Unitary Structure of Palindromes in DNA

Mehmet Ali Tibatan^{1,*} and Mustafa Sarısanman^{2,†}

¹*Department of Biotechnology, Istanbul University, 34134, Vezneciler, Istanbul, Turkey*

²*Department of Physics, Istanbul University, 34134, Vezneciler, Istanbul, Turkey*

We investigate the quantum behavior encountered in palindromes within DNA structure. In particular, we reveal the unitary structure of usual palindromic sequences found in genomic DNAs of all living organisms, using the Schwinger's approach. We clearly demonstrate the role played by palindromic configurations with special emphasis on physical symmetries, in particular subsymmetries of unitary structure. We unveil the prominence of unitary structure in palindromic sequences in the sense that vitally significant information endowed within DNA could be transformed unchangeably in the process of transcription. We introduce a new symmetry relation, namely purine-purine or pyrimidine-pyrimidine symmetries (p-symmetry) in addition to the already known symmetry relation of purine-pyrimidine symmetries (pp-symmetry) given by Chargaff's rule. Therefore, important vital functions of a living organisms are protected by means of these symmetric features. It is understood that higher order palindromic sequences could be generated in terms of the basis of the highest prime numbers that make up the palindrome sequence number. We propose that violation of this unitary structure of palindromic sequences by means of our proposed symmetries leads to a mutation in DNA, which could offer a new perspective in the scientific studies on the origin and cause of mutation.

PACS numbers: 03.65.-w, 02.30.Tb, 02.90.+p, 87.10.+e, 87.14.Gg, 87.14.G-

INTRODUCTION

Sequence specific functionality drives the proteins enhanced functions and play role as a guidance but the question of why DNA require palindromic sequences has not clear answer yet but there are only few examples in the literature on how palindromic sequences works. Nevertheless, question of 'are there any reason to determine the palindromic sequences as important and significant rather than arbitrary nucleic acid sequences?' might require more detailed answers.

Ordinarily, nucleic bases can have different properties from each other despite their structural similarities (See, Fig.1). For instance, adenine shows different aminoacid residue preferences than the guanine [1] or cytosine can present unique UV-absorption spectrum which differs from absorption spectrum of the other bases [2], additionally, guanine has the most reactive site in the DNA against reactive oxygen species [3]. Different aminoacid residue preference of the bases can be a defining feature by means of protein interaction with relative gene regions. Although nucleic basis have similar affinity to protons which directly affect the tautomeric stabilization levels, their substrate binding energetics may differ [4]. In other words, structural similarities and differences of nucleic bases provide a specific molecular recognition [1]. In this case, the molecular recognition is important in terms of information conservation because notably quadruplet and sextet palindromic sequences provide unique binding motifs especially for the transcription starter proteins or enzymes which can break the DNA strands.

Palindromic structure of DNA is nucleic acid sequences which are symmetrical with their sense and anti-sense strand. These specific sequences are suggested to enhance affinity for transcriptional regions [5] and more importantly they can involve the recombination repair mechanism because both of sense and anti-sense regions of palindromic DNA sequences are symmetrical which supports the enzymatic proteins to bind both strands at the same time. Studies show that palindromic sequences might help to homologous recombination and palindrome containing DNA regions are indented to replicate easier [6]. Moreover, previous studies have shown that these palindromic oligonucleotides have remarkable immunogenetic properties and potantiate the induction of interferon release from natural killer cells [7, 8]. However, above all the explanations regarding the functionality of palindrome, the best approach may be the ability of these specific sequences to form hairpin [9]. Nag and Kurst revealed that palindromic sequences present breakpoints during the cellular process which requires double-strand breaks such as meiosis [10]. Briefly, it has been suggested that palindromic sequences help DNA's folding process (forming hairpin structures) and recognition by nucleases, as well as that RNA viruses use palindromic sequences to dimerize the viral genome, which provides protection against enzymatic degradation by the host cell. Even, it has been noted that they cause some problems with sequencing technologies due to their tendency to fold [10–12].

However, despite all these alternative explanations, the role of palindromic sequences within our genome requires more detailed explanation because it is not clear yet how energetic functionality of such symmetric structures takes

FIG. 1: Simple ring structure of nucleobases which forms the DNA, hydrogen bonds not shown. Adenine and guanine (purines) have two-ring structure differently from cytosine and thymine (pyrimidine). Explicit unitary symmetries are displayed between adenine and thymine, and cytosine and guanine, which reflect the Chargaff's rule.

advantage through the cellular mechanism, and even how they potentially differ from the non-palindromic sequences. Therefore, the roles of palindromes should be more comprehensively uncovered in order to advance nucleobase function within the DNA. Nucleotide functions and sequence-specific protein interaction has been extensively studied, and such studies are still ongoing [13]. The electrochemical properties of the nucleobases shown in Fig. 1 may explain their functionality [12–15], but this may not be sufficient to elucidate their selective function in DNA during cellular activities.

How enzymes recognize the double strand of the palindrome, why nucleobases tend to form hairpin structure during replication, and how nucleobase alignment affects DNA stability may have an electrochemical explanation. However, the study of nucleobase behavior and thus the palindromic sequence mechanism at the quantum level will improve all these explanations and provide some insight into the energetic principles of the nucleic acid sequences [13, 14]. Therefore, we present a rudimental approach to DNA at the quantum level by revealing the palindromic sequences to better understand the nature of nucleobases, which may provide some explanation for their energetic mechanism.

In view of all these fact, we present an alternative perspective to palindromes in DNA, which reflect their unitary structure distinctively. This approach leads to the emergence of unknown symmetries of palindromes with very important functions. We propose a novel symmetry relation which has not been realized yet, which gives rise to other vitally important accompanying palindromes. In particular, we focused on specific palindromes whose properties and roles are known extensively, such as **CACGTG** depicted in Fig. 2 which is called as E-box consensus, and showed the importance of the symmetry relations we found. This palindromic sequence is normally found at the protein binding site in the breast cancer cells [15–17]. For example, since human c-myc protein has been reported to be over-expressed in tumor progression, it is responsible for cancer development [18, 19]. Additionally, E-box elements within the genome provide selectivity for transcription factors to find their target sites, so these sequences are important to bring DNA recognition specificity [16]. We propose a quantum-mechanical explanation of nucleobase symmetry in order to identify the role of a palindromic site to bind these tumorigenic proteins.

PALINDROMIC STRUCTURE OF DNA AND ITS PHYSICAL SYMMETRIES

Palindromic nature of DNA is rather significant for the maintenance of vital functions of living mechanisms. This is why we persistently witness some distinguished palindromic forms in almost all important genes. To understand its vital role when examined more closely from a physical perspective, we notice that some featured physical symmetries are encountered in these particular regions of DNA. In the process of DNA replication by a specific enzyme, initially given chain of DNA is transformed into its counterpart by means of nucleotide rules (known as Chargaff's rule). This procedure permits an inversion symmetry in the whole DNA entity. This in fact very similar to chiral symmetries of fermions in the chiral representation of Dirac type particles in physics, in which a right-handed particle exhibits similar behavior as the left-handed one [20–22]. Fig. 3 displays the inversion symmetries in **TATA**-box sequence and **CACGTG** sequence in E-box, which reflects the chiral nature of screened palindromes.

TATA-box element within DNA is another example of palindromic sequence. It is very common and well conserved genomic sequence and can be found in many species from plants to human genome. Normally, **TATA** sequence has many different variance, but **TATA** content (thymine - adenine -adenine thymine-adenine core) is constant and giving the **TATA**-box name. Primary function of the **TATA**-box is to provide protein binding region for the transcriptional process and it locates near the promoter region of the related genes. **TATA** sequence itself is short palindromic sequence and also it is core for the consensus sequence and has many variance [23]. As for all the other palindromic

FIG. 2: (Color online) Ball and stick view of the palindromic E-box consensus (**CACGTG** sequence). Dashed yellow lines indicate H-bonds between the nucleobases. This figure clearly shows the intrinsic unitary structure of **CACGTG** in the form of chiral symmetry of Dirac. Image prepared by using Avogadro software (ver. 1.2.0)

sequences, **TATA** sequence has also core properties and conserved sequence within the eukaryotic cells. Thus it is another significant example to reveal palindromic sequence importance in eukaryotic genome. Moreover it can contribute the viral contamination by helping the viral transcriptional protein involvement to the genome [24, 32].

PALINDROMES IN DNA AND THEIR UNITARY STRUCTURE

Consider a set of quantum states which consist of singlet structures $|\zeta_i\rangle$, where the index i denotes the order of singlets in the DNA sequence. In principle, $i = (1, 2)$. Now, consider another set of quantum states $|\xi_j\rangle$, which is different from the first one in basic aspects. We describe $|\zeta_1\rangle$ and $|\zeta_2\rangle$ by the nucleotides adenine **|A⟩** and thymine **|T⟩** respectively in DNA, whereas $|\xi_1\rangle$ and $|\xi_2\rangle$ represent the nucleotides guanine **|G⟩** and cytosine **|C⟩** respectively, see Fig. 1 for the structures of these nucleotides, see [25–31] for further details on quantum mechanical description of states.

We define an operator $\hat{U}$ on the set $|\zeta_i\rangle$ by the map $\hat{U} : |\zeta_i\rangle \rightarrow |\zeta_{i+1}\rangle \bmod 2$. Likewise, an operator $\hat{V}$ is defined on

FIG. 3: (Color online) Schematic presentation of palindromic form of **TATA**-box (left panel) and E-box sequence **CACGTG** (right panel), respectively. In both cases, Helicase enzyme is breaking hydrogen bonds between nucleotides during DNA replication or transcription to associate the polymerase enzyme activity.

the set $|\xi_i\rangle$ by the map $\hat{V} : |\xi_i\rangle \rightarrow |\xi_{i+1}\rangle \bmod 2$. Notice that our sets of quantum states generate a periodic structure such that they form the following closed operations,

$$|\mathbf{A}\rangle \xleftrightarrow{\hat{U}} |\mathbf{T}\rangle, \quad (1)$$

$$|\mathbf{G}\rangle \xleftrightarrow{\hat{V}} |\mathbf{C}\rangle. \quad (2)$$

In fact, the operators $\hat{U}$ and $\hat{V}$ translating quantum states within a nucleotide set correspond to an enzyme, and in principle the same enzyme operates for each nucleotide pair, it operates uncustomarily on distinct nucleotide pairs. Physically, this amounts to the noncommutativity of the operators $\hat{U}$ and $\hat{V}$, i.e. $[\hat{U}, \hat{V}] \neq 0$. Notice that $\hat{U}$ and $\hat{V}$ operators satisfy the following identity in a doubly acted accompanying singlet nucleotide system,

$$\hat{U}^2 = \hat{V}^2 = \mathbb{I}, \quad (3)$$

where $\mathbb{I}$ denotes the identity matrix such that our set is complete. This is in fact our completeness relation. This relation tells us that eigenvalues for the $\hat{U}$ and $\hat{V}$ operators are given by u'_k and v'_k

$$u'_k = v'_k = e^{\pi i k}, \quad (4)$$

where $k = 1, 2$ gives rise to the relevant eigenvalues. We realize that eigenvalues are $u'_k = v'_k = (-, +)$. Notice that both accompanying states $|\mathbf{A}\rangle - |\mathbf{T}\rangle$ and $|\mathbf{G}\rangle - |\mathbf{C}\rangle$ play the role of $|+\rangle$ and $|-\rangle$ states in case of spin states of Dirac particles. But, we understand that the action is performed by the Pauli spin matrices $\hat{\sigma}_x$ and $\hat{\sigma}_z$ in two dimensions such that the role of the corresponding enzyme is to switch the polarity (spin) of the nucleotides. This can be seen explicitly in Fig. (1). Hence, it is noted that $|\mathbf{A}\rangle$ and $|\mathbf{T}\rangle$ states are the eigenstates of $\hat{\sigma}_z$, whereas $|\mathbf{G}\rangle$ and $|\mathbf{C}\rangle$ states are the eigenstates of $\hat{\sigma}_x$ operator. Furthermore, $\hat{\sigma}_x$ operator transforms $|\mathbf{A}\rangle$ and $|\mathbf{T}\rangle$ states into each other, just as $\hat{\sigma}_z$ operator transforms $|\mathbf{G}\rangle$ and $|\mathbf{C}\rangle$ states into one another in a similar manner. These mutually periodic actions can be expressed in a clear way as follows,

$$\hat{\sigma}_x |\zeta_i\rangle = |\zeta_{i+1}\rangle, \quad \hat{\sigma}_z |\xi_i\rangle = |\xi_{i+1}\rangle. \quad (5)$$

We can briefly demonstrate all these actions in the following compact forms

$$\begin{pmatrix} \hat{\sigma}_x & 0 \\ 0 & \hat{\sigma}_x \end{pmatrix} \begin{pmatrix} |\mathbf{A}\rangle \\ |\mathbf{T}\rangle \end{pmatrix} = \begin{pmatrix} |\mathbf{T}\rangle \\ |\mathbf{A}\rangle \end{pmatrix}, \quad \begin{pmatrix} \hat{\sigma}_z & 0 \\ 0 & \hat{\sigma}_z \end{pmatrix} \begin{pmatrix} |\mathbf{G}\rangle \\ |\mathbf{C}\rangle \end{pmatrix} = \begin{pmatrix} |\mathbf{C}\rangle \\ |\mathbf{G}\rangle \end{pmatrix} \quad (6)$$

In general, we would like to find the eigenstates of the operators $\hat{U}$ and $\hat{V}$ with eigenvalues u'_i and v'_i , respectively. We realize that they are figured out to be the antisymmetric ordering of the states $|u_i\rangle$ and $|v_i\rangle$ such that $|u_+\rangle := |\mathbf{G}\rangle$, $|u_-\rangle := |\mathbf{C}\rangle$, $|v_+\rangle := |\mathbf{A}\rangle$ and $|v_-\rangle := |\mathbf{T}\rangle$. Thus, we observe that

$$\hat{\sigma}_x |u_i\rangle = u'_i |u_i\rangle, \quad \hat{\sigma}_z |v_i\rangle = v'_i |v_i\rangle, \quad (7)$$

with corresponding eigenvalues $u'_\pm = v'_\pm = \pm$. It is obvious that only way to achieve these relations is to identify $|\mathbf{G}\rangle := \frac{1}{\sqrt{2}}(|\mathbf{A}\rangle + |\mathbf{T}\rangle)$ and $|\mathbf{C}\rangle := \frac{1}{\sqrt{2}}(|\mathbf{A}\rangle - |\mathbf{T}\rangle)$, together with their counterparts $|\mathbf{A}\rangle := \frac{1}{\sqrt{2}}(|\mathbf{G}\rangle + |\mathbf{C}\rangle)$ and $|\mathbf{T}\rangle := \frac{1}{\sqrt{2}}(|\mathbf{G}\rangle - |\mathbf{C}\rangle)$. It is not difficult to observe the commutation relation of the operators $\hat{U}$ and $\hat{V}$, which is given by

$$\hat{U}\hat{V} = e^{\pi i}\hat{V}\hat{U}, \quad (8)$$

which gives rise to usual commutation relation of Pauli matrices: $\hat{\sigma}_x\hat{\sigma}_z = -\hat{\sigma}_z\hat{\sigma}_x$. Commutation of the operators $\hat{U}$ and $\hat{V}$ leads to the multi-ordering of nucleotides such that more complicated structures can be constructed in a given region of a DNA sequence. For example, a sequence involving two nucleotide sets can be fulfilled. Full set of such a binary set contains exactly 16 sequences[45]. We will not consider non-palindromic cases in this work, but will focus our attention to the palindromic cases which include only 4 members, which are $|\mathbf{AT}\rangle, |\mathbf{TA}\rangle, |\mathbf{GC}\rangle, |\mathbf{CG}\rangle$. These binary sets can be found in many DNA sequences. Notice that this set of binary palindromic structures are closed and admits a symmetry under the exchange of $\mathbf{A} \rightleftharpoons \mathbf{T}$ and $\mathbf{G} \rightleftharpoons \mathbf{C}$. This closure property is inherent to the palindromic feature. Table I shows the number of palindromic sequences in binary structures.

Palindromic Sequence	<i>Catenin</i>	<i>PMEPA</i>	<i>P50</i>	<i>NKILA</i>	<i>P65</i>
$ \mathbf{AT}\rangle$	5170	2808	9187	106	679
$ \mathbf{TA}\rangle$	4713	2037	8214	88	555
$ \mathbf{GC}\rangle$	2822	4712	3957	169	1152
$ \mathbf{CG}\rangle$	671	1811	755	76	462

TABLE I: Table depicts the numbers of binary palindromic sequences in special genes of *Catenin*, *PMEPA*, *P50*, *NKILA* and *P65*. These values are extracted from FASTA sequences obtained through NCBI database.

In any palindromic sequence consisting of even multiples of binary sequence structures, one can safely choose individual states $|\zeta_i\rangle$ and $|\xi_i\rangle$ as the basis of more complicated patterns. In case of a binary palindromic system, we can represent basis as $|\zeta_i, \zeta_{i+1}\rangle := |\zeta_i\rangle \otimes |\zeta_{i+1}\rangle$ and $|\xi_i, \xi_{i+1}\rangle := |\xi_i\rangle \otimes |\xi_{i+1}\rangle$. More clearly, these states are expressed as follows

$$|\zeta_i, \zeta_{i+1}\rangle = \begin{pmatrix} |\mathbf{A}\rangle \\ |\mathbf{T}\rangle \end{pmatrix} \otimes \begin{pmatrix} |\mathbf{T}\rangle \\ |\mathbf{A}\rangle \end{pmatrix} = \begin{pmatrix} |\mathbf{AT}\rangle \\ |\mathbf{AA}\rangle \\ |\mathbf{TT}\rangle \\ |\mathbf{TA}\rangle \end{pmatrix}, \quad |\xi_i, \xi_{i+1}\rangle = \begin{pmatrix} |\mathbf{G}\rangle \\ |\mathbf{C}\rangle \end{pmatrix} \otimes \begin{pmatrix} |\mathbf{C}\rangle \\ |\mathbf{G}\rangle \end{pmatrix} = \begin{pmatrix} |\mathbf{GC}\rangle \\ |\mathbf{GG}\rangle \\ |\mathbf{CC}\rangle \\ |\mathbf{CG}\rangle \end{pmatrix} \quad (9)$$

If we represent the operators transforming a purine and a pyrimidine states into their corresponding palindromic counterparts by $\hat{P}$ and $\hat{Q}$ respectively, we can demonstrate that

$$\hat{P}|\zeta_i, \zeta_{i+1}\rangle = |\zeta_{i+1}, \zeta_i\rangle, \quad \text{and} \quad \hat{Q}|\xi_i, \xi_{i+1}\rangle = |\xi_{i+1}, \xi_i\rangle \quad (10)$$

Thus, we can show that $\hat{P}$ operator can be written as $\hat{P} := \hat{U}_1\hat{U}_2$ for the pure purine state $|\zeta_i, \zeta_{i+1}\rangle$ and $\hat{Q} := \hat{V}_1\hat{V}_2$ for the pure pyrimidine state $|\xi_i, \xi_{i+1}\rangle$. Here, subindices 1 and 2 represent the order of the operator that acts on the binary state in the corresponding order. Since the same enzyme operates distinctively between two sequences, we may describe that $\hat{U}_i = \hat{\sigma}_x$ and $\hat{V}_i = \hat{\sigma}_z$, and therefore $\hat{P} = \hat{\sigma}_x \otimes \hat{\sigma}_x$ and $\hat{Q} = \hat{\sigma}_z \otimes \hat{\sigma}_z$. Tensor product implies the ordered operator actions. In particular, one can show Eq. 10 in matrix notation as follows

$$\begin{pmatrix} \hat{\sigma}_x \otimes \hat{\sigma}_x & 0 \\ 0 & \hat{\sigma}_x \otimes \hat{\sigma}_x \end{pmatrix} \begin{pmatrix} |\mathbf{A}\rangle \\ |\mathbf{T}\rangle \end{pmatrix} \otimes \begin{pmatrix} |\mathbf{T}\rangle \\ |\mathbf{A}\rangle \end{pmatrix} = \begin{pmatrix} |\mathbf{T}\rangle \\ |\mathbf{A}\rangle \end{pmatrix} \otimes \begin{pmatrix} |\mathbf{A}\rangle \\ |\mathbf{T}\rangle \end{pmatrix}, \quad \text{and} \\ \begin{pmatrix} \hat{\sigma}_z \otimes \hat{\sigma}_z & 0 \\ 0 & \hat{\sigma}_z \otimes \hat{\sigma}_z \end{pmatrix} \begin{pmatrix} |\mathbf{G}\rangle \\ |\mathbf{C}\rangle \end{pmatrix} \otimes \begin{pmatrix} |\mathbf{C}\rangle \\ |\mathbf{G}\rangle \end{pmatrix} = \begin{pmatrix} |\mathbf{C}\rangle \\ |\mathbf{G}\rangle \end{pmatrix} \otimes \begin{pmatrix} |\mathbf{G}\rangle \\ |\mathbf{C}\rangle \end{pmatrix}, \quad (11)$$

or, more explicitly,

$$\begin{pmatrix} \hat{\sigma}_x \otimes \hat{\sigma}_x & 0 & 0 & 0 \\ 0 & \hat{\sigma}_x \otimes \hat{\sigma}_x & 0 & 0 \\ 0 & 0 & \hat{\sigma}_x \otimes \hat{\sigma}_x & 0 \\ 0 & 0 & 0 & \hat{\sigma}_x \otimes \hat{\sigma}_x \end{pmatrix} \begin{pmatrix} |\mathbf{AT}\rangle \\ |\mathbf{AA}\rangle \\ |\mathbf{TT}\rangle \\ |\mathbf{TA}\rangle \end{pmatrix} = \begin{pmatrix} |\mathbf{TA}\rangle \\ |\mathbf{TT}\rangle \\ |\mathbf{AA}\rangle \\ |\mathbf{AT}\rangle \end{pmatrix},$$

$$\begin{pmatrix} \hat{\sigma}_z \otimes \hat{\sigma}_z & 0 & 0 & 0 \\ 0 & \hat{\sigma}_z \otimes \hat{\sigma}_z & 0 & 0 \\ 0 & 0 & \hat{\sigma}_z \otimes \hat{\sigma}_z & 0 \\ 0 & 0 & 0 & \hat{\sigma}_z \otimes \hat{\sigma}_z \end{pmatrix} \begin{pmatrix} |\mathbf{GG}\rangle \\ |\mathbf{GC}\rangle \\ |\mathbf{CG}\rangle \\ |\mathbf{CC}\rangle \end{pmatrix} = \begin{pmatrix} |\mathbf{CC}\rangle \\ |\mathbf{CG}\rangle \\ |\mathbf{GC}\rangle \\ |\mathbf{GG}\rangle \end{pmatrix}. \quad (12)$$

Notice that eigenvalues of $\hat{P}$ operator is given by $e^{i\pi(2k+1)}$ for $k = 1, 2 \bmod 2$. But, since k is an integer, this corresponds to -1 always.

Likewise, one can continue with quadruple and sextet structures. Notice that one needs to build appropriate basis for the vector space of relevant DNA sequence. In case of quadruple structure, there 256 distinct species. But not all of them are of palindromic nature. Palindromic cases are $|\mathbf{AATT}\rangle$, $|\mathbf{ATAT}\rangle$, $|\mathbf{AGCT}\rangle$, $|\mathbf{ACGT}\rangle$, $|\mathbf{TTAA}\rangle$, $|\mathbf{TATA}\rangle$, $|\mathbf{TGCA}\rangle$, $|\mathbf{TCGA}\rangle$, $|\mathbf{CATG}\rangle$, $|\mathbf{CTAG}\rangle$, $|\mathbf{CCGG}\rangle$, $|\mathbf{CGCG}\rangle$, $|\mathbf{GATC}\rangle$, $|\mathbf{GTAC}\rangle$, $|\mathbf{GGCC}\rangle$, $|\mathbf{GCGC}\rangle$, which can be figured out to be $|\zeta_i, \zeta_i, \zeta_{i+1}, \zeta_{i+1}\rangle$, $|\zeta_i, \zeta_{i+1}, \zeta_i, \zeta_{i+1}\rangle$, $|\zeta_i, \xi_i, \xi_{i+1}, \zeta_{i+1}\rangle$, $|\zeta_i, \xi_{i+1}, \xi_i, \zeta_{i+1}\rangle$, $|\xi_i, \xi_i, \xi_{i+1}, \xi_{i+1}\rangle$, $|\xi_i, \xi_{i+1}, \xi_i, \xi_{i+1}\rangle$, $|\xi_i, \zeta_i, \zeta_{i+1}, \xi_{i+1}\rangle$, $|\xi_i, \zeta_{i+1}, \zeta_i, \xi_{i+1}\rangle$. Notice that there are two types of symmetry in this palindromic set: first one is among accompanied purines and pyrimidines which are specified to be $\zeta_i \longleftrightarrow \zeta_{i+1}$ and $\xi_i \longleftrightarrow \xi_{i+1}$. Second one is among purines and pyrimidines in mixed forms, which is simply provided by $\zeta_i \longleftrightarrow \xi_i$. These symmetries are peculiar to the unitary feature of palindromic structures. We realize that first symmetry takes place between purine or pyrimidine structures such that certain cross mappings occurs. Hence, we simply call it a *pp-symmetry* (purine-pyrimidine symmetry). *pp-symmetry* is the natural symmetry that occurs usually in the any enzyme activity in transcription process of DNA. On the other hand, second symmetry relation is new and requires mapping a purine or pyrimidine phases into a its corresponding counterpart within the same phase, i.e. $\mathbf{A} \longleftrightarrow \mathbf{G}$ and $\mathbf{T} \longleftrightarrow \mathbf{C}$, that is why we call it *p-symmetry* (here, p implies purine or pyrimidine, i.e, symmetry of purine-purine or pyrimidine-pyrimidine nucleotides). *p-symmetry* is distinguished as a result of unitary structure of DNA such that important palindromic features can be attained because of this emergent symmetry relation. For instance, **TATA**-box has the accompanying *pp-symmetric* partner **ATAT** and *p-symmetric* partner **CGCG**-box, which is the minimum DNA binding element and an important palindromic sequence. The multiple **CGCG** cis-elements are found in promoters of genes such as those involved in ethylene signaling, abscisic acid signaling, and light signal perception. Diagram given in fig. 4 demonstrates the prescribed symmetry flows of singlet and doublet states.

FIG. 4: Palindromic mapping shows the flow of p- and pp-symmetries in singlet and doublet structures.

More precisely, in view of genetic perspective, it is well known that adenin transition to guanine and cytosine transition to thymine can be occurred in nature which leads to mutations. According to molecular dynamic simulations and ab initio experiments, these transitions are caused by energetic profiles of the purine and pyrimidine which are related to their different stabilities [4, 38, 39]. However, we suggest that their apparent structural similarities may be explained by means of their hidden symmetry relations as we extracted here, which are subsymmetry groups of an overall unitary symmetry in genome. In other words, repetitive palindromic sequence symmetry may be associated with adenin-guanine and thymine-cytosine structural similarities (see, Fig. 1) according to our symmetry model of unitary structure.

Moreover, it is observed that our suggested implications can be regarded as an extension to Chargaff's rules in which nucleic bases should have the same amount on each strand and have similar frequencies. Indeed, statistical studies approved these rules and each nucleic base has similar frequency at the opposite strand within DNA [40]. However, palindromic sequences have mirror symmetry in DNA, which is unique sequence-ordering because they present the same alignment in reverse order at the opposite strand. Thus, if we apply the Chargaff's parity rules to our novel approach, we can implicate that palindromic sequences with a symmetry embody a more general symmetry emphasized by Chargaff's rule. If we consider the functionality of the palindromic sequences, our proposed symmetries have purpose to controlling the conservation and/or accessibility of the genomic information, specifically help the transcriptional proteins reach the information to process [41]. With this study, we present an approach to explain the symmetrical nature of these repetitive sequences which are the part of another but more substantial symmetry in the genome of organisms. Besides, we figure out that the number of the palindromic sequence carriers (i.e., nucleic bases) can be controlled by our symmetry predictions. This provides certain palindromic structures to exist in nature, because nature allows only symmetric configurations in palindromes. As we indicate in our examples in Tables I and II, the number of palindromic sequence is drastically reduced once its sequential length increases. At that point we assume ideal conservation and accessibility of a gene region might be associated with sequential length of the palindrome. This could be related to energy considerations as well, but content of our work is beyond that level.

Table II demonstrates the number of incidence of quadruple palindromic structures corresponding to certain types of genes. Fig. 5 shows the symmetry flows of all palindromic mappings in quartet structures. Fig. 6 demonstrates the palindromic paths for binary and quadruplo structures given in Tables I and II. Notice the similar behaviors of certain genes.

Palindromic Sequence	<i>Catenin</i>	<i>PMEPA</i>	<i>P50</i>	<i>NKILA</i>	<i>P65</i>
AATT>	481	193	847	13	57
GGCC>	187	466	199	12	126
CCGG>	50	188	47	15	57
TTAA>	502	178	922	15	52
ATAT>	382	107	708	6	23
GCGC>	48	119	35	7	43
CGCG>	23	59	8	3	24
TATA>	353	85	660	1	24
AGCT>	319	357	467	14	104
GATC>	130	133	226	4	47
CTAG>	169	168	335	6	45
TCGA>	27	51	50	1	12
TGCA>	312	382	457	11	76
CATG>	279	286	488	3	62
GTAC>	132	89	238	1	21
ACGT>	50	115	92	3	10

TABLE II: Table depicts the numbers of quadruple palindromic sequences in special genes of *Catenin*, *PMEPA*, *P50*, *NKILA* and *P65*. These values are extracted from FASTA sequences obtained through NCBI database.

As can be seen from this symmetry table implying palindromic symmetries, each category contains rather crucial palindromic sequences. Another important example of short palindromic sequences in this list is 5'-**GATC**-3' regions in the E. coli genome. Mismatch nucleotide is a type of mutation in the organisms and the mismatch repair (MMR) mechanism uses the **GATC** sequences to find and repair the damaged regions on the DNA [33]. Many enzymes taking part in this complicated molecular repair mechanism. **GATC** sequences are initiators to assemble this complex molecular machinery. First, in this repair process **GATC** sequence is methylated on the leading strand (5' to 3' strand) and after this **GATC** is recognized by relevant enzymatic proteins that are capable of recognizing the breakings or mismatch regions on the DNA. **GATC** sequences are the initiating points for the MMR protein repair process, through this MMR proteins can scan the DNA strand and able to find mismatch base pair then the exonucleases begin their replacement of the mutated region [34]. Additionally, **GATC** sequences repeat themselves in every 256 nucleotides through the prokaryotic genome to facilitate the MMR proteins to scan the genome more efficiently [35]. However, in humans, **GATC** may represent methylation sites to provide chromosomal stability and control

FIG. 5: (Color Online) Diagram displays all palindromic mappings corresponding to the flow of p- and pp-symmetries in quartet structures.

FIG. 6: (Color online) Figures depict the palindromic paths for binary (left panel) and quadruple (right panel) structures provided in Tables I and II. Palindromic sequences and corresponding symmetry relations are not shown explicitly in quadruple structures because they fill the whole axis. In these plots, pink- and gray-colored zones correspond to p-symmetry transitions whereas green-colored zones correspond to pp-symmetry transition regions. Notice the similar behaviors of *Catenin* and *P50* genomes and that of *PMEPA*, *P50* and *NKILA*.

the protein synthesis [36]. In other words, **GATC** sequence presents a good methylation site in the organisms and its frequency attenuates the recognition by the relevant proteins. This methylation mechanism is very important to regulate genomic information [37].

Another consequence of p- and pp-symmetries is that quadruple structures can be obtained by employing doublets via these symmetries. For instance, $|\zeta_i, \zeta_i, \zeta_{i+1}, \zeta_{i+1}\rangle$ state can be expressed by doublets as $|\zeta_i, \zeta_i, \zeta_{i+1}, \zeta_{i+1}\rangle = |\zeta_i, \zeta_i\rangle \otimes |\zeta_{i+1}, \zeta_{i+1}\rangle$. pp-symmetry converts this quartet into its counterpart as $|\zeta_{i+1}, \zeta_{i+1}, \zeta_i, \zeta_i\rangle = |\zeta_{i+1}, \zeta_{i+1}\rangle \otimes |\zeta_i, \zeta_i\rangle$. In other words, pp-symmetry acts on doublets similar to single states, i.e. $|\zeta_i, \zeta_i\rangle \xrightarrow{pp\text{-symmetry}} |\zeta_{i+1}, \zeta_{i+1}\rangle$ and $|\zeta_{i+1}, \zeta_{i+1}\rangle \xrightarrow{pp\text{-symmetry}} |\zeta_i, \zeta_i\rangle$. Likewise, p-symmetry acts on this state to give rise to its accompanying symmetric state: $|\zeta_i, \zeta_i\rangle \xrightarrow{p\text{-symmetry}} |\xi_i, \xi_i\rangle$ and $|\zeta_{i+1}, \zeta_{i+1}\rangle \xrightarrow{p\text{-symmetry}} |\xi_{i+1}, \xi_{i+1}\rangle$. Thus, we obtain $|\zeta_i, \zeta_i, \zeta_{i+1}, \zeta_{i+1}\rangle \xrightarrow{p\text{-symmetry}} |\xi_i, \xi_i, \xi_{i+1}, \xi_{i+1}\rangle$ in quartet state. It is understood that one can generalize p- and pp-symmetries into doublets and higher order palindromic structures. Although there is no mixed forms of purines and pyrimidines in palindromic doublet states, quartet palindromic sequences allow mixed structures of doublet states. Therefore, in order to obtain all quartet palindromic sequences, one can consider all mixed combinations of $|\zeta_i, \zeta_i\rangle$,

$|\zeta_i, \xi_i\rangle$, $|\xi_i, \zeta_i\rangle$ and $|\xi_i, \xi_i\rangle$, which yields the number of distinct categories given in Table (II). In view of these symmetry relations, it seems that operator descriptions of quartet palindromic structures are more accessible.

Sextet structures deserve a special attention because of their crucial role in the transfer of important information that ensures the continuity of cell cycle. All features in view of symmetry operations can be extended to this notable palindromic structures. One can build up a sextet system by means of individual nucleotide basis of pure $|\zeta, \zeta, \zeta, \zeta, \zeta, \zeta\rangle$ states (i.e., $|\zeta_i, \zeta_i, \zeta_i, \zeta_{i+1}, \zeta_{i+1}, \zeta_{i+1}\rangle$, $|\zeta_i, \zeta_i, \zeta_{i+1}, \zeta_i, \zeta_{i+1}, \zeta_{i+1}\rangle$, $|\zeta_i, \zeta_{i+1}, \zeta_{i+1}, \zeta_i, \zeta_i, \zeta_{i+1}\rangle$, $|\zeta_{i+1}, \zeta_{i+1}, \zeta_{i+1}, \zeta_i, \zeta_i, \zeta_i\rangle$, $|\zeta_{i+1}, \zeta_i, \zeta_i, \zeta_{i+1}, \zeta_{i+1}, \zeta_i\rangle$, $|\zeta_i, \zeta_{i+1}, \zeta_i, \zeta_{i+1}, \zeta_i, \zeta_{i+1}\rangle$, $|\zeta_{i+1}, \zeta_{i+1}, \zeta_i, \zeta_i, \zeta_i, \zeta_{i+1}\rangle$, $|\zeta_{i+1}, \zeta_i, \zeta_{i+1}, \zeta_i, \zeta_{i+1}, \zeta_i\rangle$), pure $|\xi, \xi, \xi, \xi, \xi, \xi\rangle$ states which are formed by the exchange of $\zeta \leftrightarrow \xi$ and mixture of these two states. It is obvious that there are $2^6 = 64$ palindromic family of states. But notice that these peculiar states could be established by employing doublet basis. All mixed combinations of doublet states give rise to construct any sextet state. Besides, one can also use triplets as the basis of sextet structures. All their combinations consist of the members of the set: $|\zeta, \zeta, \zeta\rangle$, $|\zeta, \zeta, \xi\rangle$, $|\zeta, \xi, \zeta\rangle$, $|\zeta, \xi, \xi\rangle$, $|\xi, \zeta, \zeta\rangle$, $|\xi, \zeta, \xi\rangle$, $|\xi, \xi, \zeta\rangle$, $|\xi, \xi, \xi\rangle$. Obviously, any sextet state can be expressed as mixed combinations of doublets or triplets, this is because 6 can be factorized as $6 = 2 \cdot 3$. But for any quartet state this is $4 = 2 \cdot 2 = 2^2$. It is understood that we can obtain a formalism to construct higher order palindromic sequences by means of smaller sequences according to nucleotide basis of prime integer numbers that corresponding palindromic sequence contains. If we express the palindromic sequence by the number N , we can decompose it in terms of n prime integers as $N = P_1^{m_1} P_2^{m_2} \dots P_n^{m_n}$, where P_i is the prime integer and m_i is the corresponding power of P_i that is an integer. Therefore, one can describe a palindromic sequence with N individual basis in terms of P_i sub-basis. This could be done by n different ways so that corresponding palindromic state preserves its unitary structure. Figure 7 shows the symmetry diagram of **CACGTG** palindrome in sextet, doublet and triplet basis.

FIG. 7: Diagram displays all palindromic mappings of **CACGTG** sequence in singlet, doublet and triplet basis corresponding to the flow of p- and pp-symmetries in sextet structures.

Notice that although one can construct higher order palindromic sequences in principle, their appearance in any genome are so rare because their probability of getting together is very low. This could be explained by energy considerations in genomic sequence. But, this requires more elaborate treatment of the subject which is beyond the scope of this work.

CONCLUDING REMARKS

In this study, we aim to reveal the quantum effects of palindromic sequences in DNA in view of symmetry perspectives, especially unitary symmetry. Unitary symmetry basically points out the preserving important information of DNA inherent to the essence of continuity of cell cycle. Our approach exhibits the existence of subsymmetry groups of unitary symmetry in palindromes, which are called p- and pp-symmetries. pp-symmetry has been the commonly used natural symmetry of DNA known as Chargaff's rules. However, p-symmetry is also found in palindromes, which contribute to the unitary symmetry. By means of these symmetry relations, we can better understand the structures and formations of higher order palindromes and why they get rare with the increasing number of sequences.

Dirac notations describing the basic hallmarks of the states of palindromes stipulate new qualifications of these peculiar configurations in DNA, especially in view of the quantum aspects. It is understood that subsymmetry groups of unitary structure specify the role played by palindromes. Their distinguished behaviors are strictly controlled by symmetry rules dictated by the quantum features. It is revealed that the more we understand the symmetry properties of palindromes, the more we can learn about palindromes. Our findings propose that rare number of sequences for higher order palindromes in various genes can be explained by the low probability of the bases forming palindromes

to coalesce together. We find out that relevant bases of higher order palindromes can be constructed by any prime integer sequences that generate the palindrome.

Our approach in this study will help to reveal the further quantum nature of DNA. Palindromic examples we presented here are oncogenic associated gene sequences and contain many palindromic sequences in their genomic DNA regions. DNA molecules contain enormous amount of data which requires high level of conservation within and/or among the organisms to prevent mutations. Palindromic sequences may present unique features and has some advantages like easier enzymatic process. However, we suggest that palindromic sequences require more conservation or stabilization than other type of sequence and their symmetrical structures render them quintessential points in the manner of data transfer plus to enzymatic regulation and breaking points during the replicative process within cell division.

Here, we suggest symmetrical nature of these specific sequences renders the DNA more conservative in these points but any breaking or damaging this symmetry cause mostly carcinogenic or mutagenic results. As mentioned in the other sections in this study, we reveal the quadruplet and sextet palindrome sequence numbers in the *PMEPA*, *NKILA*, *P65*, *P50*, and β -*Catenin* genomic regions on the DNA. These molecules play important roles on the oncogenic signaling pathways and have association between them [42–44]. Quadruplet structures are more common while sextet structures more scarce. Therefore, we suggest that, longer palindromic structures have more specified roles within the coding regions on the DNA and governing the conservative significance of the sequence which it is belong.

We need to elaborate these structures by considering the mutation models which may shed light on the mutagenic and oncogenic mechanism of the mutation in point of symmetrical breaking view. Also, we need to interpret more genomic regions as much as possible in order to understand more clearly whether the palindromic sequence lengths associate with the power of conservation. Explaining the mutagenic models in Dirac notations may direct us to interpret the difference between asymmetrical and symmetrical structures of the DNA regions. Finally, with the approach we proposed in this work, we can better understand whether viral pathogenes including coronavirus has a symmetrical structure in their genome and whether they recognize palindromes in DNA easily. This can help us better understand their nature and find effective ways to avoid them.

Acknowledgement

We acknowledge the support of Scientific Research Projects Coordination Unit of Istanbul University (IU BAP) under the project number: FDK-2017-25162.

Author Contributions

Both authors contributed equally to the whole work including calculations, analysis and manuscript typing.

Competing Interests

The authors declare no competing interests.

* Electronic address: matibatan@gmail.com

† Electronic address: mustafa.sarisaman@istanbul.edu.tr

- [1] S. Usha and S. Selvaraj, “Structure-wise discrimination of adenine and guanine by proteins on the basis of their nonbonded interactions,” *J. Biomol. Struct. Dyn.*, **33**(7), 1474-1492 (2015).
- [2] M. Barbatti, A. J. Aquino and H. Lischka “The UV absorption of nucleobases: semi-classical ab initio spectra simulations,” *Phys. Chem. Chem. Phys.*, **12**(19), 4959-4967 (2010).
- [3] M. K. Shukla and J. Leszczynski, “Tautomerism in nucleic acid bases and base pairs: a brief overview,” *WIREs Comput. Mol. Sci.* **3**, 637-649 (2013).
- [4] R. Wu and T. B. McMahon, “Investigation of proton transport tautomerism in clusters of protonated nucleic acid bases (cytosine, uracil, thymine, and adenine) and ammonia by high-pressure mass spectrometry and ab initio calculations,” *J. Am. Chem. Soc.* **129**(3), 569-580 (2007).
- [5] M. Berry, A. M. Nunez and P. Chambon, “Estrogen-responsive element of the human pS2 gene is an imperfectly palindromic sequence,” *Proc. Natl. Acad. Sci. USA*, **86**(4), 1218-22 (1989).

- [6] D. R. Leach, E. A. Okely and D. J. Pinder, "Repair by recombination of DNA containing a palindromic sequence," *Mol. Microbiol.* **26**(3), 597-606 (1997).
- [7] S. Iho, T. Yamamoto, T. Takahashi and S. Yamamoto, "Oligodeoxynucleotides containing palindrome sequences with internal 5'-CpG-3' act directly on human NK and activated T cells to induce IFN-gamma production in vitro," *J. Immunol.* **163**(7), 3642-52 (1999).
- [8] T. Yamamoto, S. Yamamoto, T. Kataoka and T. Tokunaga, "Lipofection of synthetic oligodeoxyribonucleotide having a palindromic sequence of AACGTT to murine splenocytes enhances interferon production and natural killer activity," *Microbiol. Immunol.* **38**(10), 831-6 (1994).
- [9] V. Raykov, M. E. Marvin, E. J. Louis and L. Maringele, "Telomere Dysfunction Triggers Palindrome Formation Independently of Double-Strand Break Repair Mechanisms," *Genetics*, **203**(4), 1659-68 (2016).
- [10] D. K. Nag and A. Kurst, "A 140-bp-long palindromic sequence induces double-strand breaks during meiosis in the yeast *Saccharomyces cerevisiae*," *Genetics*, **146**(3), 835-47 (1997).
- [11] Y. F. Huang, S. C. Chen, Y. S. Chiang, T. H. Chen and K. P. Chiu, "Palindromic sequence impedes sequencing-by-ligation mechanism," *BMC Syst. Biol.* **6**, Suppl 2:S10 (2012).
- [12] D. Cain, O. Erlwein, A. Grigg, R. A. Russell and M. O. McClure, "Palindromic sequence plays a critical role in human foamy virus dimerization," *J. Virol.* **75**(8), 3731-9 (2001).
- [13] J. Guenthoer, S. J. Diede, H. Tanaka, X. Chai, L. Hsu, S. L. Tapscott, et al, "Assessment of palindromes as platforms for DNA amplification in breast cancer," *Genome Res.* **22**(2), 232-45 (2012).
- [14] J. Sponer, J. Leszczynski and P. Hobza, "Electronic properties, hydrogen bonding, stacking, and cation binding of DNA and RNA bases," *Biopolymers.*, **61**(1), 3-31 (2001).
- [15] R. Alex, O. Sozeri, S. Meyer and R. Dildrop, "Determination of the DNA sequence recognized by the bHLH-zip domain of the N-Myc protein," *Nucleic Acids Res.* **20**(9), 2257-63 (1992).
- [16] Y. Chen and O. I. Olopade, "MYC in breast tumor progression," *Expert Rev. Anticancer Ther.* **8**(10), 1689-98 (2008).
- [17] H. Chen, H. Liu and G. Qing, "Targeting oncogenic Myc as a strategy for cancer treatment," *Signal Transduct Target Ther.* **3**, 5 (2018).
- [18] Y. Fallah, J. Brundage, P. Allegakoen and A. N. Shajahan-Haq, "MYC-Driven Pathways in Breast Cancer Subtypes," *Biomolecules*, **7**(3), 53 (2017).
- [19] L. C. Lustig, D. Dingar, W. B. Tu, C. Lourenco, M. Kalkat, I. Inamoto, et al, "Inhibiting MYC binding to the E-box DNA motif by ME47 decreases tumour xenograft growth," *Oncogene*, **36**(49), 6830-7 (2017).
- [20] G. Senjanovic and R. N. Mohapatra, "Exact left-right symmetry and spontaneous violation of parity," *Phys. Rev. D*, **12**(5), 1502 (1975).
- [21] R. N. Mohapatra and J. C. Pati, "Natural" left-right symmetry," *Phys. Rev. D*, **11**(9), 2558-2561 (1975).
- [22] M. Peskin and D. Schroeder, "An Introduction to Quantum Field Theory," Westview Press, (1995).
- [23] K. Kiran, S. A. Ansari, R. Srivastava, et al., "The TATA-Box Sequence in the Basal Promoter Contributes to Determining Light-Dependent Gene Expression in Plants," *Am. Soc. of Plant Biologist*, **142**, 364-376 (2006).
- [24] R. Tjitro, L. A. Campbell, L. Basova, et al., "Modeling the Function of TATA Box Binding Protein in Transcriptional Changes Induced by HIV-1 Tat in Innate Immune Cells and the Effect of Methamphetamine Exposure," *Frontiers in Immunology*, **9**, 3110 (2019).
- [25] J. Schwinger. *Quantum Mechanics, Symbolism of Atomic Measurements*. Springer-Verlag Berlin Heidelberg, (2001).
- [26] J. Schwinger, "The Algebra of Microscopic Measurement," *Proceedings of the National Academy of Sciences of the United States of America*, **45**(10), 1542-1553, (1959).
- [27] J. von Neumann, "Mathematical Foundations of Quantum Mechanics," Princeton University Press, (1955).
- [28] U. Fano, "Description of States in Quantum Mechanics by Density Matrix and Operator Techniques," *Reviews in Modern Physics*, **29**, 74 (1957).
- [29] P. A. M. Dirac, "Principles of Quantum Mechanics," Oxford university Press, Oxford, (1958).
- [30] C. Cohen-Tannoudji, B. Diu and F. Laloë, "Quantum Mechanics," Wiley, Three volumes, (1977).
- [31] R. P. Feynman, "The Feynman Lectures on Physics" 3 (2nd ed.). Addison-Wesley, (2005).
- [32] M. Grace, M. B. Chandrasekharan, T. Hall, et al., "Sequence and Spacing of TATA Box Elements Are Critical for Accurate Initiation from the β -Phaseolin Promoter," *J. of Biological Chem.* **279**/9, 8102-8110 (2004).
- [33] G. M. Li, "Mechanisms and functions of DNA mismatch repair," *Cell Res.* **18**, 85-98 (2008).
- [34] K. Sneppen, and S. Semsey, "Mismatch repair at stop codons is directed independent of GATC methylation on the *Escherichia coli* chromosome," *Sci. Rep.* **4**, 7346 (2014).
- [35] M. G. Marinus and A. Løbner-Olesen, "DNA methylation," *EcoSal Plus*, **6**(1), (2014).
- [36] H. Xie, M. Wang, M. Bonaldi, V. Rajaram, W. Stellpflug, C. Smith, et al., "Epigenomic analysis of Alu repeats in human endometrium," *Proceedings of the National Academy of Sciences*, **107**(15), 6952-6957, (2010).
- [37] J. A. Shapiro and R. von Sternberg, "Why repetitive DNA is essential to genome function," *Biol. Rev. Camb. Philos. Soc.* **80**(2), 227-50 (2005).
- [38] S. Tolosa, J. A. Sansón and A. Hidalgo, "Theoretical Study of Adenine to Guanine Transition Assisted by Water and Formic Acid Using Steered Molecular Dynamic Simulations," *Frontiers in Chemistry*, **7**, 414 (2019).
- [39] S. Tolosa, J. A. Sansón, A. Hidalgo, "Structural and thermodynamic studies of cytosine to thymine conversion in gas and solution phases using steered molecular dynamic simulations," *Journal of Molecular Liquids*, **278**, 61-69 (2019).
- [40] G. Cristadoro, M. Degli Esposti and E. G. Altmann, "The common origin of symmetry and structure in genetic sequences," *Sci. Rep.* **8**, 15817 (2018).
- [41] Y. Oh, C. Lee, K. Baek, W. Kim, J. Yoon, K. Han and N. Cho, "An element with palindromic structure is required for

- the expression of TBP (TATA box-binding protein) gene in *Drosophila melanogaster*,” *Mol. Cells*, **9(6)**, 673-677 (1999).
- [42] M. A. Tibatan and E. Ö. Uçar, “NKILA regulates HSP90 α , NF- κ B subunits and β -catenin expressions in the MCF-7 cells,” *Molecular Biology Reports*, **48**, 4563-4571 (2021).
- [43] J. Deng, S. A. Miller, H. Y. Wang, W. Xia, Y. Wen, B. P. Zhou, Y. Li, S. Y. Lin, and M. C. Hung, “Beta-Catenin Interacts with and Inhibits Nf-Kappa B in Human Colon and Breast Cancer,” *Cancer Cell* **2(4)**, 323-34 (2002).
- [44] J. M. Dijkstra, and D. B. Alexander, “The ‘ Nf-K B Interacting Long Noncoding Rna’ (Nkila) Transcript Is Antisense to Cancer-Associated Gene Pmepa1,” *F1000Research*, **4**, 96 (2015).
- [45] These are explicitly $|AA\rangle, |AT\rangle, |TA\rangle, |TT\rangle, |GG\rangle, |GC\rangle, |CG\rangle, |CC\rangle, |AG\rangle, |AC\rangle, |TG\rangle, |TC\rangle, |GA\rangle, |GT\rangle, |CA\rangle$ and $|CT\rangle$.