala181 and Arg187. His172 is in a vertical conformation with the channel, allowing water to pass through. Further down the channel, the radius near CSII is at 1.30 angstroms. The residues Tyr149, Phe75, and His66 provide enough room for passage of water molecules to occur.

Next we investigate the changes that occur to the tetramer when two CaM molecules bind to it. The combined use of SMD trajectories and HOLE visuals shows that the path the water molecules take doesn't change for any of the four monomers when CaM is bound to AQP0 versus when the tetramer is in the open conformation. All of the HOLE

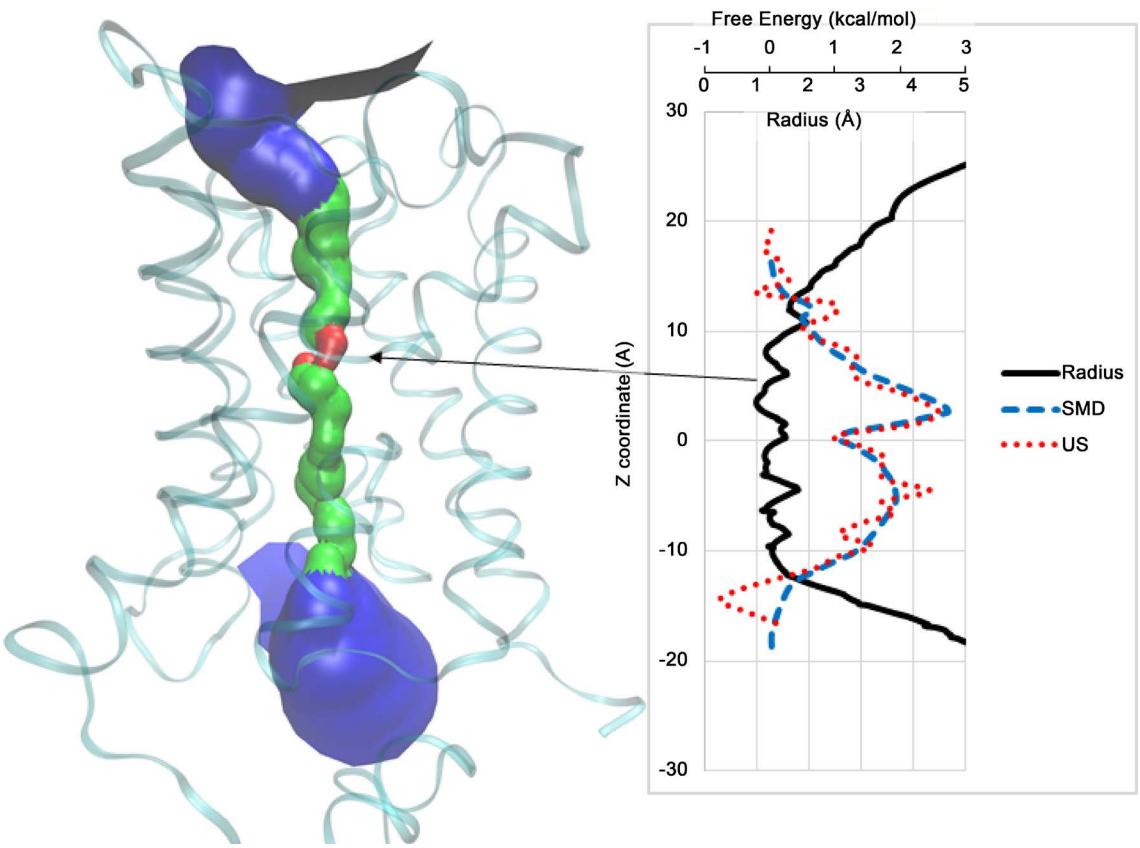


Figure 2. Measured energy profiles for water conduction, and the pore radius calculations (Right) along with the HOLE visualization of the open conformation of an AQP0 monomer with no CaM bound to the system (Left). The black solid line represents the pore radius. The blue-dashed and red-dotted lines represent SMD and umbrella sampling free energy calculations along the pore, respectively.

plots reveal the same pathway. However, the overall conductivity decreases due to the smaller channel radius and less total channel volume. This decrease in conductivity occurs at the second constriction site (CSII), as well as the ar/R site.

Figure 3 shows the major points of constriction along the monomer when it is bound directly to CaM. The radius decreases significantly at several points along the channel. HOLE analysis and radius data give the radius at CSII as only 0.97 Å. At this point in the channel residues Try149, Phe75, and His66 all come together in close proximity to restrict the passage of water. However, the most significant point of constriction in the channel occurs near residues Ala181, Arg187, His172, and Met183. The channel radius decreases to 0.665 Angstroms. This small radius is partly due to His172 being in a horizontal conformation, and the constriction of the ar/R site.

Since only half of the channel is directly bound to CaM, we wanted to see whether or not there are any allosteric changes to the shapes of the two closed-open monomers. To date, there hasn't been much insight into whether or not the closed-open monomers differ from the closed-closed monomers. Our investigation shows that the closed-open monomers are very restrictive to the passage of water in both the SMD simulations as well as equilibration of the protein along the initial portion of the channel. When comparing between the closed-open and the open conformation, the data makes it clear that there is less permeability in the closed-open, with a higher energy barrier existing along the ar/R site in the closed-open conformation.

Analysis of the closed-open conformation of AQP0 reveals a distinct contrast with

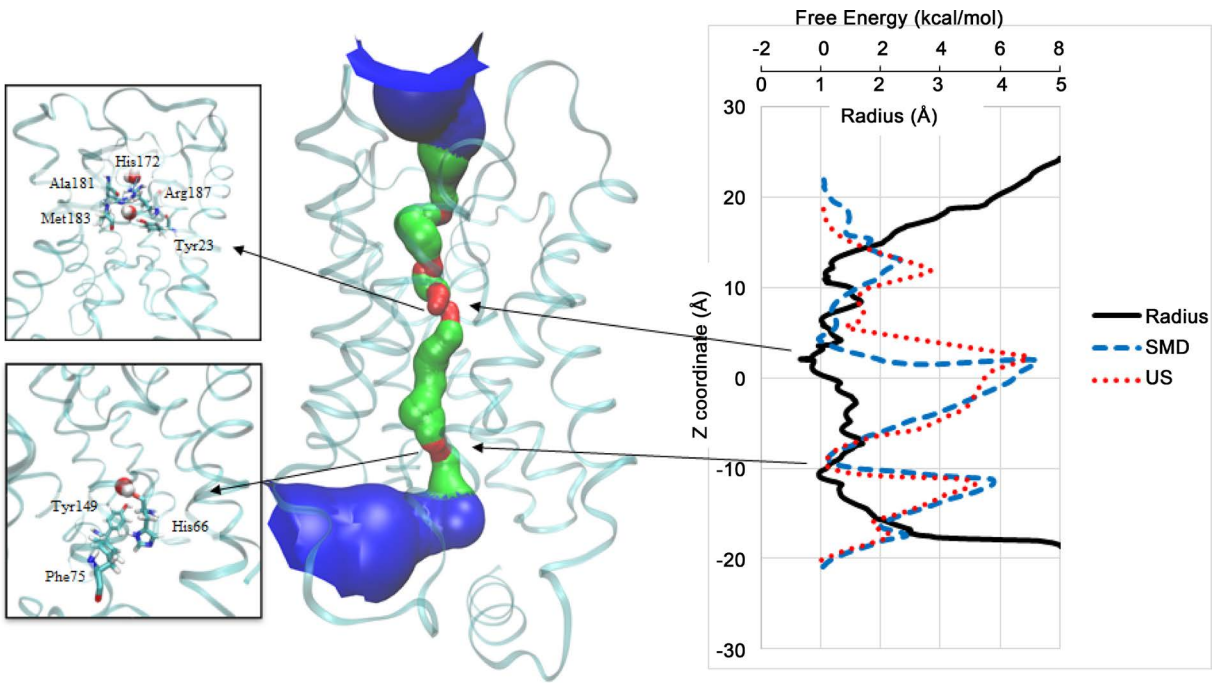


Figure 3. Measured energy profiles for water conduction, and the pore radius calculations (Right) along with the HOLE visualization of the closed-closed conformation of an AQP0 monomer (Middle). The ar/R and CSII residues orientations are shown in the boxes (Left). The black solid line represents the pore radius. The blue-dashed and red-dotted lines represent SMD and umbrella sampling free energy calculations along the pore, respectively.

the open conformation. We predicted the closed-open channel would be similar to the open channel, and it is more permeable at CSII when compared to the closed-closed conformation. Figure 4 shows that the energy barrier around residues His66, Phe75, and Tyr149 is around 2.5 kcal per mole, in comparison to the 5 kcals per mole the closed-closed conformation had at CSII. So it appears that there may indeed be a lesser degree of allosteric modulation that occurs at CSII in the closed-open conformation. Since it is not directly covered by calmodulin, there may be less strain on the protein on the two closed-open conformations, and the residues at CSII may not be brought into as close of proximity as the closed-closed monomers bound directly to CaM.

During SMD trajectories, it has been noted that His172 serves as a sort of "plug" when it is oriented horizontally. The water molecule spends a large portion of time above these residues until the orientation of His172 shifts into a vertical position, allowing for the passage of the water to the rest of the channel. This gating mechanism is visible in Figure 5. Note how the water molecule is blocked initially, and in the second conformation the molecule was allowed passage through the ar/R site. In the closed-closed and the closed-open conformations the radii at this point are 0.665 and 0.95 Angstroms, respectively. The discrepancy between the last two conformations is due to the fact that His172 was vertical in the closed-open monomer. The fluctuation between His172 between a horizontal and vertical orientation may have an effect on

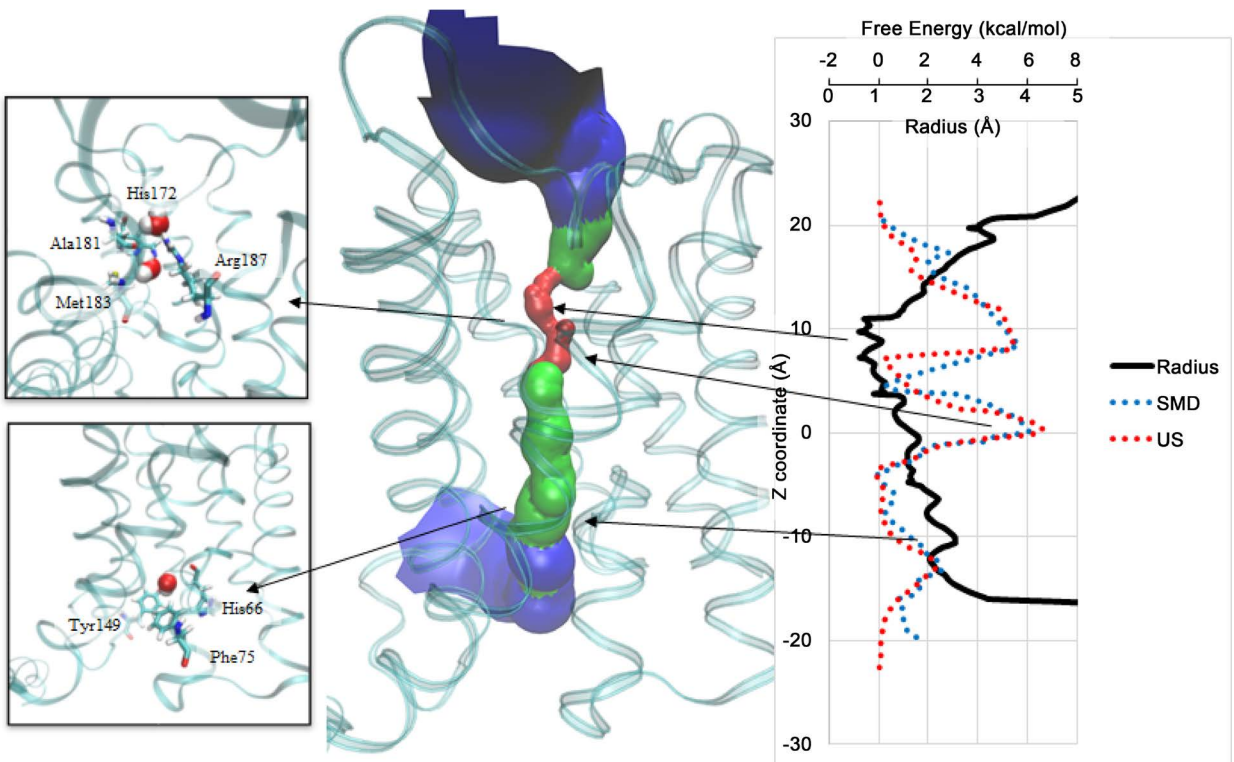


Figure 4. Measured energy profiles for water conduction, and the pore radius calculations (Right) along with the HOLE visualization of the closed-open conformation of an AQP0 monomer (Middle). The ar/R and CSII residues orientations are shown in the boxes (Left). The black solid line represents the pore radius. The blue-dashed and red-dotted lines represent SMD and umbrella sampling free energy calculations along the pore, respectively.

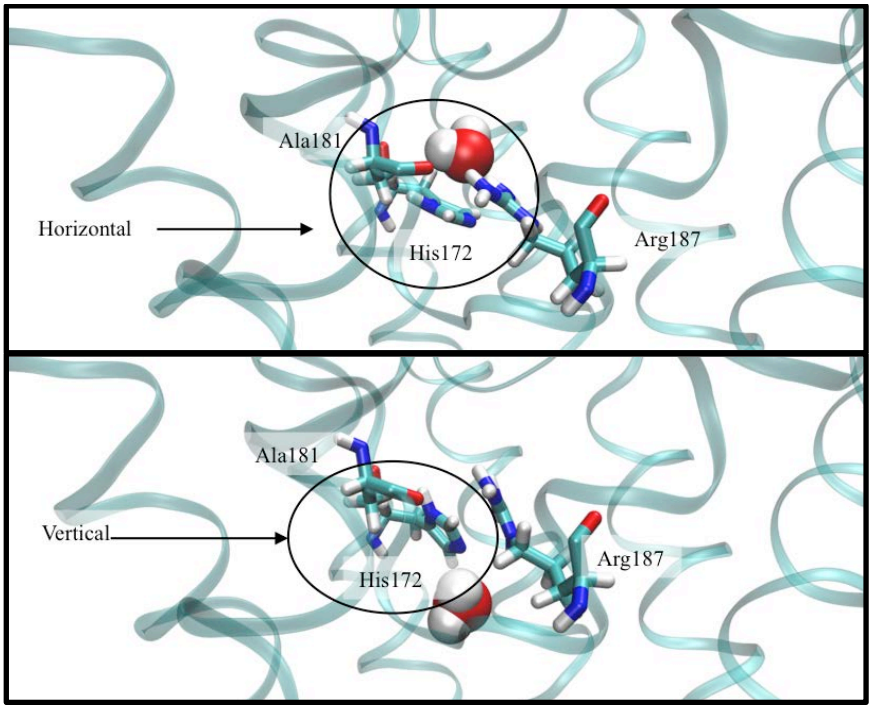


Figure 5. The dynamic motion of the His172 observed to facilitate the water conduction Steered Molecular Dynamics. The closed-closed (Top) and closed-open (Bottom) monomers show distinct difference.

water permeability between the channels and may be the deciding factor in how restrictive the monomer is at the ar/R site.

3.2. Multiple Residues Take Part in Stabilizing CSII

Figure 6 highlights RMSD data collected from a 30 ns equilibration in both the closed and open conformations. Open RMSD values are shown in orange, and closed values are in blue. The values were taken when compared to the first frame, and the atoms on the whole residue were averaged. All of the listed residues, Tyr149-His66-Phe75 are part of CSII. The end result shows that the His66 and Phe75 shifts more on closed configuration to stabilize the CSII site.

However, His172, Arg187, and Ala181, all involved with the ar/R restriction site showed no difference on RMSD analysis for closed and open models.

Figure 6 shows that that the binding of CaM protein stabilizes the AQP0 tetramer allosterically and results in a shift in position in residues that make up the CSII site of the channel. This result confirms other studies claiming that the main form of regulation of AQP0 by CaM is done via restriction of CSII [42].

4. Conclusion and Discussion

Our investigation into the AQP0 tetramer was meant to expand previous studies done and to explore the differences between the closed-closed and closed-open conformations of the protein when CaM is bound. From analysis of HOLE plots and the energy

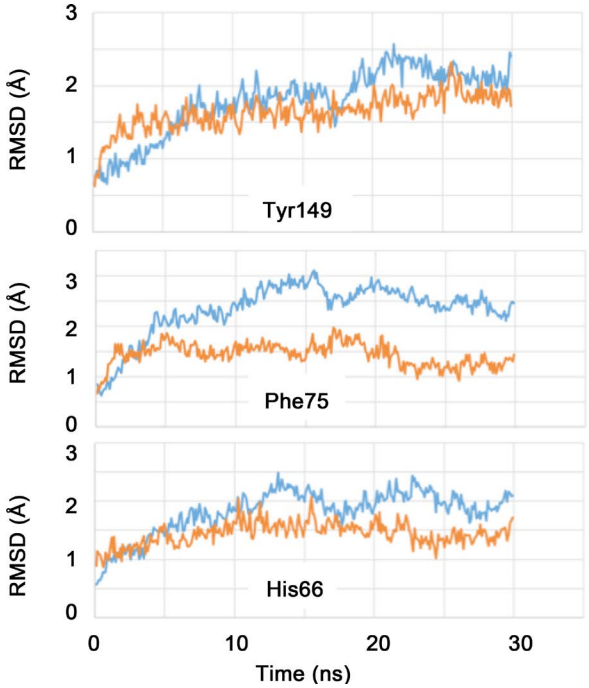


Figure 6. The RMSD distributions of the three critical residues in CSII site, Tyr149-Phe75-His66, during 30 ns equilibration run. Blue line represents the CaM bound AQP0, Orange line is AQP0 tetramer with no CaM. The coordinates are compared to the first frame, and all four monomers results were averaged.

profiles, there appears to be a significant difference between the closed-closed and the closed-open conformation in the channels. Both of these conformations have the same regions of constriction at CSI (ar/R site), but differential constriction at CSII. Energy profiles showed that the closed-closed conformation reached a peak of around 5 kcal/mol at CSII, while the value for CSII for the closed-open conformation was around 2.5. Both the HOLE plots and the energy profiles show that the different conformations in the AQP0-CaM complex may be differentially modulated, with the most restriction of water occurring in the closed-closed conformations.

Something of potential interest is the large energy peaks at the top of the channel near the ar/R site as opposed to CSII. In all conformations the energy peak is higher at the ar/R site, even when CaM is bound. The ar/R site appears to have the highest energy barrier for all channels and conformations.

Analysis of the RMSD fluctuations of the side chains involved in the ar/R site and the allosterically modulated CSII showed that the binding of CaM didn't necessarily reduce the internal thermodynamic fluctuations of side chains in the channel. Two simulations produced conflicting results, and we cannot determine whether or not CaM stabilizes the interior of the channel by reducing these fluctuations. Previous literature has suggested that CaM reduces RMSD fluctuations when it is bound. Our findings are inconsistent with this assertion. But possible interpretation may be that binding of CaM causes the side chains involved in CSII to adopt a certain conformation (a closed conformation) during the course of equilibration rather than having lower fluctuations.

Further investigation should be done, especially with longer equilibration periods.

More individual frames should be taken over the course of simulations to produce a variety of HOLE visuals for comparison. This should paint a more accurate picture of whether or not there is a clear difference between these two conformations. It would have also been useful if more SMD trajectories were produced. One problem encountered was that in a few cases the water molecule would venture off target, even though the channel was more permeable in the traditional path of water.

In conclusion, the energy profiles show that CSII has a higher energy barrier in closed-closed as opposed to closed-open. Our preliminary findings indicate that they may be significantly different. We also note the importance of His66 and Phe75 in regards to the reduced radius at CSII when CaM is bound, while previous literature focused only on Tyr149. Future investigation should flesh out and fully characterize these differences, but to date no one has determined if these channels have differential permeability. It is assumed that all channels have the same dynamics in the closed AQP0 tetramer.

Acknowledgements

The authors give our special thanks to Extreme Science and Engineering Discovery Environment (XSEDE) for the computational time awarded to our project. Dean of Coe College, via The Ella Pochobradsky Endowment Award, funded this research. Authors also would like to give special thanks to the Coe College Chemistry Department for their generous support throughout this project.

References

[1] Agre, P., Sasaki, S. and Chrispeels, M.J. (1993) Aquaporins: A Family of Water Channel Proteins. *American Journal of Physiology—Renal Physiology*, **265**, F461.

[2] Gonen, T. and Walz, T. (2006) The Structure of Aquaporins. *Quarterly Reviews of Biophysics*, **39**, 361-396. https://doi.org/10.1017/S0033583506004458

[3] Chauvigne, F., Zapater, C., Stavang, J., Taranger, G., Cerda, J. and Finn, R. (2015) The pH Sensitivity of Aqp0 in Tetraploid and Diploid Teleosts. *The FASEB Journal*, **29**, 2172-2184. https://doi.org/10.1096/fj.14-267625

[4] Knepper, M.A. (1994) The Aquaporin Family of Molecular Water Channels. *Proceedings of the National Academy of Sciences of the United States of America*, **91**, 6255-6258. https://doi.org/10.1073/pnas.91.14.6255

[5] Nielsen, S., Frøkiær, J., Marples, D., Kwon, T.-H., Agre, P. and Knepper, M.A. (2002) Aquaporins in the Kidney: From Molecules to Medicine. *Physiological Reviews*, **82**, 205-244. https://doi.org/10.1152/physrev.00024.2001

[6] Badaut, J., Lasbennes, F., Magistretti, P.J. and Regli, L. (2002) Aquaporins in Brain: Distribution, Physiology, and Pathophysiology. *Journal of Cerebral Blood Flow and Metabolism*, **22**, 367-378.

[7] Murata, K., *et al.* (2000) Structural Determinants of Water Permeation through Aquaporin-1. *Nature*, **407**, 599-605.

[8] Harries, W.E.C., Akhavan, D., Miercke, L.J.W., Khademi, S. and Stroud, R.M. (2004) The Channel Architecture of Aquaporin 0 at a 2.2—A Resolution. *Proceedings of the National*

Academy of Sciences of the United States of America, **101**, 14045-14050. https://doi.org/10.1073/pnas.0405274101

[9] de Groot, B.L. and Grubmuller, H. (2005) The Dynamics and Energetics of Water Permeation and Proton Exclusion in Aquaporins. *Current Opinion in Structural Biology*, **15**, 176-183. https://doi.org/10.1016/j.sbi.2005.02.003

[10] Costello, M.J., McIntosh, T.J. and Roberrson, J.D. (1989) Distribution of Gap Junctions and Square Array Junctions in the Mammalian Lens. *Investigative Ophthalmology and Visual Science*, **30**, 975-989.

[11] Sui, H.X., Han, B.-G., Lee, J.K., Walian, P. and Jap, B.K. (2001) Structural Basis of Water-Specific Transport through the AQP1 Water Channel. *Nature*, **414**, 872-878.

[12] Ren, G., Reddy, V.S., Melnyk, P. and Mitra, A.K. (2001) Visualization of a Water Selective Pore by Electron Crystallography in Vitreous Ice. *Proceedings of the National Academy of Sciences of the United States of America*, **98**, 1398-1403.

[13] Petrova, R., Schey, K., Donaldson, P. and Grey, A. (2015) Spatial Distributions of AQP5 and AQP0 in Embryonic and Postnatal Mouse Lens Development. *Experimental Eye Research*, **132**, 124-135. https://doi.org/10.1016/j.exer.2015.01.011

[14] Mulders, S.M., *et al.* (1995) Water Channel Properties of Major Intrinsic Protein of Lens. *The Journal of Biological Chemistry*, **270**, 9010-9016. https://doi.org/10.1074/jbc.270.15.9010

[15] Mathias, R.T., Rae, J.L. and Baldo, G.J. (1997) Physiological Properties of the Normal Lens. *Physiological Reviews*, **77**, 21-50.

[16] Chandy, G., Zampighi, G.A., Kreman, M. and Hall, J.E. (1997) Comparison of the Water Transporting Properties of MIP and AQP1. *The Journal of Membrane Biology*, **159**, 29-39. https://doi.org/10.1007/s002329900266

[17] Donaldson, P., Kistler, J. and Mathias, R.T. (2001) Molecular Solutions to Mammalian Lens Transparency. *News in Physiological Sciences*, **16**, 118-123.

[18] Debdutta, R. and Abraham, S. (1979) Human Lens Membrane: Comparison of Major Intrinsic Polypeptides from Young and Old Lenses Isolated by a New Methodology. *Experimental Eye Research*, **28**, 353-358. https://doi.org/10.1016/0014-4835(79)90097-6

[19] Bok, D., Dockstader, J. and Horwitz, J. (1982) Immunocytochemical Localization of the Lens Main Intrinsic Polypeptide (MIP26) in Communicating Junctions. *The Journal of Cell Biology*, **92**, 213-220. https://doi.org/10.1083/jcb.92.1.213

[20] Takemoto, L., Takehana, M. and Horwitz, J. (1986) Covalent Changes in MIP26K during Aging of the Human Lens Membrane. *Investigative Ophthalmology and Visual Science*, **27**, 443-446.

[21] Gonen, T., Cheng, Y.F., Kistler, J. and Walz, T. (2004) Aquaporin-0 Membrane Junctions Form upon Proteolytic Cleavage. *Journal of Molecular Biology*, **342**, 1337-1345. https://doi.org/10.1016/j.jmb.2004.07.076

[22] Gonen, T., Cheng, Y.F., Sliz, P., Hiroaki, Y., Fujiyoshi, Y., Harrison, S.C. and Walz, T. (2005) Lipid-Protein Interactions in Double-Layered Two-Dimensional AQP0 Crystals. *Nature*, **438**, 633-638. https://doi.org/10.1038/nature04321

[23] Ball, L.E., Garland, D.L., Crouch, R.K. and Schey, K.L. (2004) Post-Translational Modifications of Aquaporin 0 (AQP0) in the Normal Human Lens: Spatia and Temporal Occurrence. *Biochemistry*, **43**, 9856-9865. https://doi.org/10.1021/bi0496034

[24] Berry, V., Francis, P., Kaushal, S., Moore, A. and Bhattacharya, S. (2000) Missense Mutations in MIP Underlie Autosomal Dominant "Polymorphic" and Lamellar Cataracts Linked to 12q. *Nature Genetics*, **25**, 15-17. https://doi.org/10.1038/75538

[25] Francis, P., Berry, V., Bhattacharya, S. and Moore, A. (2000) Congenital Progressive Polymorphic Cataract Caused by a Mutation in the Major Intrinsic Protein of the Lens, MIP (AQP0). *British Journal of Ophthalmology*, **84**, 1376-1379. https://doi.org/10.1136/bjo.84.12.1376

[26] Geyer, D.D., Spence, M.A., Johannes, M., Flodman, P., Clancy, K.P., Berry, R., *et al.* (2006) Novel Single-Base Deletional Mutation in Major Intrinsic Protein (*MIP*) in Autosomal Dominant Cataract. *American Journal of Ophthalmology*, **141**, 761-761.e4. https://doi.org/10.1016/j.ajo.2005.11.008

[27] Yu, Y.B., *et al.* (2007) A Novel MIP Gene Mutation Associated with Autosomal Dominant Congenital Cataracts in a Chinese Family. *Molecular Vision*, **13**, 1651-1656.

[28] Nemeth-Cahalan, K.L., Kalman, K. and Hall, J.E. (2004) Molecular Basis of pH and Ca^{2+} Regulation of Aquaporin Water Permeability. *The Journal of General Physiology*, **123**, 573-580. https://doi.org/10.1085/jgp.200308990

[29] Nemeth-Cahalan, K.L. and Hall, J.E. (2000) pH and Calcium Regulate the Water Permeability of Aquaporin 0. *The Journal of Biological Chemistry*, **275**, 6777-6782. https://doi.org/10.1074/jbc.275.10.6777

[30] Varadaraj, K., Kumari, S., Shiels, A. and Mathias, R.T. (2005) Regulation of Aquaporin Water Permeability in the Lens. *Investigative Ophthalmology and Visual Science*, **46**, 1393-1402. https://doi.org/10.1167/iovs.04-1217

[31] Louis, C.F., Hogan, P., Visco, L. and Strasburg, G. (1990) Identity of the Calmodulin-Binding Proteins in Bovine Lens Plasma Membranes. *Experimental Eye Research*, **50**, 495-503. https://doi.org/10.1016/0014-4835(90)90038-V

[32] Girsch, S.J. and Peracchia, C. (1991) Calmodulin Interacts with a C-Terminus Peptide from the Lens Membrane Protein MIP26. *Current Eye Research*, **10**, 839-849. https://doi.org/10.3109/02713689109013880

[33] Kalman, K., Nemeth-Cahalan, K.L., Froger, A. and Hall, J.E. (2008) Phosphorylation Determines the Calmodulin-Mediated Ca^{2+} Response and Water Permeability of AQP0. *The Journal of Biological Chemistry*, **283**, 21278-21283. https://doi.org/10.1074/jbc.M801740200

[34] Lindsey Rose, K.M., Wang, Z., Magrath, G.N., Hazard, E.S., Hildebrandt, J.D. and Schey, K.L. (2008) Aquaporin 0-Calmodulin Interaction and the Effect of Aquaporin 0 Phosphorylation. *Biochemistry*, **47**, 339-347. https://doi.org/10.1021/bi701980t

[35] Reichow, S.L. and Gonen, T. (2008) Noncanonical Binding of Calmodulin to Aquaporin-0: Implications for Channel Regulation. *Structure*, **16**, 1389-1398. https://doi.org/10.1016/j.str.2008.06.011

[36] Reichow, S.L., Clemens, D.M., Freites, J.A., Németh-Cahalan, K.L., Heyden, M., Tobias, D.J., *et al.* (2013) Allosteric Mechanism of Water-Channel Gating by Ca^{2+}-Calmodulin. *Nature Structural & Molecular Biology*, *Advanced Online Publication*, **20**, 1085-1092. https://doi.org/10.1038/nsmb.2630

[37] Sudhakar Badu, Y., Sack, J.S., Greenhough, T.J., Bugg, C.E., Means, A.R. and Cook, W.J. (1985) Three-Dimensional Structure of Calmodulin. *Nature*, **315**, 37-40. https://doi.org/10.1038/315037a0

[38] Zhang, M.J., Tanaka, T. and Ikura, M. (1995) Calcium-Induced Conformational Transition Revealed by the Solution Structure of Apo Calmodulin. *Nature Structural Biology*, **2**, 758-767. https://doi.org/10.1038/nsb0995-758

[39] Kuboniwa, H., Tjandra, N., Grzesiek, S., Ren, H., Klee, C.B. and Bax, A. (1995) Solution Structure of Calcium-Free Calmodulin. *Nature Structural Biology*, **2**, 768-776. https://doi.org/10.1038/nsb0995-768

[40] Chou, J.J., Li, S., Klee, B. and Bax, A. (2001) Solution Structure of Ca^{2+}-Calmodulin Reveals Flexible Hand-Like Properties of Tis Domains. *Nature Structural Biology*, **8**, 990-997. https://doi.org/10.1038/nsb1101-990

[41] Gold, M.G., Reichow, S.L., O'Neill, S.E., Weisbrod, C.R., Langeberg, L.K., Bruce, J.E., *et al.* (2012) AKAP2 Anchors PKA with Aquaporin-0 to Support Ocular Lens Transparency. *EMBO Molecular Medicine*, **4**, 15-26. https://doi.org/10.1002/emmm.201100184

[42] Reichow, S., Clemens, D., Freites, J.A., Nemeth-Cahalan, K.L., Heyden, M., Tobias, D., Hall, J. and Gonen, T. (2013) Allosteric Mechanism of Water-Channel Gating. *Nature Structural & Molecular Biology*, **20**, 1085-1092. https://doi.org/10.1038/nsmb.2630

[43] Humphrey, W., Dalke, A. and Schulten, K. (1996) VMD—Visual Molecular Dynamics. *Journal of Molecular Graphics*, **14**, 33-38. https://doi.org/10.1016/0263-7855(96)00018-5

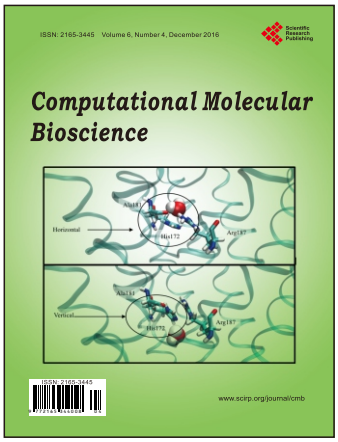
[44] Brooks, B.R., Bruccoleri, R.E., Olafson, B.D., States, D.J., Swaminathan, S. and Karplus, M. (1983) CHARMM: A Program for Macromolecular Energy, Minimization, and Dynamics Calculations. *Journal of Computational Chemistry*, **4**, 187-217. https://doi.org/10.1002/jcc.540040211

[45] Phillips, J.C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R.D., Kale, L. and Schulten, K. (2005) Scalable Molecular Dynamics with NAMD. *Journal of Computational Chemistry*, **26**, 1781-1802. https://doi.org/10.1002/jcc.20289

[46] Palmieri, B. and Ronis, D. (2007) Jarzynski Equality: Connections to Thermodynamics and the Second Law. *Physical Review*, **75**, Article ID: 011133. https://doi.org/10.1103/physreve.75.011133

[47] Smart, O.S., Goodfellow, J.M. and Wallace, B.A. (1993) The Pore Dimensions of Gramidicin A. *Biophysical Journal*, **65**, 2455-2460. https://doi.org/10.1016/S0006-3495(93)81293-1

Call_for_Papers



Computational Molecular Bioscience

ISSN: 2165-3445 (Print) ISSN: 2165-3453 (Online)
http://www.scirp.org/journal/cmb

***Computational Molecular Bioscience* (*CMB*)** is an international journal dedicated to the latest advancement of Computational Molecular Bioscience. The goal of this journal is to provide a platform for scientists and academicians all over the world to promote, share, and discuss various new issues and developments in different areas of Computational Molecular Bioscience. All manuscripts must be prepared in English, and are subject to a rigorous and fair peer-review process. Accepted papers will immediately appear online followed by printed hard copy. The journal publishes original papers including but not limited to the following fields:

- *Ab Initio* and Density Functional Calculations of Biomolecules
- Atomistic and Coarse Grained Molecular Dynamics
- Combined Computational and Experimental Studies of Biomolecular Interactions
- Combined Quantum Mechanical and Molecular Mechanical Methods (QM/MM)
- Computational Chemistry of Biomolecules, Ligands and Drugs
- Computational Drug Design: Structure-Based; Ligand-Based; Rational; *De Novo*
- Computational Modelling of Biomolecular Structures Interactions and Processes
- Computational Systems Biology and Chemistry
- Development and Applications of Monte Carlo Methods
- Development and Design of New Biological and Chemical Databases and Data Mining Techniques
- Development, Testing and Applications to Biomolecular Systems
- Enzymatic Reaction Mechanisms and Inhibition
- High Performance Computing in Molecular and Biomolecular Sciences
- Ligand Binding and Free Energy Calculations
- Modelling of Membrane Processes and Protein-Membrane Interactions
- Modelling Protein Structure, Conformational Dynamics an Interactions
- Molecular Mechanics, Force Field Development and Evaluation
- Molecular Visualizations and Data Analysis
- Multilevel Computational Simulations
- Nucleic Acids Structure, Dynamics and Interactions with Ligands
- Protein Folding
- Protein Ligand Docking New Algorithm, Codes and Applications
- Protein-Nucleic Acids Interactions
- Quantitative Structure-Activity Relationships (QSAR)
- Semiempirical Electronic Structure Calculations
- Structural Bioinformatics and Homology Modelling

We are also interested in: 1) Short Reports—2-5 page papers where an author can either present an idea with theoretical background but has not yet completed the research needed for a complete paper or preliminary data; 2) Book Reviews—Comments and critiques.

Website and E-Mail

http://www.scirp.org/journal/cmb **E-mail: cmb@scirp.org**

What is SCIRP?

Scientific Research Publishing (SCIRP) is one of the largest Open Access journal publishers. It is currently publishing more than 200 open access, online, peer-reviewed journals covering a wide range of academic disciplines. SCIRP serves the worldwide academic communities and contributes to the progress and application of science with its publication.

What is Open Access?

All original research papers published by SCIRP are made freely and permanently accessible online immediately upon publication. To be able to provide open access journals, SCIRP defrays operation costs from authors and subscription charges only for its printed version. Open access publishing allows an immediate, worldwide, barrier-free, open access to the full text of research papers, which is in the best interests of the scientific community.

- High visibility for maximum global exposure with open access publishing model
- Rigorous peer review of research papers
- Prompt faster publication with less cost
- Guaranteed targeted, multidisciplinary audience

Scientific Research Publishing

Website: http://www.scirp.org

Subscription: sub@scirp.org

Advertisement: service@scirp.org