BAGHAEE RAVARI, SOODEH, Ph.D. Ionic Modulation of QPX Stability as a Nano - Switch Regulating Gene Expression in Neurons. (2016).
Directed by Dr. Ethan Will Taylor. 120 pp.

G-quadruplexes (G-QPX) have been the subject of intense research due to their unique structural configuration and potential applications, particularly their functionality in biological process as a novel type of nano–switch. They have been found in critical regions of the human genome such as telomeres, promoter regions, and untranslated regions of RNA. About 50% of human DNA in promoters has G-rich regions with the potential to form G-QPX structures. A G-QPX might act mechanistically as an ON/OFF switch, regulating gene expression, meaning that the formation of G-QPX in a single strand of DNA disrupts double stranded DNA, prevents the binding of transcription factors (TF) to their recognition sites, resulting in gene down-regulation. Although there are numerous studies on biological roles of G-QPXs in oncogenes, their potential formation in neuronal cells, in particular upstream of transcription start sites, is poorly investigated. The main focus of this research is to identify stable G-QPXs in the 97bp active promoter region of the choline acetyltransferase (ChAT) gene, the terminal enzyme involved in synthesis of the neurotransmitter acetylcholine, and to clarify ionic modulation of G-QPX nanostructures through the mechanism of neural action potentials. Different bioinformatics analyses (in silico), including the QGRS, quadparser and G4-Calculator programs, have been used to predict stable G-QPX in the active promoter region of the human ChAT gene, located 1000bp upstream from the TATA box. The results of computational studies (using those three different algorithms) led to the

identification of three consecutive intramolecular G-QPX structures in the negative strand (ChAT G17-2, ChAT G17, and ChAT G29) and one intramolecular G-QPX structure in the positive strand (ChAT G30). Also, the results suggest the possibility that nearby G-runs in opposed DNA strands with a short distance of each other may be able to form a stable intermolecular G-QPX involving two DNA complementary strands (ds ChAT G21).

Formation of G-QPX structures, by blocking the availability of the transcription factor binding site (TFBS) on double stranded DNA, can interfere with transcriptional activation. This suggests that there is competition between TFBS binding to dsDNA and the conversion to high order non-B form secondary structures (G-QPXs) in the active promoter region. TFBS mapping analysis of the active promoter region of the human ChAT gene revealed that it contains multiple consensus AP-2α and Sp1 binding sites and consensus sites for other TF, including multiple sites for GR-alpha, Pax-5, p53 and GC box proteins.

To get a better understanding of how modulation of G-QPX structures might affect the ChAT promoter activity, an artificial GFP reporter vector (modified GFP) was constructed, synthesized and used for reporter gene measurement. As known human ChAT promoter activators, nerve growth factors (HNGF and TGB) and cytokines (IL-β and TNF-α) were used for activation of the artificial promoter driving GFP. Also, the G-QPX stabilizing drug TMPYP4 and aconitine, a Na+ channel opening drug, were used as G-QPX stability modulating factors. It was observed that aconitine potentiated the action of the transcriptional activator NGF, suggesting that the effect of sodium is contrary to

that of TMPYP4, i.e., that an increase in promoter activity may be due to instability of G-QPX structures in a high Na+ environment, which results in melting these structures, enabling dsDNA formation required for the binding of TF to their recognition sites for initiation of transcription. The results were confirmed in several independent sets of experiments, using GFP reporter gene measurement by plate reader, by flow cytometry and using fluorescent microscopy.

Moreover, quantitative RT-PCR was conducted to evaluate the effect of the same factors under similar conditions on the actual ChAT mRNA expression. It was observed that TMPY4 knocked down the ChAT mRNA expression by 87%, suggesting that G-QPX stabilization inhibits promoter activity as expected and that aconitine along with HNGF increases ChAT mRNA expression up to 2.8 fold. Aconitine-mediated influx of Na+ ions, possibly by inhibiting the formation of stable G-QPX structures, resulted in an

Unique G-QPX structures can be stabilized with certain metal cations or small cationic molecule ligands such as TMPYP4, through occupying the space between the layers of G-tetrads. Although G-QPX are reported to have high stability in potassium solution, the diversity of G-QPX structures (due to diversity in sequence and size of G-runs, sequence and size of loops) will lead to diversity in physical behavior of G-QPX structures. Therefore, to get a clear image of folding topology and stability of identified G-QPX structures, physical studies including CD spectroscopy and AFM imaging were conducted. CD results showed that the identified ChAT G-QPX structures formed a hybrid, stable configuration in potassium environment (10mM) while being instable in sodium solution (100mM). AFM imaging demonstrated star-shaped structures (involving

clusters of DNA strands) due to incubation with TMPYP4, where a greater number of these G-rich sequences have converted to G-QPX structures.

The results of both an artificial engineered reporter gene system and actual ChAT mRNA expression (in vitro), plus physical characterization studies, strongly support the novel hypothesis that a neural action potential ionic mechanism regulates G-QPX formation/ deformation in the promoter region, due to movement of monovalent cations across the membrane, which is consistent with gene silencing and expression during neuronal resting and firing.

IONIC MODULATION OF QPX STABILITY AS A NANO-SWITCH REGULATING GENE EXPRESSION IN NEURONS

by

Soodeh Baghaee Ravari

A Dissertation Submitted to
the Faculty of The Graduate School at
The University of North Carolina at Greensboro
in Partial Fulfillment
of the Requirement for the Degree
Doctor of Philosophy

Greensboro
2016

Approved by

_____
Committee Chair

DEDICATION

To those whose efforts led to advancement in sciences and technologies

And

To my loving husband, Mahdi for always being there

APPROVAL PAGE

This dissertation written by Soodeh Baghaee Ravari has been approved by the following committee of the Faculty of The Graduate School at The University of North Carolina at Greensboro.

Committee Chair _____
Ethan Will Taylor, Ph.D.

Committee Members _____
James G. Ryan, Ph.D.

_____
Dennis R. LaJeunesse, Ph.D.

_____
Christopher Kepley, Ph.D.

7/1/2016
Date of Acceptance by Committee

7/1/2016
Date of Final Oral Examination

ACKNOWLEDGEMENTS

I would like to thank my advisor Dr. Ethan Will Taylor for giving me the opportunity to expand my knowledge. I am thankful for his guidance and suggestions which have helped me throughout the time I was part of his team.

I would like to thank Dr. James G. Ryan, Dean of Joint School of Nanoscience and Nanoengineering and one of my advisory committee members for his guidance. I would like to thank Dr. Dennis R. LaJeunesse and Dr. Christopher Kepley for being part of my work and serving in my advisory committee.

I am thankful to my parents and husband who have encouraged and supported me during my studies.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER I

INTRODUCTION TO G-QUADRUPLEX STRUCTURES

This dissertation is focused on identification of G-quadruplex (G-QPX) nanoswitches on the promoter region of the human ChAT gene, and the assessment of their effect on regulating cellular process, in particular transcriptional activity under varied conditions. This chapter provides some background of G-QPX motifs formation, stability and their diversity in the whole genome, and explains neural communication and neural signaling.

1. 1 G-QPX Structures, Formation, Stability and Diversity

The G-QPX structure is nucleic acid DNA/ RNA non B form secondary structure, containing guanines self-associate (G-tetrad). The building blocks of G-QPX are two or more planar tetrad which stacks by hydrogen binding. The planar tetrad is formed by association of four guanine bases through Hoogsteen binding and their arrangements create a central cavity. Since the central cavity of tetrad has onionic charge due to existence of four oxygen atoms (O6), cationic charge is essential for stabilizing the structure (Fig.1). It has been suggested that cationic molecules and metal cations such as K^+ and Na^+ acts as strong molecules in stabilizing G-QPX structures. However, in most

G-QPX studies it has been observed that K^+ solution is highly favor G-QPX formation and stability in compare to other metal cations such as Na^+ and Li^+ [1]. Depending on the orientation of the strands in G-QPX structures, they can be described as parallel, antiparallel, and hybrid (Fig.1) [2, 3]. G-runs are connecting through the loops which could be varied in length and sequence. Typically, the loops length is between 1-9 base pairs but they could be long up to 26 base pairs [4].

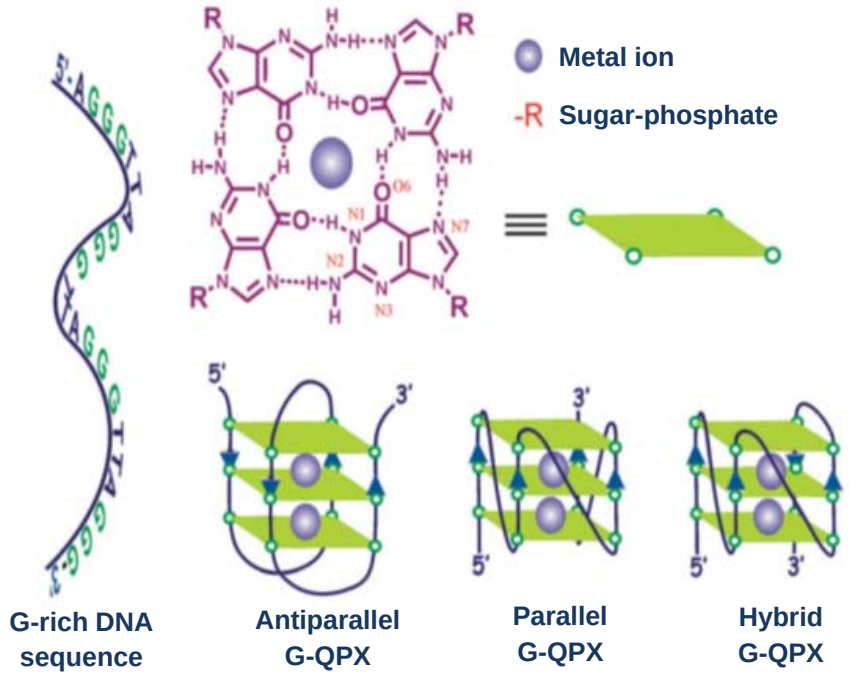


Figure 1. Structure of G-QPXs and their Folding Topologies [5].

G-QPX structures are polymorphic which can form various topologies as a combination of different number of nucleic acid strands. They can form intermolecular or intramolecular structures. Intramolecular G-QPX requires one single strand containing four or more runs of guanine which linked by loops (Fig.2-A), whereas intermolecular G-

QPX structures can arise from involvement of G-runs from multiple strands, two or four strands (Fig. 2-B). It has been observed that most of the G-QPXs of genome fold into intramolecular structures within double stranded (ds) DNA (Fig. 2-C) [2]. However, recently it has been reported hybrid G-QPX structure can form between DNA and RNA strands (Figure2.D) [6].

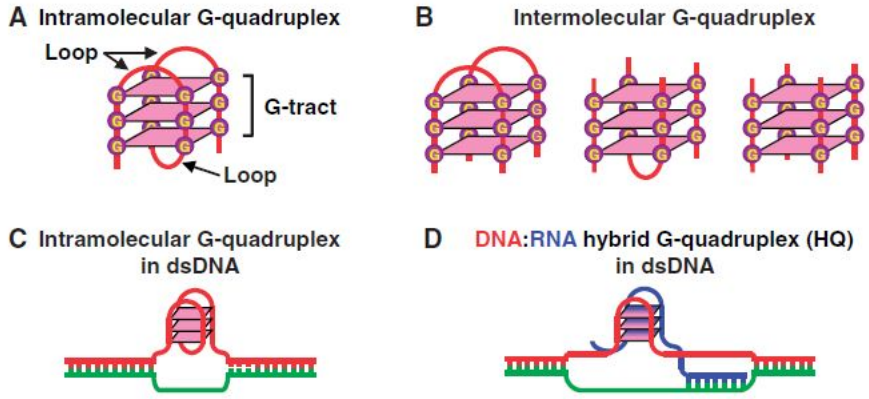


Figure 2. Different Examples of G-QPX Structures. A) An intramolecular G-QPX with three loops and three tetrad layers. B) Intermolecular G-QPX composed of different number of nucleic acids strands (two, three and four). C) An intramolecular G-QPX within double stranded DNA (ds DNA). D) Intermolecular DNA: RNA hybrid G-QPX in double stranded DNA (ds DNA) [6].

The structural diversity of G-QPX with different folding topology is extended into the following parameters: variation in the sequence of G-runs and G-run length, variation in the number of oligonucleotide strands involved and their orientations, variation in the length and sequence of the loops, and finally environmental factors such as presence of small cationic ligands molecules or metal cations. The combination of these parameters brings out the behavior complexity of distinct G-QPX structure towards its ligands binding affinity, stability and its functionality in biological process. Several biophysical

methods have been used to characterize G-QPX structures such as X- ray diffraction (XRD), circular dichroism (CD), and nuclear magnetic resonance (NMR). One of the common techniques is CD spectroscopy, which directly monitor molecules with chiral structures like G-QPX under various environmental conditions such as presence of metal cations, small binding molecules, different temperature and pH. Generally, CD spectra demonstrates the conformations of DNA, for instance a CD spectrum with a negative peak of 265nm and a positive peak of 295nm described as antiparallel structure, whereas a negative and positive peak at 240 and 260, respectively described as parallel structure [7]. Small DNA binding molecules and metal cations have a critical impact on conformational transition of G-QPX. Indeed, the type of monovalent cations and their concentration has a large effect on stabilization/destabilization of G-QPXs.

1.2 Potential G-QPX in the Human Genome

G-QPX structures have been shown to exist in the critical region of the human genome. Different computational methods (in silico analysis) support existence of G-rich regions in the human genome. Attention in study of G-QPX motifs is rising due to their important role in cellular functions [8-10]. It has been extensively suggested that these unique nanostructures are located in promoter regions, telometric region, and 5` untranslated regions [11]. Some of these sites may lead to genome stability due to altering DNA/RNA architecture, resulting interference in the normal physiological and pathological processes like replication, transcription and translation. They also have potential to cause genomic instability which lead to aging, and various genetic diseases.

Recently, it has been reported that about G-QPX structures exist in about 90% of human DNA replication and frequently near the origin of replication (Fig. 3 A-C). DNA analysis of different cancer types have shown that G-QPX has high potential formation in telomerase as being a novel target for anti-aging [12]. Moreover, G-QPX has been observed in the 5` UTR of mRNAs which repress the translation activity (Fig.3-D) [11, 13].

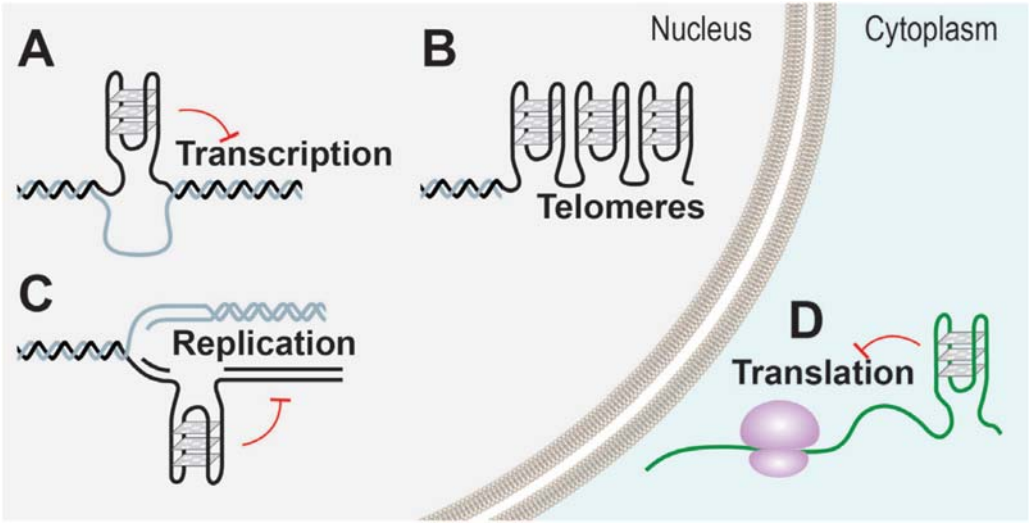


Figure 3. Possible Locations of Formation of G-QPX Motifs in Cells. Genome sequencing analysis revealed that G-QPX forms in the critical regions of the genome, distributed to promoters, 5`untranslated regions and telomeres. In the nucleus, G-QPXs can form in G-rich sequences when double stranded DNA separated and single stranded become free to form unique conformation, during A) transcription, C) replication and B) in telemetric regions. In cytoplasm, outside of nucleus, G-QPXs can form at the mRNA level D) and regulate translation activity [14].

1.2.1 G-QPX in transcription regulation

About 50% of the human genes have G-rich regions in their promoter to form G-QPX motifs, resulting in gene expression regulation (Fig.4) [14]. The first G-QPX in the promoter region, has been found on promoter of C-MYC oncogenes and it has been

demonstrated that addition of G-QPX stabilizing ligands such as TMPYP4 significantly affect transcription, in vivo system [15]. Interestingly, G-QPX structures mostly observed on the promoter of regulatory genes or oncogenes rather than tumor repressor gene [16].

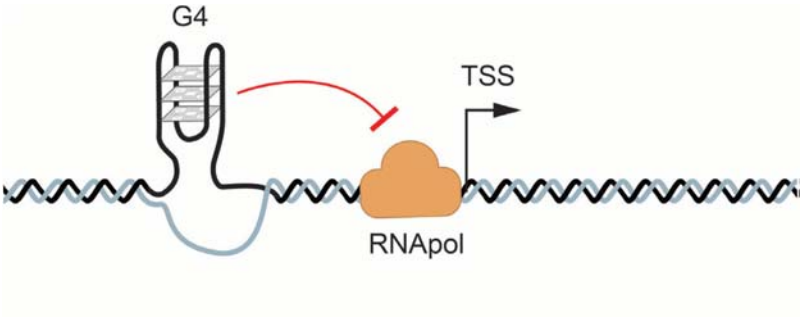


Figure 4. G-QPXs in Transcription. G-QPX exists in approximately in about 50% of the human genes in promoter regions. Formation of G-QPXs prevents paring of RNA polymerase and disrupt transcription initiation [14].

1.2.2 Targeting DNA G-QPX scaffolds with small ligands molecules

As mentioned before, the planner G-tetrad creates a negatively polarized central cavity which can be occupied by metal cations or small cationic ligand molecules. Hence, cationic charge is essential in G-QPX stabilization. Among metal cations used to study G-QPX stability, K^+ ion is known to be one of the most effective ones. Additionally, interaction of cationic small ligand molecules with the planner tetrad via π-π stacking and electrostatic interactions, can lead to form a highly stabilize G-QPX structures [17]. The structural feature and the side chains of the ligand allow the molecule to bind to G-QPX with high affinity. For instance, porphyrin (TMPYP4) highly binds to the human telomeric G-QPX or the human MET promoter, resulting an enhancement of the stability of G-QPX formation [18, 19]. Although TMPYP4 is known as a typical small ligand

molecule in stabilizing of various G-QPX structures in different region of human genome, there are other ligands that have been identified as a G-QPX stabilizing compound such as the perylenes PIPER, BRACO-19, RHSP4 and several tetrasubstituted naphthalene diimide [20-24]. In addition, ligands can intercalate between the G-tetrads through bindings to the grooves of the G-QPXs and forming a class of G-QPX stabilizing agents with high selectivity to G-QPX structures. For instance, SYUIQ-FM05 as a derivative of quindoline has been used to stabilize the specific G-QPX in the promoter region of BCL2, regulating transcriptional activity through blocking the binding of wt1 protein (Wilms' tumor suppressor gene) to a G-rich region in the promoter [25]. In another word, ligands that recognize and bind to G-QPXs are promising new strategies in anti-cancer drug development. Importantly, the topology and loop length of G-QPX are varying from one to another so the binding of small molecule ligands to G-QPX motifs should vary. Therefore, considering the ligand interactions with different types of loops, grooves and structural folding is important in utilizing G-QPX role in drug design.

1.3 Introduction to Neurons and Neurotransmitters

1.3.1 Neuronal communications

The neuron or nerve cell is the key component in the structure and function of the nervous system. It transfers information through electrochemical signals. The neuron is composed of three components which are the dendrites, the cell body and the axon. Dendrites are thin fibers that extend from the cell body in tree like branches, which receive information from other neurons. The cell body carries out the neuron's cellular functioning. Finally, the axon is a long, thin fiber that transmits information from the cell

body to other neurons, forming part of the nerve signaling communication system. Nerve signals often travel over long distances in the body to rapidly convey signals between cells, through a complex communication system in which dozens of neurons can be involved in a single circuit. There are two mechanism involved in transmission of the nerve signals: first, an electrical signal that conveys the information within cells; second, chemical signals for communication between cells, through conversion of electrical signals to chemical signals the action of small molecules called neurotransmitters.

1.3.2 Signal transmission within nerve cells

The mechanism of transmission within the neuron is based on differences in membrane potential that exist between inside and outside of the cell (Fig.5). It is created by uneven distribution of ions such as K^+ (higher inside the cell under resting conditions) vs. Cl^- and Na^+ (higher outside the cell). At rest, the cell membrane potential is always electrically negative on the inside relative to the outside. This is called the resting potential in non-excitable cells and is also the resting state in excitable cells like neurons when they are not exited. Movement of ions is mediated by trans membrane proteins that function as ion channels in the cell membrane, which open or close in response to changes in the trans membrane potential or via their response to activation by neurotransmitters. When the electrical potential across the cell membrane is depolarized during excitation, it triggers the opening of additional nearby ion channels, which propagates the depolarization down the axon, as an action potential. During this event, when the ion channels are open, sodium is flooding into the cells and potassium is moving out, down their chemical gradients.

The action-potential mechanism is based on movement of ions, changing in permeability of sodium and potassium across the cell membrane, due to the opening/closing of membrane ion channels (Fig. 5). When the membrane potential reaches threshold potential, sodium channel opens, allowing influx of sodium ions into the cells which this cause depolarization, resulting opening more sodium channels. Flow of sodium ions into cells across membrane creates high positive voltage around (+60mV), so to get equilibrium potential for sodium, potassium channels begin to open at the peak of depolarization to reach the equilibrium. When repolarization occurs, both potassium and sodium channels close but sodium channels return to their activated states.

1.3.3 Signal transmission between cells

Communication between neurons occurs at the synaptic cleft, which is a site where one neuron communicates with another, called the pre- and postsynaptic neurons, respectively. In this signal transmission event, the presynaptic cell's electrical signal is converted to a chemical signal via neurotransmitter release into the synaptic cleft, which occurs when an action potential reaches the axon tip at the synapse. Intercellular communication is thus chemical, not electrical, and is achieved through the binding of neurotransmitters to their receptors located on the postsynaptic neuronal cell. This may activate or inhibit the triggering of an action potential in the postsynaptic cell, depending on how the receptors are coupled to ion channels in that cell.

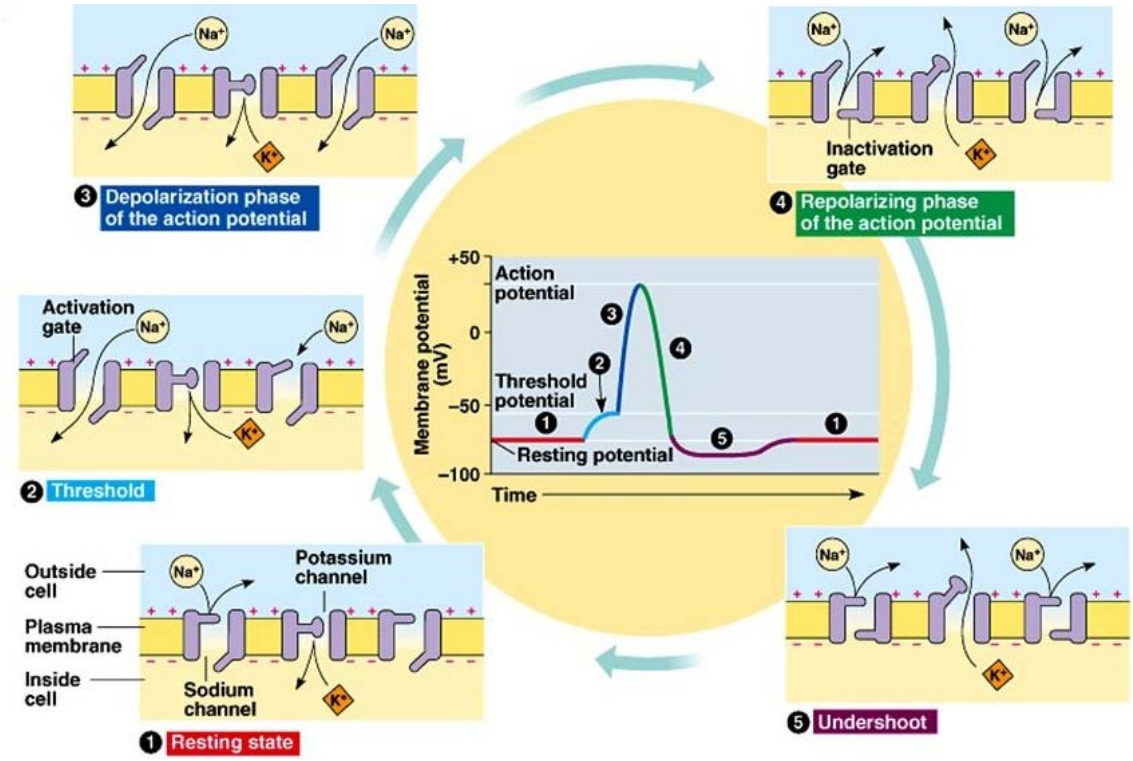


Figure 5. Changes in Movement of Sodium and Potassium Ions across the Membrane during Neuronal Action Potential. 1) both Na^+ and K^+ channels are closed, membrane`s resting potential. 2) Stimulations open some of Na^+ channels, Na^+ flow into the cell and membrane depolarized. The action-potential happens if the depolarization reaches the threshold. 3) Most of the Na^+ channels open (activation gates) due to depolarization, but K^+ channels remain closed. Inside membrane is more positive to the outside due to influx of Na^+ ions. 4) Most of Na^+ channels close (inactivation gates), Na^+ influx blocking, whereas K^+ channels open leading to efflux of K^+ ions outside of cells which makes the inside negative again. 5) Both gates are closed again and membrane returns to resting state. [https://dundeemedstudentnotes.wordpress.com]

1.3.4 Acetylcholine (ACh) as a neurotransmitter

ACh is an important neurotransmitter in both the mammalian brain and the peripheral nervous system. It is the terminal neurotransmitter in the parasympathetic division of the autonomic nervous system, which makes synaptic connections to multiple organs, glands and smooth muscles, e.g., acting to slow the heart in opposition to the action of adrenaline. It is also released at the neuromuscular junction, where efferent nerves release ACh to induce the contraction of skeletal muscle fibers. It also has multiple roles in the brain, and is implicated in various CNS pathologies such as Alzheimer's disease and Parkinson's.

The biosynthesis of ACh from choline and acetate involves a single enzyme, choline acetyl transferase, or ChAT. This makes ACh particularly appropriate for investigations pertaining to the regulation of neurotransmitter biosynthesis.

1.3.5 Neurotransmitter depletion and the regulation of neurotransmitter biosynthesis

Under normal conditions, neuronal firing leads to the release of ACh, enabling its action at ACH receptors on the postsynaptic neuron or muscle cell (Fig. 6). In order to distinguish this signal from the resting state, excess ACH in the synaptic cleft is removed by the enzyme acetyl cholinesterase, which splits ACH back into acetate and choline. Biochemical efficiency is achieved by recycling the choline by means of reuptake into the presynaptic neuron, which involves active transport via a transporter protein, also known as a “reuptake pump”. Because of limitations of biological efficiency, and the fact that energy is required, it remains necessary for the presynaptic neuron to synthesize

additional neurotransmitter molecules. The question of how this neurotransmitter biosynthesis is regulated is the focus of this research proposal. Particularly when neurons are stimulated at a very high firing rate - e.g., during extreme physical exertion such as running, one would expect that neurotransmitter depletion would require activation of ChAT gene expression in the nucleus of the affected neurons, in order to replenish the depleted ACh.

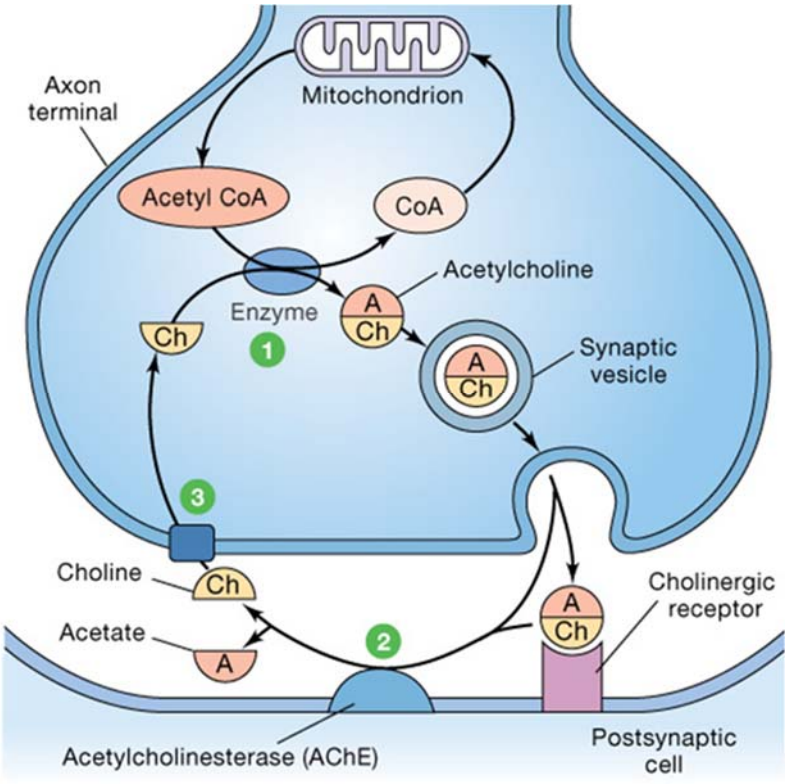


Figure 6. Synthesis and Recycling of Acetylcholine (ACh) at the Synapse. 1) Ach synthesis from choline and acetyl CoA. 2) Ach break down at the synaptic cleft by enzyme called acetyl cholinesterase. 3) To make more Ach, choline transport back to the axon terminal.

1.4 A Novel Hypothesis for the Regulation of ACh Synthesis by G-QPX

The basis of our hypothesis is the idea is that there is a genetic switch in neurons, so that when a nerve fires excessively, the sodium let into the cell during action potentials could accumulate in the nucleus, where it could act to destabilize G-QPX in gene promoter regions. Because the presence of such G-QPX would normally prohibit transcriptional activation (Fig. 7), their destabilization in the presence of excess sodium ions would be expected to activate transcription of the genes for the enzymes required to make more neurotransmitters. Thus, using acetylcholine synthesis as a paradigm, this

proposal will focus on the stabilization versus destabilization of G-quadruplexes in the promoter region of the ChAT gene as a potential regulatory mechanism for ChAT gene expression.

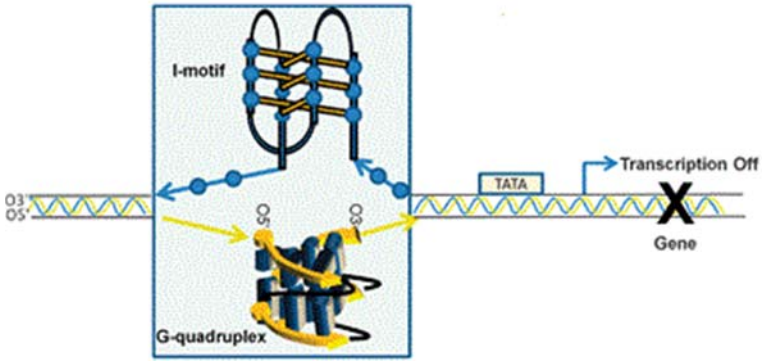


Figure 7. Schematic Representation of G-QPX Formation, upstream of Transcription Start Site and its Effects on Cellular Function. The duplex DNA unwinds upstream from the transcription start site, which makes the G-rich single strand available to fold into a G-QPX, which leads to down-regulation of transcription because the transcription factors (DNA-binding proteins) that normally activate gene expression are unable to bind [26].

CHAPTER II

BIOINFORMATICS STUDIES TO PREDICT G-QPX IN THE ACTIVE PROMOTER REGION OF THE ChAT GENE

2.1 Abstract

G-QPX nanostructures are found in critical regions of that human genome such as telomeres, promoter regions and 5'UTR of mRNA. In the promoter regions, they have functionality in up or down regulation of transcriptional activity. The mechanism of regulation could occur through disruption of dsDNA in which individual single strands contain G-runs, which when separated can form G-QPX structure, and thereby prevent binding of TBFS to their recognition sites for transcription initiation. Thus, they have become the focus of intensive interest due to their biological significant. Most G-QPX formation has been observed in the promoters of oncogenes. Different computational tools (various algorithms) have been developed to analyse DNA sequences and predict the potential to form stable G-QPX. Although the roles of G-QPX as regulatory elements have been studied in various human genes, their regulatory roles have not been investigated in the human nervous system to date. Thus, the focus of this work is identification of stable G-QPX structures in the promoter region of the choline acetyltransferase (ChAT) gene; it's protein product is involved in neurotransmitter

synthesis (the terminal enzyme for the synthesis of acetylcholine). A long fragment (97 nucleotides) with high CG content in the human ChAt gene promoter, 1000bp upstream of the TSS, has been chosen as an active promoter region. Bioinformatics studies using three different computational tools (QGRS, quadparser, and G4 calculator) found multiple intramolecular G-QPX structures in the active promoter region of the human ChAT gene, three consecutive G-runs in the negative strand (ChAT G17, ChAT G17-2, and ChAT G30) and one set of G-runs in the positive strand (ChAT G29). Here, for the first time we propose that G-QPX can form in approximately the same location in opposite strands. Also, we propose that intermolecular G-QPX could be formed between the two strands of duplex DNA (two DNA complementary strands) due to the short distance between G-runs in both DNA strands. Bioinformatic analysis to identify the TBFS on the active promoter region of the human ChAT gene showed that AP2-α and Sp1 have multiple recognition sites, in positions close to the identified G-QPXs. Transcription binding analysis further support the hypothesis of a significant function for G-QPX nanoswitches in cellular process.

2.2 Introduction

The common image of DNA is a B-form double helix, stabilized by a series of hydrogen bonds between two complementary strands of DNA as Watson and Crick proposed while definitive evidence has revealed that a unique non-Watson-Crick form of DNA base pairing interaction may form in certain sequences, giving a structure called a G-tetrad. Stacking of 2 or more G-tetrads gives rise to what is known as a G-quadruplex (G-QPX). These G-QPX structures are formed in guanine–rich regions through non-

covalent interactions and can be described as parallel or antiparallel [27]. If all of the strands contributing runs of G bases to the structure have the same 5`-3` orientation, the G-QPX is termed parallel, while if one or more strands contributing runs of guanine bases has a 5`-3` orientation opposite to the other runs of guanine bases, the structure is called antiparallel. Many investigators have reported that G-QPX can adopt various folding topologies; in some cases, even a single DNA sequence can form different alternative conformations of G-QPX [28].

In general, G-QPX consist of multiple G-tetrads (also called G quartets) stacked vertically. Each tetrad is formed by association of four guanine bases. The planar tetrad is a square shape and is stabilized by Hoogsteen base pairing, yielding hydrogen bonds between guanine residues (for a total of eight hydrogen bonds per tetrad). The unique tetrad structure creates a central cavity which can be occupied by monovalent cations such as (K^+, Na^+, Li^+) but only when two tetrads sandwiching the metal in between the tetrad layers . The presence of metal cations strongly influences the stability and folding topology of the quadruplexes [29]. In particularlys, K^+ is known to stabilize QPX, whereas at least in some cases, Na^+ plays a role in destabilizing the structure. Although it has been shown that stability of G-QPX is linked to the radii of the metal cations (increasing radii up to a point leads to increasing stability of G-QPX) other physiochemical conditions may also need to be favorable for the formation of G-QPX [26].

After sequencing the entire human genome, it has been found that G-QPX can form in critical regions such as telomeres, promoter regions, introns and untranslated

(UTR) regions of mRNAs [30]. Formation of G-QPX in these regions plays an important role in many cellular functions. G-QPX in telomeres inhibit the activity of telomerase which up–regulates the proliferation of most cancers cells [31]. In promoter regions of oncogenes, it has been observed that their formation regulates transcriptional activity (increasing or decreasing activity in various systems) [32, 33]. G-QPX formation in UTR regions has been found to control gene expression through decreasing translational activity [34, 35]. Therefore, G-QPXs have been the focus of intensive interest as an important class of targets for drug design and development, due to their biological significance and anticancer potential.

The first putative G-QPX sequence was discovered in the chicken beta - globin promoter [36], followed by the discovery of many sequences able to form intramolecular G-QPX, which were found in various genes including c-MYC [37, 38], VEGF [39, 40], HIF-1α [41], Ret [42], K& H RAS [43, 44], Bcl-2 [45, 46], c-Kit [47-50], PDGF-A [51] and c-Myb [52], hTERT [53, 54] and Rb [55]. Obviously, the diversity in the sequence of G-runs leads to diversity of folding G-QPX structures, where not only parallel and anti-parallel topology occurs, but also mixed parallel and anti-parallel folding is possible, as has been shown by NMR and CD spectroscopy. The structural diversity of G-QPX within promoter regions also is extended into the variation in G-run length, the length and sequence of the loops and finally the G-pairing arrangements. It has been proven that the type of monovalent cations and their concentration has a large effect on stabilization/destabilization of G-QPX. Circular Dichroism (CD) spectroscopy has been used to show that K^+ has a higher effect on stabilizing G-QPX compared to Na^+ [18, 56].

Moreover, several G-QPX stabilizing compounds have been identified to date. Examples include the porphyrin TMPyP4 [21], the perylenes PIPER [24], BRACO-19 [57], and several tetrasubstituted naphthalene diimide compounds [22].

In parallel with experimental studies, various computational tools have been developed to predict which sequences will form G-QPX in DNA [58-60]. There are a variety of algorithmic rules that can be used to predict the existence of G-QPX [61, 62]. Bagga and coworkers designed a site called QGRS mapper [63] which has adjustable parameters for looking at sequences for predicting the G-QPX. Quad finder [64], Quad Base [65] and other sites are also offered using essentially the same algorithm as QGRS. Lastly in this category, a new tool has been developed to identify G-QPX in a specific sequence (quadparser), which also offers a database of predicted G-QPX in the human and other genomes (Quad DB [66]). This tool has provided a QuadPredict server which can predict thermal stability of G-QPX in a sequence. Thermal stability of G-QPX can vary in the presence of monovalent cations. Finally, Maizel's lab [67, 68] has used a different approach to predict G-QPX in which a G4 calculator is built that looks at the propensity for G-QPX formation within a sequence region, rather than the prediction of individual G-QPX structure.

Although the role of G-QPX as regulatory elements has been studied in various human genes, their regulatory roles have not been investigated in human nervous system to date. We hypothesize that there is a genetic switch in neurons for the activation of genes involved in the synthesis of neurotransmitters, which could be depleted from the nerve endings when a nerve fires excessively. Under such conditions, the sodium let into

the cell during action potentials could accumulate in the nucleus, where it could act to destabilize G-QPX in gene promoter regions. Because the presence of such G-QPX would normally prohibit transcriptional activation, their destabilization in the presence of excess sodium ions would be expected to activate transcription of the genes for the enzymes required to make more neurotransmitters.

Thus, here we will focus on studying the role of G-QPX in the promoter region of a specific enzyme involved in neurotransmitter synthesis, choline acetyltransferase (ChAT), the terminal enzyme for the synthesis of acetylcholine. The aim is to assess the possibility of a regulatory mechanism for ChAT gene expression involvingthe stabilization versus destabilization of G-QPXs in the promoter region of the ChAT gene.

2.3 Materials and Methods

2.3.1 Comparative sequence analysis to select functional gene promoter region

Genetic information was explored by investigating well-cited literature with precise, clear and trustable results, available from the National Center for Biotechnology Information website, which allows the analysis of the promoter sequence to obtain a specific region of interest. Candidate gene has been used for initial sequencing analysis which only express in human neuronal cells with their enzymatic activity presented in Table1.

Table 1. Candidate Genes Involved in Neurotransmitter Synthesis with their Activities.

Gene terminology	**Enzymatic activity**
DOPA decarboxylase	Monoamine NT synthesis
Monoamine oxidase (both A & B isoforms)	NT degradation
Choline acetyltransferase (ChAT)	Monoamine NT synthesis
Serotonin Transporter (SERT)	NT reuptake (ie: recycling)

2.3.2 Bioinformatics analysis to predict G-QPX

After selecting a DNA region of interest (an active promoter region containing G-rich sequences), analysis is performed using different open access online tools to predict G-QPXs formation and their stability. QGRS mapper, Quadparser and G4 calculator are typical algorithms (three independent software tools) that have been used in this study, a summary of these methods are presented in Table 2.

Table 2. Methods used for Predicting Specific G-QPXs.

Terminology	G-runs length	Specificity	Year
Quadruplex G-rich sequences (QGRS)	2,3,or 4	Length restriction, scoring system	2004
Putative quadruplex sequences (PQS)	3+	Adjustable parameters in quadparser	2005
G4 potential(G4P)	2-5	Density of G- runs in a region	2006

2.3.2.1 QGRS Mapper (http://bioinformatics.ramapo.edu/QGRS/analyziz.php)

QGRS mapper is one of the few freely available web based programs which facilitate the study of G-QPX by prediction of their formation and stability in DNA and mRNA. It has been used to find the putative G-rich sequences in untranslated regions of the human genome as well as in translated regions [69]. The program is designed in such

a way that it identifies sets of potential candidates within the nucleotide sequences that have the capability of forming G-QPX. For instance, a motif composed of four guanine and three loops can be identified by searching $G_xN_{y1}G_xN_{y2}G_xN_{y3}G_x$ motif through the sequence. Once the candidates are identified (G-rich sequences), the algorithm used in the program calculates a factor called a G-score [60, 63]. Moreover, the algorithm behind QPXs recognition is efficient enough to not only identify the conserved G-QPXs across nucleotide sequences but also offer strategies to include overlapping G-QPXs [69]. Sequences with higher G-scores are predicted to be more stable and their formation is more probable compare to low G-score sequences. The two main parameters influencing the G-score predicted for a sequence are G-tetrads and loop characteristics. The loop size between G-tracts is variable, although it is allowed to be zero - base length [60]. A greater number of G-tetrads with shorter loop length will increase the G-score. Another factor which increases the G-score is evenly sized loops, due to improve stability of the GQPX.

Although introducing the concept of scoring for each sequence seems interesting to predict the most likely stable sequence, while there has not been any experimental data to verify the attributes based on the scoring system. Importantly, the scoring process is not clear to know how its produce, how the ranking list related to stability/ disability of predicted structure [61, 70]. Therefore, it is clear that the QGRS program should be used for GPXs prediction and the scoring values should be considered with caution.

2.3.2.2 Quadparser

The putative quadruplex sequence (PQS) approach offered by Huppert and Balasubramanian as a free online downloadable QPX-prediction tool, which was accessed at http://www.quadruplex.org . It has been provided a core database, adjustable quadparser setting and prediction stability server to identify the specific G-QPX motifs on either strand. The program designed to be used for different searching pattern like considering various loop and G-runs lengths to extract all possible QPXs which fitted desired context. Similar to QGRS, same algorithms has been applied in this approach to identify the G-QPX motifs without the scoring judgment.

2.3.2.3 G4 Calculator

Eddy and Maizels [68] they discovered a computational program called "G4 potential" (G4P) to predict which sequence has a potential to form G-QPXs based on the density of guanine runs in sequence. They took different approach and looking for four or more runs of three or more guanine within a fixed length sequence in a sliding window instead of looking for individual QPX formation [62, 71]. Therefore, the big flaw of the program is that it does not locate the precise position of G4 motifs. On the other hand, it get round of any problems arising from choosing between multiple potential QPXs structures within sequences.

2.3.3 Prediction of transcription factor binding motifs

Generally, QPXs upstream of promoter regions results in transcriptional changes due to the equilibrium effect between a folded and unfolded form of these sequences. One possible mechanism could be the QPX formation blocking the transcription machinery. This could involve preventing transcription factor binding sites (TFBS) binds to the regulatory elements and start transcription, OFF state transcription machinery (Fig.8). Proteins and drugs interacting – QPX might effects expression level. It is important to note that QPX sequences neighboring TFBS results in either decrease or increase in gene expression.

We used the TRANSFAC system, which is linked to the Transcription Element Search System (TESS) and PROMO database to scan our active promoter sequence for reported transcription factor binding sites (TFBS).

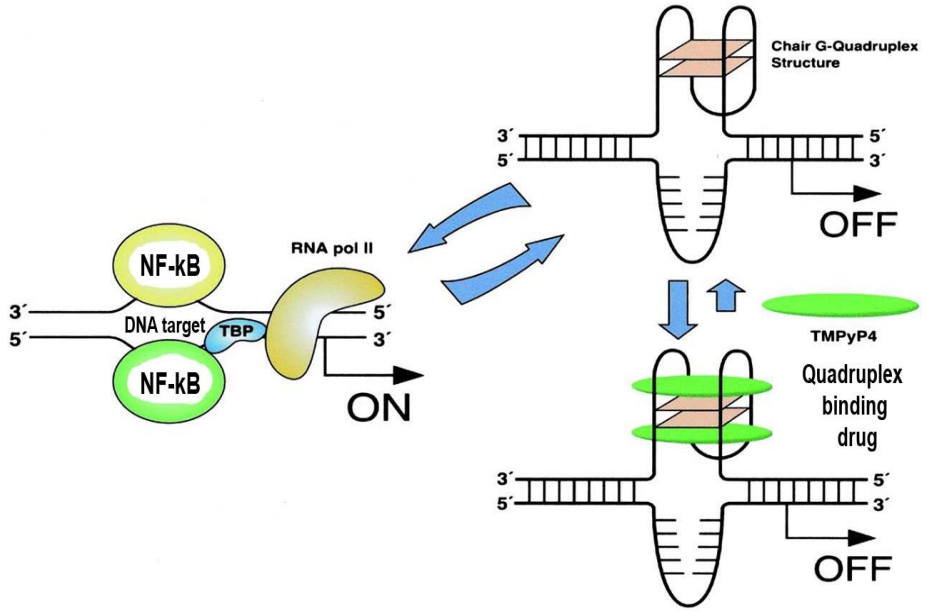


Figure 8. Effects of QPXs on Transcription Activity as Genetic Switch. There is a competition between a DNA double helix that can bind to TFBS (in this case NF-κB) OR formation of G-QPXs after separation of two complementary strands into DNA single strands. ON state meaning activation of gene expression and OFF state meaning inhabitation of gene expression which can be further "locked" off by QPX stabilizing drugs like TMPYP4.

2.4 Results

2.4.1 Identification of conserved sequences upstream the ChAT promoter region with high effect on gene expression from experimental literature

Choline acetyltransferase (ChAT) was chosen as a paradigm, because information about the functional portion of its promoter region has been well studied, and it is the single key enzyme required for Ach synthesis from precursors in the mammalian cell.

From previous publications, genomic information about the promoter region of ChAT was localized and discovered by Hersh, et.al. [72],was chosen as a candidate for further analysis (Fig. 9). The Hersh group observed a significant increase in basal

expression of the gene when nucleotides -925 to -1007 were cloned into the pGuP.PA8 vector. Increased expression was more than 30 fold and approximately five fold when this region cloned in the positive and reverse orientation, respectively, in the pGuP.PA8 vector. This observation indicates the presence of an important regulatory element in this region, which they emphasize, reflects the necessity of the enhancer elements for efficient gene transcription. Potentially active ChAT promoter match were identified via nucleotide BLAST searches in the NCBI website.

2.4.1.1 Reference sequences used in computational studies and oligonucleotide design

Genbank accession numbers (NC_000010.11; Homo sapiens chromosome 10, Homo sapiens choline O-acetyltransferase (ChAT)). The relevant sequence range is shown below.

Selected positive orientation of ChAT gene (plus/plus, 100% identity from NCBI).

```
  1 ggctgtcacccacggtcaccaaggagcaccatgctcccctcagcccaggatagaccctct
 61 tttccaggcctagcgcagagcccggggatgccgcccggggggagcctgaggacccgctcca
121 gctaggcacgccaggcccgccctttgaggacacgccccacaccagcctcagagctctga
181 ggtgcctgggctgagcttccttcagaccagaatcccgccccgttgaggctttgagaaag
241 gagtaggagc c
```

In the upper panel, blue and red highlights show the positions of -980 and -906, respectively, as presented in the Hersh paper. Yellow highlighting indicates the boundaries of a region spanning ~80 bases that showed a significant impact on promoter

activity (numbered between -925 to -1007). The CG nucleotide pairs, which may be involved in gene regulation based on Hersh group suggestion, are shown in pink italics.

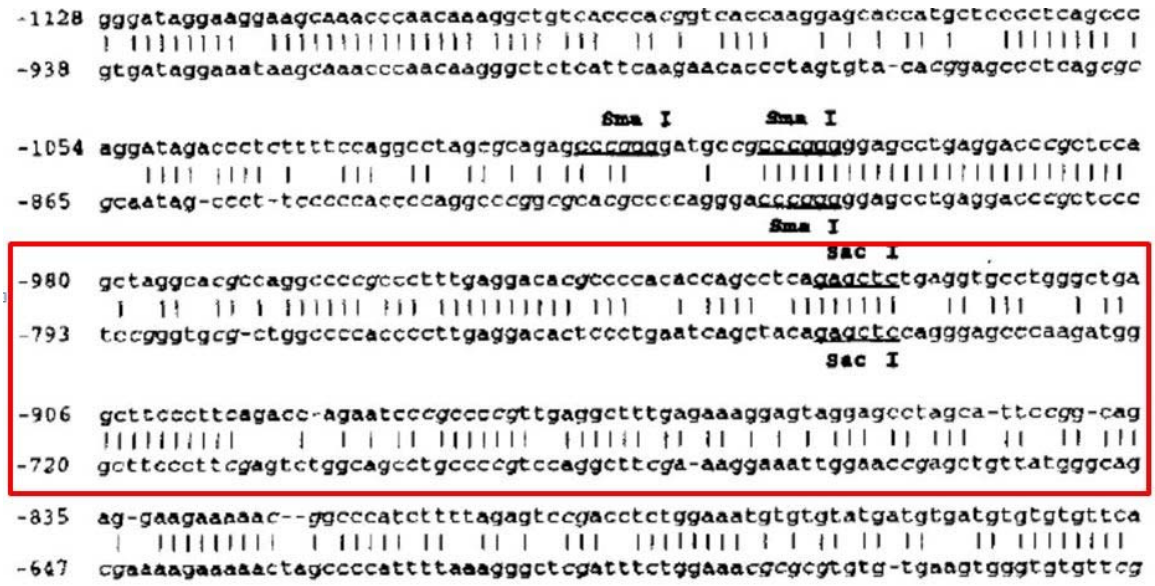


Figure 9. Promoter Region of ChAT Gene from Hersh, 1992 (Comparison of the Nucleotide Sequence of the 5' Flanking Sequences of the Human and Porcine ChAT Gene). Upper and lower sequences given in lowercase letters are presented as the human and porcine sequence, respectively. The red box is the region covered between positions -925 to -1007 shown by Hersh et.al as an active promoter of the human ChAt gene which increase transcriptional activity significantly [72].

2.4.2 Conserved G-rich region regulates ChAT promoter activity

From a bioinformatics study, the active promoter region of the ChAT gene shown above has strong potential to form multiple putative G-quadruplexes. These unique guanine rich sequences were identified by several independent computational methods as detailed below. Quadruplex prediction carried out on both single strands (positive and negative orientation) individually. Various different input parameters considered based

on each program but only results presented which met the criteria. Unmatched results (considering various input parameters) are not presented.

2.4.2.1 QGRS Mapper

The program found one and multiple non – overlapping conserved G-rich sequence on the positive (5`→3`) and negative (3`→5`) orientations of the active promoter region of the ChAT neurotransmitter gene, respectively (Table 3, 4). In the negative strand, six different non - overlapping QPXs with various topologies, length, position and G-score were predicted by QGRS mapper. Out of six different sequences, the sequence structure and position of three consecutive G-QPXs was most promising for further study, their positions are highlighted in color in Table 4.

Table 3. G-QPX Prediction Result for ChAT Gene on the Positive Strand from QGRS. One G-QPX with 25 nucleotides was identified in the promoter region. The number in the left column corresponds to the position of the first base of the potential G-QPX detected by this computational tool.

Position	Length	QGRS	G-Score
86	25	GGATGCCGCCCGGGGGAGCCTGAGG	13

Sequence of the positive orientation of ChAT promoter gene with labeled guanine bases of G-QPX predicted is shown below.

5′GGCTGTCACCCACGGTCACCAAGGAGCACCATGCTCCCCTCAGCCCAGGATAGACCC TCTTTTCCAGGCCTAGCGCAGAGCCCGGGGATGCCGCCCGGGGGAGCCTGAGGACCCGC TCCAGCTAGGCACGCCAGGCCCCGCCCTTTGAGGACACGCCCCACACCAGCCTCAGAGC TCTGAGGTGCCTGGGCTGAGCTTCCCTTCAGACCAGAATCCCGCCCCGTTGAGGCTTTG AGAAAGGAGTAGGAGCC 3′

Table 4. G-QPX Prediction Results for ChAT Gene on the Negative Strand from QGRS. Three different colored positions (84, 110, 133) are the selected G-QPX sequences. The numbers in the left column correspond to the position (in the promoter sequence) of the first base of the potential G-QPX detected by this computational tool.

Position	Length	QGRS	G-Score
30	23	GGGGCGGGATTCTGGTCTGAAGG	19
84	13	GGCTGGTGTGGGG	18
110	13	GGGCGGGGCCTGG	18
133	25	GGAGCGGGTCCTCAGGCTCCCCCGG	18
182	26	GGCCTGGAAAAGAGGGTCTATCCTGG	16
213	30	GGGGAGCATGGTGCTCCTTGGTGACCGTGG	20

Sequence of the negative orientation of ChAT promoter gene with three different sets of labeled guanine bases of G-QPX predicted is shown below.

3′GGCTCCTACTCCTTTCTCAAAGCCTCAACGGGGCGGGATTCTGGTCTGAAGGGAAGC TCAGCCCAGGCACCTCAGAGCTCTGAGGCTGGTGTGGGGCGTGTCCTCAAAGGGCGGGG CCTGGCGTGCCTAGCTGGAGCGGGTCCTCAGGCTCCCCCGGGCGGCATCCCCGGGCTCT GCGCTAGGCCTGGAAAAGAGGGTCTATCCTGGGCTGAGGGGAGCATGGTGCTCCTTGGT GACCGTGGGTGACAGCC 5′

2.4.2.2 Quadpaerser

Similar results were obtained using Quadparser for both strands, which validate the QGRS results. Since the program has various adjustable settings, varied inputs parameters can be used to extract all possible QPXs, some of which did not meet the criteria and importantly, were not verified by the other programs used. Unmatched results have not been shown. Adjustable input parameters are: type of base, number of bases in repeats, repeats in sequence, minimum gap size, and maximum gap size.

```
OUTPUT: ChAT promoter(positive strand)

0-26     4:1:1      GGGGATGCCGCCCGGGGGAGCCTGAGG

Found    1:1:1  overlapping:sequences:lines
0-26     4:1:1      GGGGATGCCGCCCGGGGGAGCCTGAGG


OUTPUT: ChAT promoter(negative strand)

7-19      3:2:1     GGCTGGTGTGGGG
32-44     3:2:1     GGGCGGGGCCTGG
56-84     5:4:2     GGAGCGGGTCCTCAGGCTCCCCCGGGCGG

Found     8:4:3 overlapping:sequences:lines
7-19      3:2:1     GGCTGGTGTGGGG
32-44     3:2:1     GGGCGGGGCCTGG
56-84     5:4:2     GGAGCGGGTCCTCAGGCTCCCCCGGGCGG
```

2.4.2.3 G4 Calculator

The program evaluates runs of guanines in a sliding window and calculates the percentage of windows searched that meet the desired criteria. For both strands, various criteria have been studied. Only on the negative strand the program found potential regions with sufficient density of guanine runs to form G-quadruplex structures (Fig.10). The calculator could not identify any region where G-QPX motifs may lie on the plus strand.

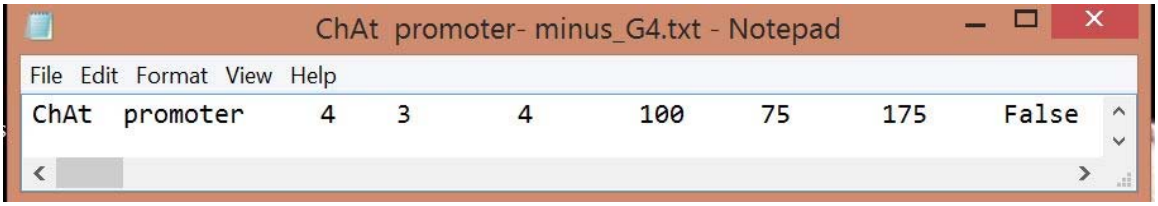


Figure 10. Result from G4 Calculator of the Active Promoter Region of the Negative Strand of the ChAT Gene. First column contains the file name (ChAT promoter). Following columns with this header information contains: the number of G-runs that met criteria (4), the number of C-runs that met criteria (3), the total number of windows searched (4), the percentage of windows containing G- runs that met criteria (100), the percentage of windows containing C- runs that met criteria (75), the sum of two percentages, both C and G runs (175), if the sequence contained a block of unknown sequence (Ns) (False means no significant runs of Ns).

2.5 Predicted TFBS on the Active Selected Promoter Region of ChAT Gene

From the open access TFBS prediction tools and database Sp1 and AP2-α has been chosen as enhancer elements. From PROMO database various TFBS were found in the selected active promoter region of ChAT gene. However, among them, Sp1 and AP2-α in agreement with the results of TRANSFAC analysis (Fig. 11 and 12) has been nominated for efficient transcription of the ChAT gene (Table. 5 and 6). Sp1 (zinc finger TFBS) and AP2-α bind to the consensus DNA-binding sequences which are 5'(G/T) GGGCGG (G/A) (G/A) (C/T)3', (GC box element) and 5'GCCNNNGGC3', respectively.

Table 5**.** Predicted TFBS in the Active Promoter Region of the ChAT Gene on the Positive Strand, Identified using the PROMO Database. The locations of Sp1 and AP2-α are presented in separate colors, pink and blue, respectively.

Sequence	10	20	30	40	50	60	70	80	90	100
TFBS	0, 1, 2, 3, 4, 18	8	1, 5, 6	5, 7, 8	2, 3, 5, 8, 9, 17	1, 2, 3, 4, 5, 11, 12, 14, 15, 16	2, 3, 5, 7 9, 10, 13		5,8	

Each number represent the beginning of a TFBS, 0: XBP-1, 1: E2F-1, 2:Pax-5, 3: p53, 4: Sp1, 5: GR-alpha, 6: RXR-alpha, 7:TFII-I, 8: AP2-α, 9: AhR/Arnt, 10: C/EBPbeta, 11: LEF-1, 12:TCF-4, 13: AR, 14: SRY, 15: TCF-4E, 16: GR, 17: ENKTF-1, 18: PEA3.

Table 6. Predicted TFBS in the Active Promoter Region of the ChAT Gene on the Negative Strand, Identified using the PROMO Database. Sp1 and AP2-α are presented in separate colors, pink and blue, respectively

Sequence	10	20	30	40	50	60	70	80	90	100
TFBS	0, 1	2, 11, 12, 15	0, 3, 4, 5, 6, 16, 17, 18	5,1, 10, 11, 12, 13	0, 1, 7, 15	3, 10	0, 1, 14	11, 12, 13	8, 9, 10	

Each number represent the beginning of TFBS, 0: GR-alpha, 1: AP2-α, 2: AR, 3: TFII-I, 4: SRY, 5: TCF-4E, 6: GR, 7: ENKTF-1, 8: PEA3, 9: XBP-1, 10: E2F-1, 11: Pax-5, 12: p53, 13: Sp1, 14: RXR-alpha, 15: AhR/Arnt, 16: C/EBPbeta, 17: LEF-1, 18: TCF-4.

matrix identifier	position (strand)	core match	matrix match	sequence (always the (+)-strand is shown)	factor name
V$AP2_Q6	9 (-)	0.992	0.982	cgcccGGGGGag	AP-2
V$GC_01	50 (-)	1.000	0.975	aggccCCGCCcttt	GC box
V$SP1_Q6	51 (-)	1.000	0.970	ggccCCGCCcttt	Sp1
V$SP1_01	52 (-)	1.000	0.954	gccCCGCCct	Sp1

Total sequences length=97
Total number of sites found=4
Frequency of sites per nucleotide=0.041237

Figure 11. TRANSFAC Analysis Result of Active Promoter Region of ChAT Gene on the Positive Strand.

matrix identifier	position (strand)	core match	matrix match	sequence (always the (+)-strand is shown)	factor name
V$SP1_Q6	30 (+)	1.000	0.970	aaagGGCGGggcc	Sp1
V$GC_01	30 (+)	1.000	0.975	aaagGGCGGggcct	GC box
V$SP1_01	32 (+)	1.000	0.954	agGGCGGggc	Sp1
V$AP2_Q6	73 (+)	0.992	0.982	ctCCCCCgggcg	AP-2

Total sequences length=92
Total number of sites found=4
Frequency of sites per nucleotide=0.043478

Figure 12. TRANSFAC Analysis Results for the Negative Strand of the Active Promoter Region of the ChAT Gene.

2.6 Conclusion and Discussion

Computational methods have been used extensively to better understand the role of G-QPXs in biological systems, e.g., their functions in telometric, UTR and promoter regions. In this study, we have evaluated computationally the possibility that G-QPX may be present in the active promoter region of the ChAT gene, where they could act as nano-switches to control the expression of the ChAT enzyme for neurotransmitter synthesis. The active promoter region of ChAT gene was discovered by Hersh, et.al [72]. They observed that there are a high number of CG dinucleotides in the promoter regions of the human ChAT gene which might play an important role in the regulation of ChAT gene expression via methylation. Indeed, Schlissel and Brown have noted that susceptibility of CG dinucleotide pairs to methylation has a crucial impact on transcriptional regulation [73].

To get a better sense of G-QPX significance and stability, quite a long fragment of promoter sequence has been chosen to strongly confirm the hypothesis. To date,

different numbers of algorithms have been used to identify the particular sequences with the capability to form G-QPX structures. Broadly they have been classified into specific and general algorithms [61]. Specific algorithms predict individual G-QPX structures while general algorithms identify regions which are likely to have high potential to form G-QPX structures. Although some of these motifs are accepted and widely used but there is no sufficient experimental evidence of their formation. It has been suggested that QPX motifs with two tetrads are expected to be more common than three, four or more tetrads.

Using several open access G-QPX prediction online tools, we performed a bioinformatics analysis of the active promoter region of ChAT gene. This analysis identified multiple G-QPXs in the close proximity to TFBS, where they could modulate gene expression via control of G-QPX stability (e.g., by flux of Na^+ vs. K^+ ions). The overall selected active promoter fragment of the human ChAT gene was 97 nucleotides which contain three consecutive G-runs in the negative strand and one set of G-runs in positive strand [Fig.13- bottom, Table.7]. A schematic for multiple G-QPX that can form upstream of the TSS of the human ChAT gene is presented in (Fig 13- bottom). To ensure that G-rich sequences have a potential to form QPX structures and their precise location in the selected active promoter fragment of the human ChAT gene, only non - overlapping predicted G-QPXs have been chosen. Potential overlapping G-QPXs were predicted in the both positive and negative orientations of the active promoter fragment of human ChAT gene (data not shown).

Moreover, from the literature, it has been described that the enhancer/ silencer regulatory elements are required for efficient transcription of the gene/ turning off the

gene expression in cells, respectively. Therefore, to further validate the discovery of G - QPXs in both strands, different bioinformatic studies performed to evaluate the TFBS positions on the active promoter region of the human ChAT gene. It was concluded that Sp1 and AP2-α transcription factors bind to their recognition sites which are close or same position of selected G-QPXs. It was demonstrated that the transcriptional activity of the rat ChAT gene increases in NGF-dependent (Nerve Growth Factor) fashion [74, 75]. In 1991, Ibanez and Persson suggested that enhancement of the rat ChAT gene expression by NGF occurs through AP-1 signaling [75].A similar outcomes was presented by Hersh, *et.al,* where a significant increase in transcriptional activity of the human ChAT gene in PC12 cells was seen in the presence of NGF [72]. CG-rich DNA consensus recognition sites for AP-2 showed notable homology to TGF-inducible (Transforming Growth Factor) transcription binding complex [76]. Additionally, it has been observed that Sp1- binding complex formation, function as crucial complex in accurate transcription initiation [77]. They have observed that 5' Sp1 consensus binding site containing CG-rich regions plays a critical role in maximal transcription.

The CG pairs in the active promoter region of ChAT gene, have been suggested involve in transcription regulation [72] which they are positioned close to the G-QPX predicted results (Fig.13- top, pink highlight) and further confirm important function of G-QPX nanostructures in cellular process.

In conclusion, in this study we have shown that the 97 nucleotide long active promoter region of ChAT unique within the human genome has a potential to fold into

multiple G-QPX structures featuring various loop lengths. Based on predicted G-QPX sequences, it can be suggested that ChAT G17 and ChAT G 17-2 should each have one zero loop (the end and middle loop, respectively). Therefore, assuming the idea of both structures adopting a novel zero loop conformation is valid, it can be concluded that they exhibit antiparallel or hybrid folding topology.

Here, for the first time, we present evidence for formation of consecutive G-QPXs upstream of the TSS in the promoter region of an enzyme involved in neurotransmitter synthesis. Also, we propose, for the first time, that G-QPXs might be formed in opposed strands of a natural DNA duplex at approximately the same location (Fig.12-bottom). Interestingly, visualizing G-runs in both DNA strands with a short distance brought up a new approach that G-QPX can form an intermolecular structure within two DNA complementary strands (Fig.14)

Although, the conclusion (predict and locate potential G-QPX motifs upstream of TSS) is supported by various bioinformatics analysis, experimental evidence is required to determine whether these unique structures fold into actual QPX structures, and how they may function in a biological system.

Table 7. Selected G-QPX Predicted Structures in the Active Promoter Region of the Human ChAT Gene.

Name	Strand	Sequence (5`→3`)
ChAT G17	Negative	GAGGCTGGTGTGGGGCG
ChAT G17-2	Negative	AAGGGCGGGGCCTGGCG
ChAT G30	Negative	CTGGAGCGGGTCCTCAGGCTCCCCCGGGCG
ChAT G29	Positive	GGGGATGCCGCCCGGGGGAGCCTGAGGAC

5′TCAGAGCTCTGAGGCTGGTGTGGGGCGTGTCCTCAAAGGGCGGGGCCTGGCGTGCCT
AGCTGGAGCGGGTCCTCAGGCTCCCCCGGGCGGCATCCCC3′

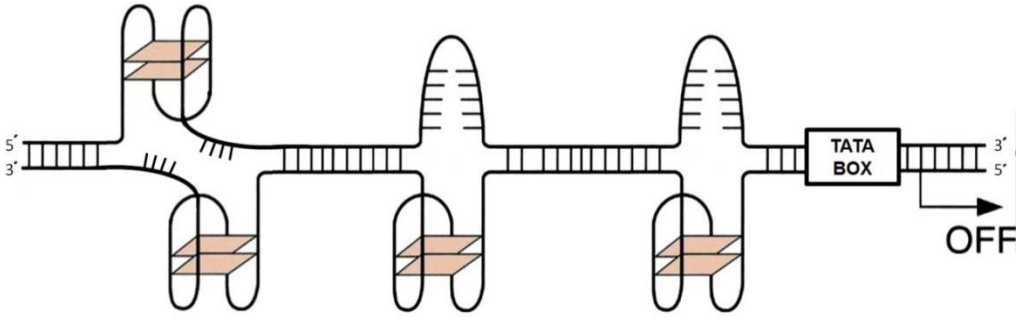


Figure 13. Schematic of Predicted Potential G-QPX in the Active Promoter Region of the Human ChAT Gene. Top) overall selected active promoter sequence of the human ChAT gene, shown as - strand 5'→ 3'(97 nucleotides), G-runs in colored fonts show the positions of potential G-QPX formed in both strands (green, pink, brown represent G-QPX structures in the negative strand, the blue C bases represent a potential G-QPX structure in the positive strand), locations which showed high promoter activity are highlighted in yellow (sequence match with location -925 to -1007, from Hersh, 1992), pink highlights are CG nucleotide pairs, which may be involved in gene regulation via methylation (Hersh suggestion, 1992). Bottom) proposed model of multiple G-QPX formation upstream of the TSS of the human ChAT gene.

5`GGGTCCTCAGGCTCCCCCGGG3`
3`CCCAGGAGTCCGAGGGGGCCC5`

Figure 14. G-QPX Sequences Forming Intermolecular Structures within DNA Duplex.

CHAPTER III

IN VITRO ASSESSMENT OF G-QPX MODULATING PROMOTER ACTIVITY

3.1 Abstract

The active promoter region of the human ChAT gene (97 cases long), is 1000 bp upstream of TSS, with a high CG base content with a high potential to form multiple G-QPX. Formation / deformation of G-QPX nanostructures play important roles in influencing transcriptional activity. G-QPX structure contain cavities between G-tetrads which can be occupied by alkali metal ions such as K^+ and Na^+. It has been observed that G-QPX structures showed the highest stability in K^+ solution. Also, small molecule ligands like TMPYP4 can stabilize G-QPX formation and further keep the gene in an off state, and inhibit gene expression due to formation of G-QPX structures in promoter regions. In silico analysis showed the first evidence of regulatory G-QPXs in the active promoter region of the human ChAT gene. Precisely, the active region encompasses AP2-α and Sp1 sites contain G-runs which involved in formation of multiple putative G-QPX structures. Wild type GFP reporter gene plasmid (AcGFP1-N3) was used to design and construct an artificial system (GFP reporter gene containing only active ChAT promoter region- AcGFP1-N3-modified). SP1 and Ap2-α are TFBS on the active promoter of the human ChAT gene which upregulate ChAT gene transcription

activity. Therefore, in general, formation of G-QPX in the same location or neighboring the TFBS recognition sites is expected to lead to downregulation of the transcription activity. NGFs (HNGF and TGF) and cytokines (TNF-α and IL-β) have been used as activators, facilitating signaling pathways terminating in the TFs AP2- α and Sp1. Aconitine as a Na^+ - dependent ion channel opening drug and TMPYP4 have been used as G-QPX destabilizing and stabilizing compounds, respectively. The results of modified reporter gene measurement showed highest GFP expression after 24 h treatment of transfected cells (neuroblastoma cell line, IMR-32) with HNGF in the presence of aconitine. In contrast, the GFP measurement using plate reader showed the lowest intensity following treatment with TMPYP4 (pre-DNA and post-transfection treatment). The obtained results were further supported using flow cytometry and EVOS fluorescence microscopic imaging. In vitro analysis using qRT-PCR, along with use of the same treatments to modulate ChAT mRNA expression, confirmed that HNGF and aconitine enhance the ChAT mRNA expression 2.4 fold and that TMPYP4 knocked down the ChAT mRNA expression by 87%. The results support the hypothesis that formation of G-QPX (with TMPYP4 treatment as G-QPX stabilizer) blocks binding of TFs (AP2-α and Sp1) to their binding sites, resulting in an inhibition of transcriptional activity of the ChAT gene, whereas an influx of Na^+ ions (as alkali metal ion G-QPX destabilizer) into the cells through aconitine treatment melts the G-QPX structures and increases ChAT gene expression. No cytotoxicity was observed after exposing the cells to various conditions used in the experiments.

3.2 Introduction

Human ChAT (choline acetyltransferase) gene located on chromosome 10, with NCBI reference sequence number of NC_000010.11, and is linked in the synthesis of neurotransmitter acetylcholine. The active promoter region of the human ChAT gene is located in the untranslated region (-1007 to -925), upstream of the TSS. It has been suggested that this region is a TATA-less, with highly CG - rich sequences which implying the importance of these sequences in transcription activity [72]. The DNA in such CG - rich regions capable of forming non-B form secondary structures called G-QPX. G-QPX forms from runs of guanine through Hoogsteen hydrogen bonds between four guanines in a planner fashion. The planner tetrad has a square shape structure with a central cavity which can be occupied by monovalent cations. Multiple planner G-tetrads are stacked vertically to form G-QPX structure. G-tetrads interaction could be occurs with one strand folding upon itself or involvement of multiple strands which these folding topologies are relevant to intramolecular or intermolecular isoform, respectively. G-runs are connecting through the loops which could be varied in length and sequence with either parallel, antiparallel or hybrid configuration. Typically, the loops length is between 1-9 base pairs but they could be long up to 26 base pairs [4]. Occupying the central cavity, between G-tetrad stacks, with metal cations such as (K^+, Na^+, Li^+) showed significant impact on stability of G-QPX structures. Particularly, G-QPX structures showed the highest stability in K^+ environment [49]. As mentioned before, CG - rich regions forming G-QPXs play an important role in human genome. These regions were clustered around 5`untranslated regions, telomerase and in promoter region upstream of

TSS. It has been shown that usually in promoter regions they formed in nuclease hypersensitivity elements [15, 78, 79]. Unwinding of these CG rich regions neighboring transcription binding factors (TFBS) can promote formation of G-QPX secondary structures and repress transcription activity. Onel et.al in their G-QPX studie`s, have reported that G-QPX structure in the P1 promoter of the human BCL2 gene were positioned in proximity to several TFBS and has been suggested that Sp1 and AP2- α binds to this region [80].

Study of promoter activity of the 5` flanking region of human ChAT gene in PC12 cells by hersh *et.al* showed that the transcription activity was enhanced in both strands orientation in presence of NGF [72]. In expression of ChAT gene in rat, it has been suggested by Ibanez and Persson, 1991 that NGF acted through an AP-1 TFBS [75]. Previous publication has been reported that Sp1 TFBS upstream of TSS acts as a determination of Transcription initiation site. Sp1 has been functioned as crucial protein in initiation of transcription form the TATA-less promoter sequences contain CG-rich regions. Moreover, TGF-α protein showed exclusive homology to CG-rich DNA recognition site of AP2 [76]. This means that the protein (TGF-α) remarkably recognized by AP2 (transcriptional regulator which controls gene expression). Figure 15 shows the signaling path of Sp1 and AP2-α synthesis.

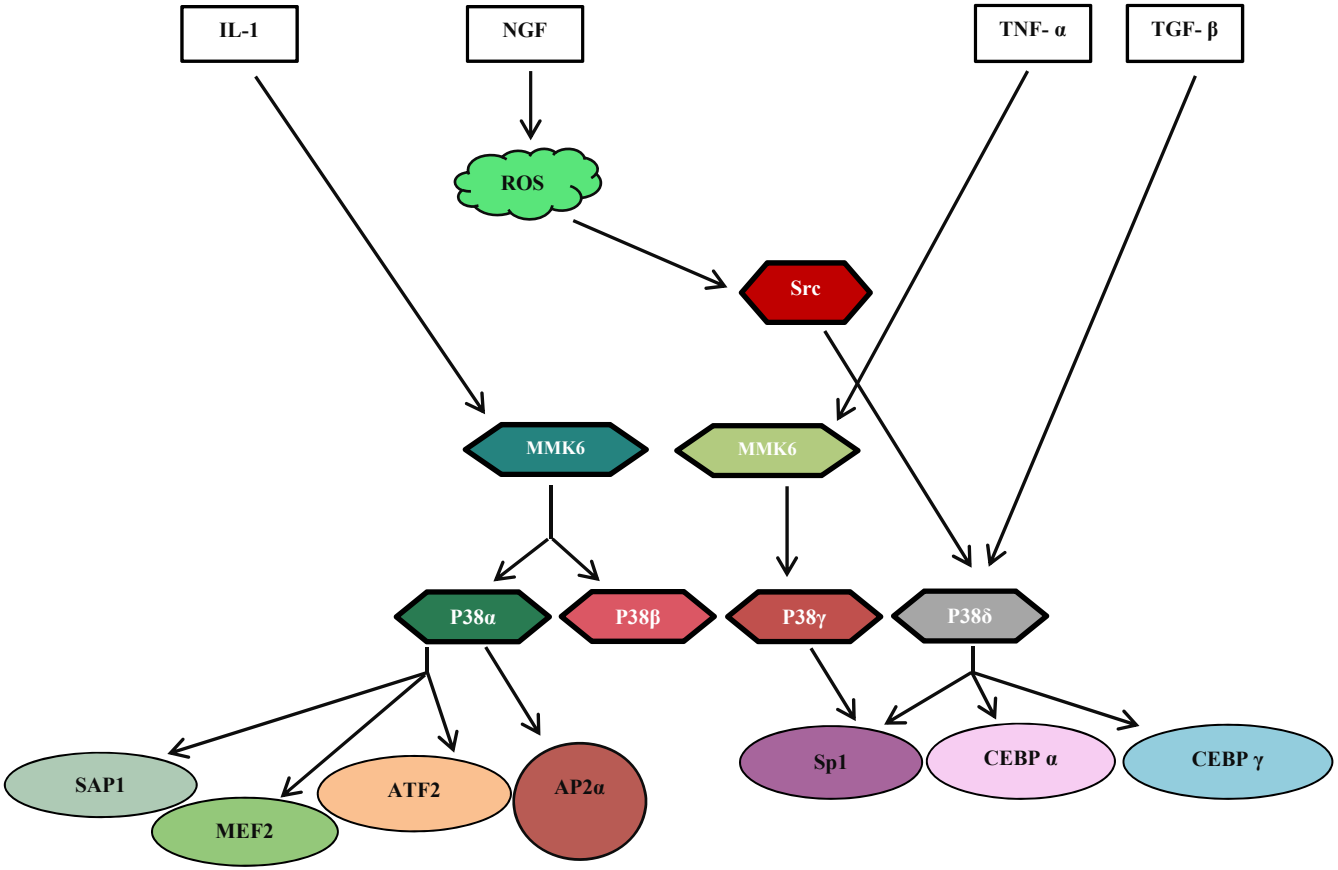


Figure 15. Signaling Path of Sp1 and AP2-α Synthesis. NGF, TGF, IL-beta and inflammatory cytokines like TNF-α are cell signaling proteins that facilitate synthesis of activator proteins (Sp1 and AP2-α) which are proteins that promote expression of the genes through regulating transcription machinery.

It is possible that G-QPX in the ChAT promoter involved in ChAT transcription activity just as the G-QPXs in BCL2, c-myc, c-kit, and k-RAS, and VEGF promoter of many other oncogenes [39, 80, 81]. Increase investigation of G-QPX sequences within a promoter region of oncogenes has led to offer the role of these sequences as regulatory regions and potent targets for anticancer drug discovery [3, 16]. There is no data available on the presence of G- QPXs in the promoter of neuron cells; in particular genes involve neuron signaling/ communications. Here we show the first evidence of regulatory G-QPXs in the active promoter region of the human ChAT gene. Precisely, the active region encompasses AP2-α and Sp1 sites contain G-runs which involved in formation of multiple putative G-QPX structures. These new G-QPX structures function as a transcription repressor. Ligands stabilizing the G-QPX conformation greatly fold and stabilize these high order structures (G-QPXs). In general, ligands deplete promoter activity which indicates a regulatory role ChAT G-QPXs in neurotransmitter synthesis. The development of specific G-QPX stabilizing compounds could have great potential as therapeutics targets.

3.3 Materials and Methods

3.3.1 DNA oligonucleotides

DNA oligonucleotides, HPLC purified, were purchased from IDT (Table1), which were stored at−20 °C. DNA concentration was measured by using nanodrop ND - 1000 ultramicro UV−Vis spectrophotometer (Thermo Fisher Scientific).

3.3.2 Cell culture

IMR-32 (Human neuroblastoma) cells (ATCC, cat#: CCL-127) were grown at 37°C, under 5% CO_2 in Eagle's Minimum Essential Medium, EMEM, 1X (low glucose and 2.2 g/L sodium bicarbonate with L-glutamine and phenol red) (Thermo Fisher Scientific). This media was supplemented with 10% Fetal Bovine Serum. Cells used in the assays were at a passage number less than 10 (Appendix A).

3.3.3 Plasmid construction

To construct the plasmid, the DNA sequence of active promoter region of ChAT gene, ds ChAT 137 sequence (from -960 to -980), containing consecutive runs of guanine, TATA box and flanked with restriction sites for enzymes PciI and Hind III (NEB) were synthesis from IDT, and amplified by using PCR (Appendix B). The total volume of PCR reaction mixture contained 2 μL PCR buffer (10X), 500 μmol/L dNTP mix, 0.15 μM each primer (shown in Table 1(p-ChAT S & p-ChAT A), 1 μL DNA template, 1 U Taq DNA polymerase, and nuclease-free water. The protocol used for PCR includes a denaturing cycle of 2 min at 95 °C, 35 cycles of PCR (95 °C for 30 s, 54 °C for 30 s, 72 °C for 1 min), and then 72°C for 10 min, followed by holding at 4 °C. The PCR then loaded in 1% agarose gel. After electrophoresis, the gel was imaged by using UV-shadowing.

In order to insert the DNA sequence (active promoter region) into the vector, plasmid (pAcGFP1-N3) was incubated in 37 °C water bath for 2 h to ensure that all of the vectors were cut by the enzyme PciI and HindIII. Then the native cut plasmids and DNA

sequences were ligated T4 DNA ligase (NEB) at 16 °C for 1 h. pAcGFP1-N3- modified plasmids were obtained by transformed plasmids into MJ109 cells. The plasmid has been further validated by DNA sequencing.

3.3.4 Transfection and reporter GFP-gene expression

The potential effects of G-QPX formation up stream of TSS, against the ChAT promoter activity was studied by transfections of IMR-32 (human neuroblastoma) cell line. IMR-32 cells were maintained in EMEM medium with 10% FBS (Thermo Fisher Scientific) at 37 °C, under 5% CO2. Transfection were carried out on 96-well tissue culture plates when cells were 70-80% confluent. Then IMR-32 cells were tryptinized at 37°C, under 5% CO_2 for 5-7 minutes, followed by seeding onto 96 well plates at a density of 3 x 10^4 cells /well. After 24 h incubation at 37°C, under 5% CO_2 (See appendix B), the transfections were operated according to the manufacturer's protocol, both Transfast transfection reagent (Promega) & Lipofectamine 3000 reagent (Invitrogen). For lipofectamine 3000, cells were transfected using 100 ng of DNA (AcGFP1-N3-modified reporter plasmid) and 0.3 µl lipofactamine 3000 for each well, at a ratio of 1:1 plasmid DNA and lipofectamine3000, in OptiMEM medium. Before adding the mixture to the micro-plates, the DNA mixed reagents were incubated at room temperature for 5 minutes. For Transfast transfection reagent, cells were transfected using 100 & 200ng of DNA (pAcGFP1-N3-modified reporter plasmid) and 0.3 & 0.6µl transfact reagent for each well respectively, at a ratio of 1:1 plasmid DNA and transfast reagent, in complete medium. After 1h of transfection, different conditions were added to the cells, varied transcription activators (TNF-α: 10 ng/ml , HNGF: 10 ng/ml , IL-β: 1nM , TGF: 10ng/ml), drug

dependent – Na^+ channels (Aconotine: 0.15 μM), incubated at 37 °C, under 5% CO_2. To study the effect of G-QPX stabilizing drug (TMPYP4: 10 μM), DND was incubated with drug at room temperature for 2h before addition of transfection reagents. The GFP expression of pAcGFP1-N3-modified plasmid in IMR- 32 cells under varied conditions in compare to wild type pAcGFP1-N3 were assayed by using plate reader and fluorescence imaging. Experiments were repeated at least three times in six replicates.

Table 8. Oligonucleotide Sequences have been used for Amplifying Inserted Fragment to Construction pAcGFP1-N3- Modified Plasmid and In vitro qPCR Studies.

Name	Sequence (5` →3`)
p-ChAT S	AAAAACATGTGCTTGTCGACTCAGAGCTCTGAG
p-ChAT A	GGTCAAGCTTAGCATTGCTTATATAGACCTGGG
qChAT S	TGTGGTCCCAAATGGGTATG
qChAT A	TGCAGCTGTGAAAGCTAGAG
USB1(h.g)	Ref seq: NM-000181(IDT)
ds ChAT 137	GCTTGTCGACTCAGAGCTCTGAGGCTGGTGTGGGGCGTGTCCT CAAAGGGCGGGGCCTGGCGTGCCTAGCTGGAGCGGGTCCTCA GGCTCCCCCGGGCGGCATCCCCAGGTCTATATAAGCAATGCT AAGCTTGACC

3.3.5 Cell cytotoxicity

Cytotoxicity of IMR- 32 cells transfection followed by various conditions (individual or mixed) as mentioned before was tested using MTT and MTS proliferation assays (Appendix A). MTT cell proliferation assay kit (Invitrogen) involves the conversion of yellow soluble tetrazole (3-(4,5-dimethylthiazol-2-yl)-2,5-ditetrazolium bromide) to insoluble purple product in living cells. Adding the second solution DMSO dissolves the insoluble purple product into a colored solution. The colored solution can be detected using spectrometer at an absorbance of 570 nm. MTS cell proliferation assay,

CellTiter 96® AQ_{ueous} one solution kit (Promega) contains a solutions of a novel tetrazolium compound [3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium] and the electron coupling reagent, phenazine ethosulfate; (PES). Similar to MTT assay, MTS produce a soluble product by only sue of one solution and eliminating the DMSO solubilizing step. It a one-step solution assay and the colored solution can be detected using a spectrometer at an absorbance 490nm.

Cells were seeded at a density of 3.5 x 10^4 cells/ well in a 96 well plate and incubated for 24 hours at 37°C, under 5% CO_2 . On day 2, Cells were transfected to pAcGFP1-N3-modified plasmid and treated with various concentrations (TNF-α: 10 ng/ml, TGF: 10 ng/ml, HNGF: 10 ng/ml, IL-β: 1 nM, AC: 0.15 μM, (HNGF: 10 ng/ml + AC: 0.15 μM)) for both the MTT and MTS assays. Untransfected cells and same volume of DMSO were used as a control.

3.3.6 Quantitative real-time RT-PCR

The IMR-32 cells were seeded onto 6 well plates at a density of 18x 10^4 cells /well on day one. After 24 h incubation at 37°C, under 5% CO_2, the cells were treated (on day two) with following conditions: Aconitine (0.15 μM) in the presence of HNGF(10 ng/ml), TGF(10 ng/ml), TMPYP4 (10 μM). Control sample include IMR-32 cells, which did not give any treatment, only expose to the normal complete media (90% EMEM, 10% FBS). After 4 h of treatment, cells were lysed for RNA isolation. For each experiment the total RNA was isolated using promega total isolation kit (Promega). RNA concentration was measured by using nanodrop ND -1000 ultramicro UV–Vis

spectrophotometer (Thermo Fisher Scientific). Power SYBR Green RNA-to-C_T 1-Step Kit (Applied bioscience) was used to perform one- step RT-PCR with SYBR Green reagent for quantification analysis on a real-time PCR (Appendix B). qRT-PCR was run on ABI 7500 fast equipment (software version 2.3) with conditions as manufacturer`s protocol, 48^0 C for 30 min (reverse transcription step), 95^0 C for 10 min (Activation of AmpliTaq polymerase), 95^0 C for 15 seconds (Denature) and 60^0C for 60 seconds (Anneal/ extension).Primers used for real time RT-PCR were: ChAT primers (qChAT s & qChAT A), USB1 primers (NM-000181). The mRNA expression of ChAT gene in each sample was normalized relative to expression of internal housekeeping gene (USB1) and ChAT gene in untreated cells to obtain the fold change. Data analysis for calculating the relative expression was done by using Paffl's method [82]. Experiments were run in triplicate with internal technical and one-way ANOVA analysis was used to determine statistical significance.

3.3.7 Flow cytometer analysis

IMR-32 Cells were seeded in 6-well plates (0.4×10^6 cells/well) the day before transfection (Day1). Next day (day2), the cells were transfected with the pAcGFP1-N3-modified construct and AcGFP1-N3 parent plasmid using Transfast transfection reagent (Promega). One hour after transfection, wells were replaced with pre-heated fresh complete media with various additives including TNF-α: 10 ng/ml, HNGF: 10 ng/ml, IL-β: 1 nM, TGF: 10 ng/ml, Aconotine: 0.15 μM, and Kcl: 10nM. TMPYP4 (G-QPX stabilizing drug): 10 μM was pre-treated with DNA at room temperature, 2hr before adding transfection reagent. For analysis of transfected cells, after 24hr, cells were

trypsinized, washed with 1% PBS, followed by re - suspension in 400 µl of 0.1% BSA in PBS. Total of 10,000 gated events were collected for each sample. Data were analyzed using FACS DIVA version 6.1.3 software.

3.4 Results

3.4.1 Synthetic substitute active ChAT promoter region to GFP- reporter gene

The close neighboring of identified TFBS and predicted G-QPXs forming position compelled us to explore the relationship between selected TFBS (Sp1 and AP2-α) and multiple G-QPX forming sequences in the active promoter region of the human ChAT gene. As shown in Hersh study, the CG- rich regions should be the key DNA recognition element of TFBS that bind to the ChAT active promoter sequence and effectively regulate transcription activity. These regions are located in the same position or near of selected TFBS, in particular Sp1 and AP2-α (Fig.13-Top, pink highlights). Thus, we cloned the engineered ChAT promoter in to pAcGFP1-N3 and create a novel reporter vector, pAcGFP1-N3 modified (Fig. 16 -B & C). To generate the practical ChAT promoter composition, specific sequences such as TATA box and restriction enzymes recognition sites have attached to the active ChAT promoter sequence which are required for beginning of the transcriptional machinery (Fig. 16-A).

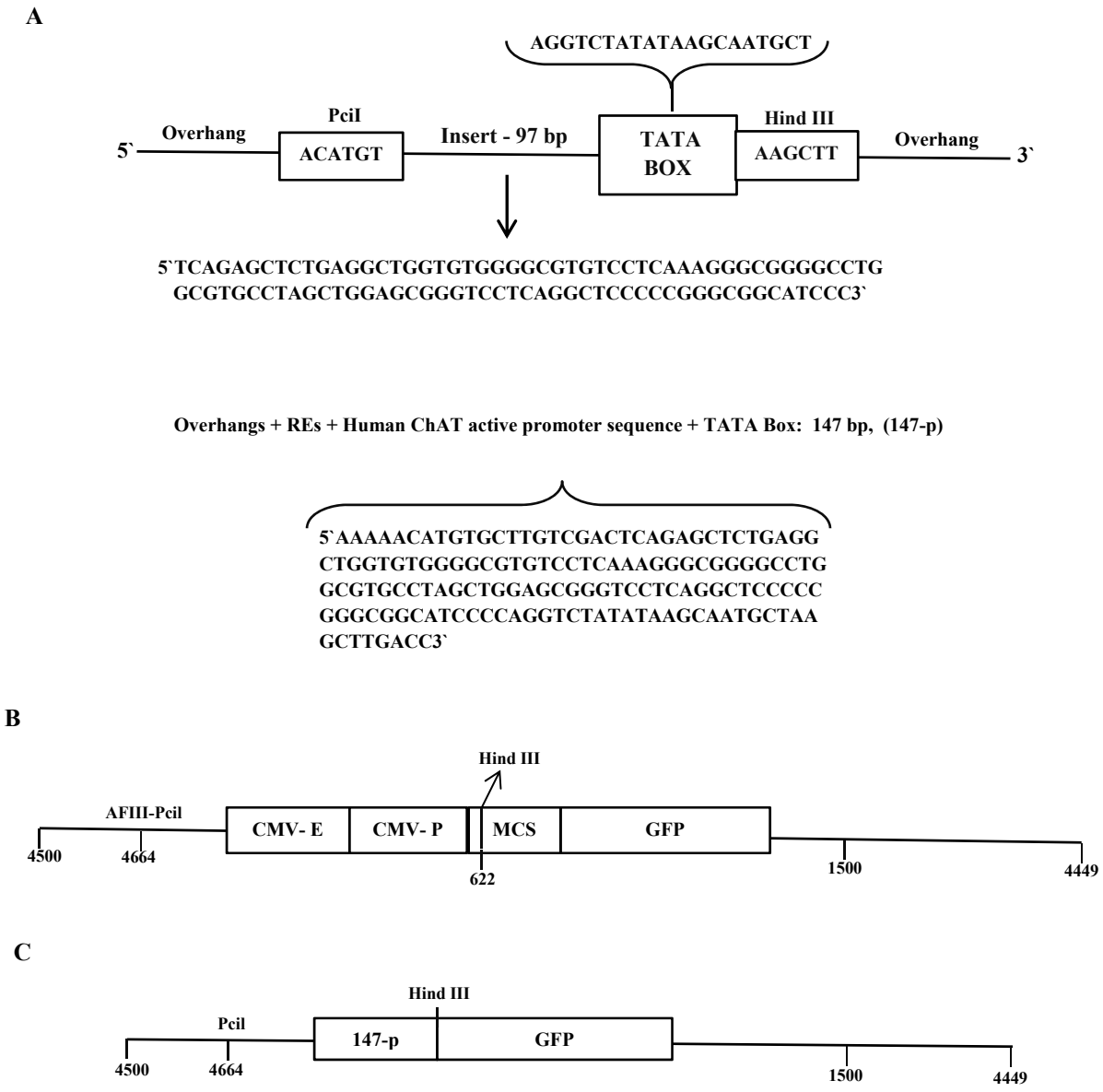


Figure 16. Schematic Representation of Engineered ChAT Promoter and pAcGFP1-N3-Modified Plasmid. A) Engineered ChAT promoter (147 bp) containing overhangs, Res, active promoter sequence of ChAT gene (97 bp), and designed TATA box (20 bp). B) Schematic of wild type pAcGFP1-N3 (4722 bp) containing enhancer elements (CMV-E) an promoter sequence (CMV-P) up stream of Multi cloning sites (MCS). C) pAcGFP1-N3-modified (4189 bp) containing engineered ChAT promoter inserted between Hind III and PciI.

3.4.2 Measuring reporter expression using activators and aconitine

Our reporter system is a modified GFP reporter that is used to examine the regulatory role of the active promoter of ChAT gene with a high potential of forming G-QPXs which needs to be analyzed. To determine the reporter gene transcription efficiency two different transfection reagents has been used in presence of different additives. Thus, for both reagents, maximum of expression level observed after 24 h treatment but the Transfast transfection showed more dominant promoter activity through GFP expression (Fig.17). There was a recognizable difference in transfection efficiency, using transfast reagent.

To study the effect of TFBS, we directly compare the GFP reporter system containing the same backbone but in the presence of varied transcription activators, were transfected into the neuroblastoma cell line (IMR-32). Expression was measured in cells after several exposure times to get the highest level. Background activity, was measured and subtracted from GFP expression level of each transfected sample. Untransfacted cells and transfected cells with wild type GFP vector (pAcGFP1-N3) used as negative and positive controls, respectively. As shown in Figure. 18, all samples have highest expression level after 24h treatment. Cells showed the highest GFP expression level in presence of aconitine with HNGF compares to base GFP expression level, M condition (exposure to normal complete medium without additives). The GFP expression level followed by transfected cells exposed to medium containing transcription activators (HNGF, TGF, TNF-α, IL-β), and dug – dependent Na^+ channel (aconitine: 0.015 μM).

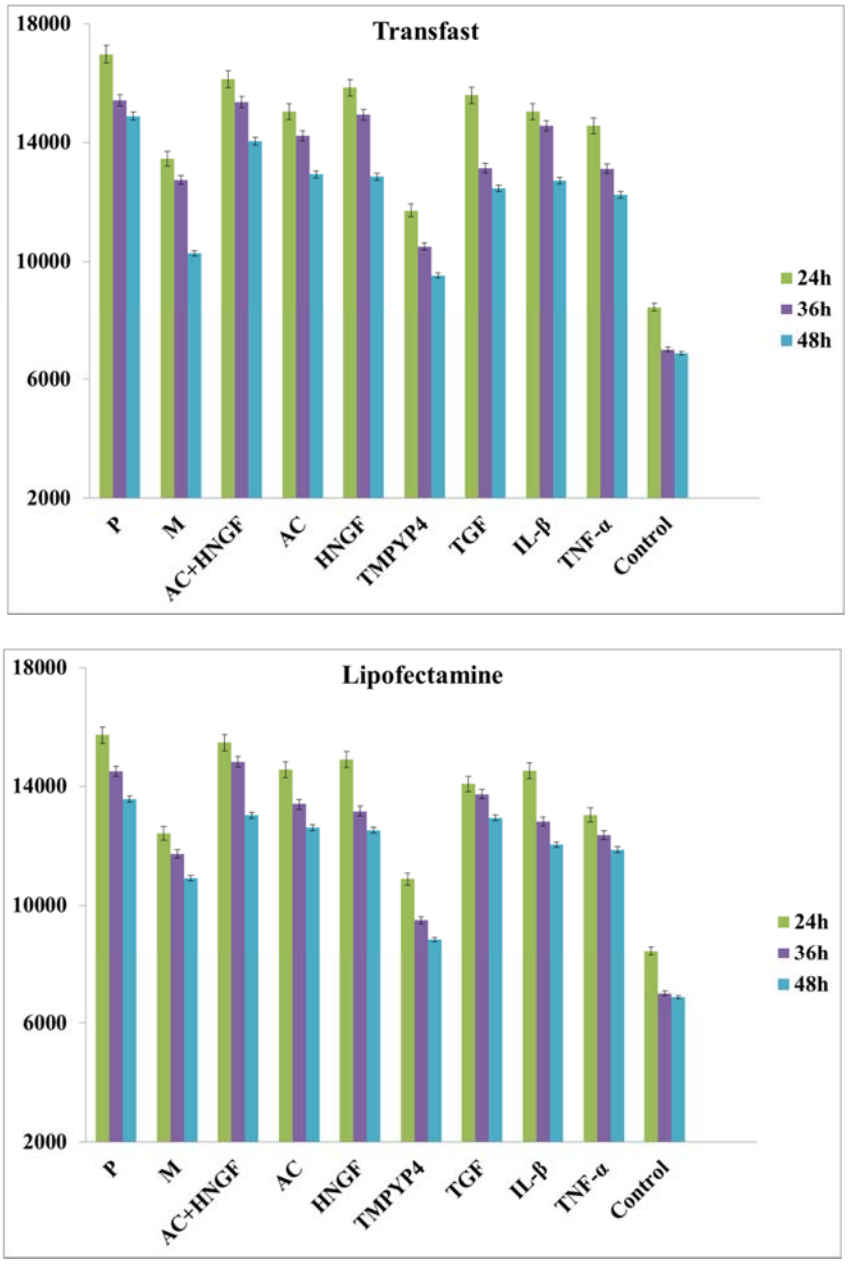


Figure 17. Selection of Transfection Reagents and Optimization of the Transfection Efficiency. IMR-32 cells were transfected with pAcGFP1-N3-modified construct using two different transfection reagent kits: Top) Transfast, Bottom) Lipofectamine. Un-transfected and transfected cells with parent plasmid used as negative and positive (P) controls, respectively. . M condition is base expression level which cells expose to normal complete medium without any additives. The GFP protein was detected at emission (474nm) and excitation (515nm) using a spectrometer. The fluorescent values represent the mean of two separate experiments of four readings and the corresponding standard errors of the means (s.e.m.).

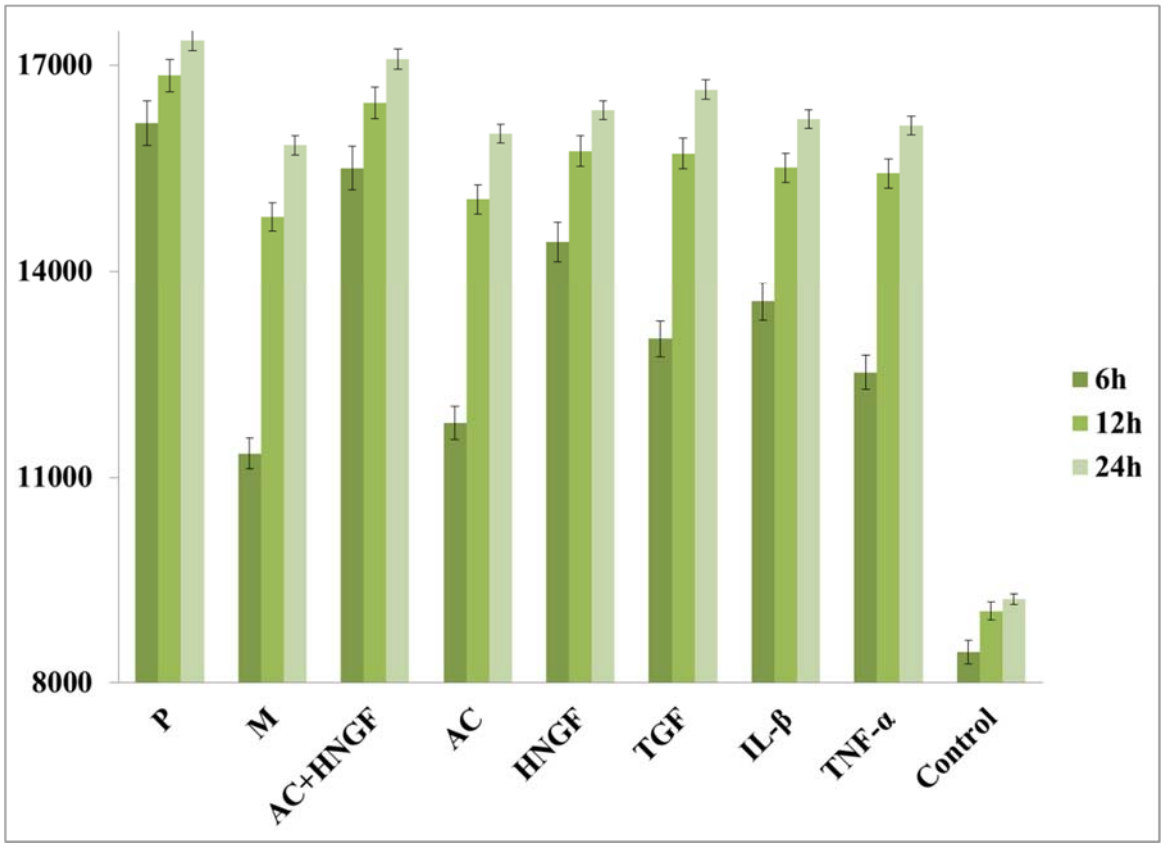


Figure 18. The GFP Expression Level of the pAcGFP1-N3-Modified Construct in the Presence of Varied Environmental Conditions. IMR-32 cells were transfected with pAcGFP1-N3-modified construct using following conditions as an expression activator and repressor: AC, (AC+HNGF), HNGF, TGF, TNF-α, IL-β. M condition is base expression level which cells expose to normal complete medium without any additives. Un-transfected and transfected cells with parent plasmid (pAcGFP1-N3) used as negative and positive (P) controls, respectively. The GFP protein was detected at emission (474nm) and excitation (515nm) using a spectrometer. The fluorescent values represent the mean of three separate experiments of four readings and the corresponding standard errors of the means (s.e.m.).

3.4.3 Block the TFBS binding and stabilizing G-QPX structures with small stabilizer molecule, TMPYP4

To verify the relationship between the small ligand binding molecules and G-QPX structures, TMPYP4 has been used, which was identified as G-QPX stabilizer. The interaction between G-tetrad planner and TMPYp4 could block transcription machinery through stable QPX structure formation or blocking the binding position of TFBS. Based on the results, transfection was conducted with modified vector pretreated with TMPYP4 to identify the binding ability of small ligand (TMPYP4) with G-QPX structures up stream of TSS. As shown in figure.19, THE GFP expression level decreased significantly not only compares high expression level in presence of aconitine with HNGF but also in compression with the base level of GFP expression.

To further show that first, G-QPX structures are strongly stabilize with TMPYP4, seconds, aconitine with HNGF significantly increase the expression level due to influx of ions and their effect on G-QPX formation, and third, the effect of transcription activator on the level of expression, the fluorescence GFP expression was directly visualized using the EVOS. As is shown in figure 20, similar level of expression was observed in IMR-32 cells transfect with pAcGFP1-N3-modified construct compare to wild type. While, significant decrease in the GFP expression was detected in cells, when DNA was pre-incubated with TMPYP4. TGF exposure treatment after transfection leads to increase the GFP expression level more than base level (only complete medium without any additive).

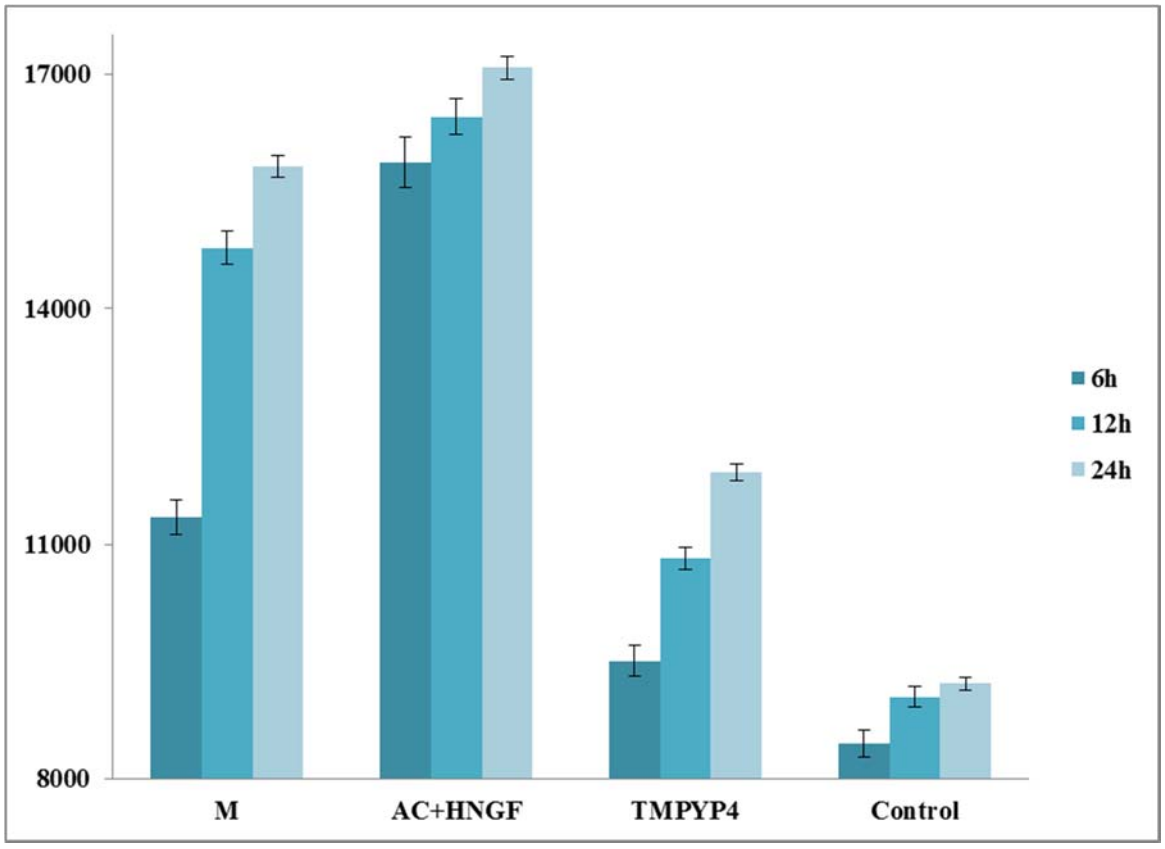


Figure 19. The Effect of Ligand Stabilizing G-QPX (TMPYP4) on the GFP Expression Level of the pAcGFP1-N3-Modified Construct. IMR-32 cells were transfected with 2h pre-treated pAcGFP1-N3-modified construct with TMPYP4.Un-transfected cells as negative control and transfected cells with parent plasmid (pAcGFP1-N3) and pAcGFP1-N3-modified construct (exposure to HNGF+Aconitine), both used as positive controls. M condition is base expression level which cells expose to normal complete medium without any additives. The GFP protein was detected at emission (474nm) and excitation (515nm) using a spectrometer. The fluorescent values represent the mean of three separate experiments of four readings and the corresponding standard errors of the means (s.e.m.). There was a significant difference between the means of TMPYP4 treatment and positive controls.

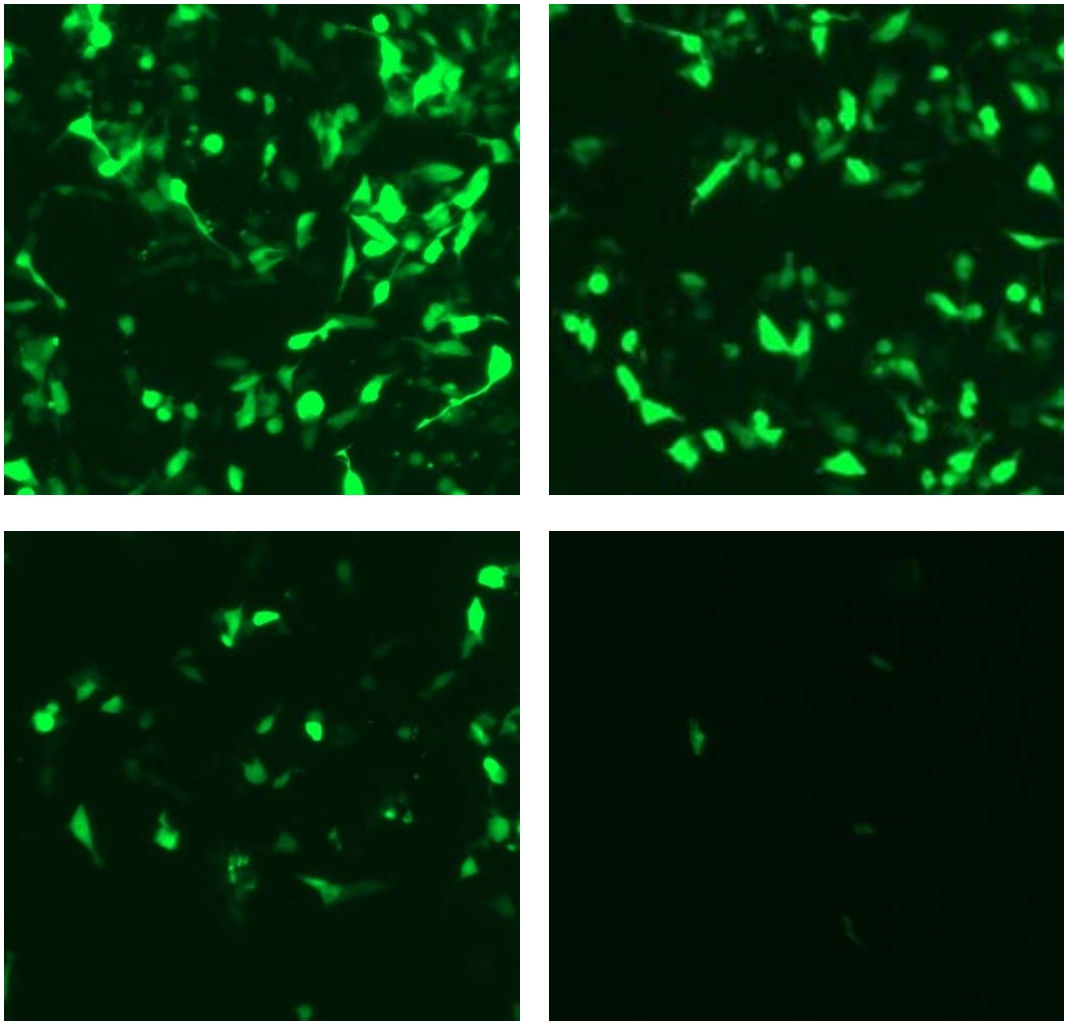


Figure 20. Visualization of the Fluorescence GFP Expression in Transfected Cells using the EVOS FL Microscope. Cells were transfected and visualized after 24h incubations. Top-left) cells transfected with wild type GFP vector, Top-right) cells transfected with modified GFP vector and exposed to aconitine with HNGF, Bottom-left) cells transfected with modified GFP vector and exposed to TGF, Bottom-right) cells transfected with modified GFP vector and pre- incubation with TMPYP4 G-QPX stabilizer. All the images were taken at 20× magnification.

3.4.4 Cytotoxicity analysis

Our cytotoxicity analysis showed there was no cytotoxic effect using either the MTT or MTS cell proliferation assays using Transfast reagent (Fig.21). There was no inhibition of cell growth compared to the untreated cells (Data not shown).

3.4.5 Flow cytometry analysis

Reporter expression was further studied using flow cytometry considering all previous conditions. The results were the same as expression level using plate reader (Fig. 22 and 23). Therefore, it has been observed that TMPYP4 as G-QPX stabilizing compound and aconitine as Na^+ - dependent channel have high influence on stabilizing and melting G-QPX structures, respectively, also, it could be suggested that they lead to a significant reduction and enhancement in transcription activity. In each treatment, the P3 population shows the maximum GFP expression of modified plasmid relative to wild type GFP vector under specific condition, presented in figures 24, 25.

Figure.26 shows the percentage of mean GFP fluorescence intensity using flow cytometerty which the promoter activity was affected by presence or absence of G-QPX stabilizing/ destabilizing molecules or metal cations. Thus, there is a clear evidence for competition of TFBS and G-QPX formations upstream of TSS which directly related to distribution of the GFP fluorescence intensity in the transfected population.

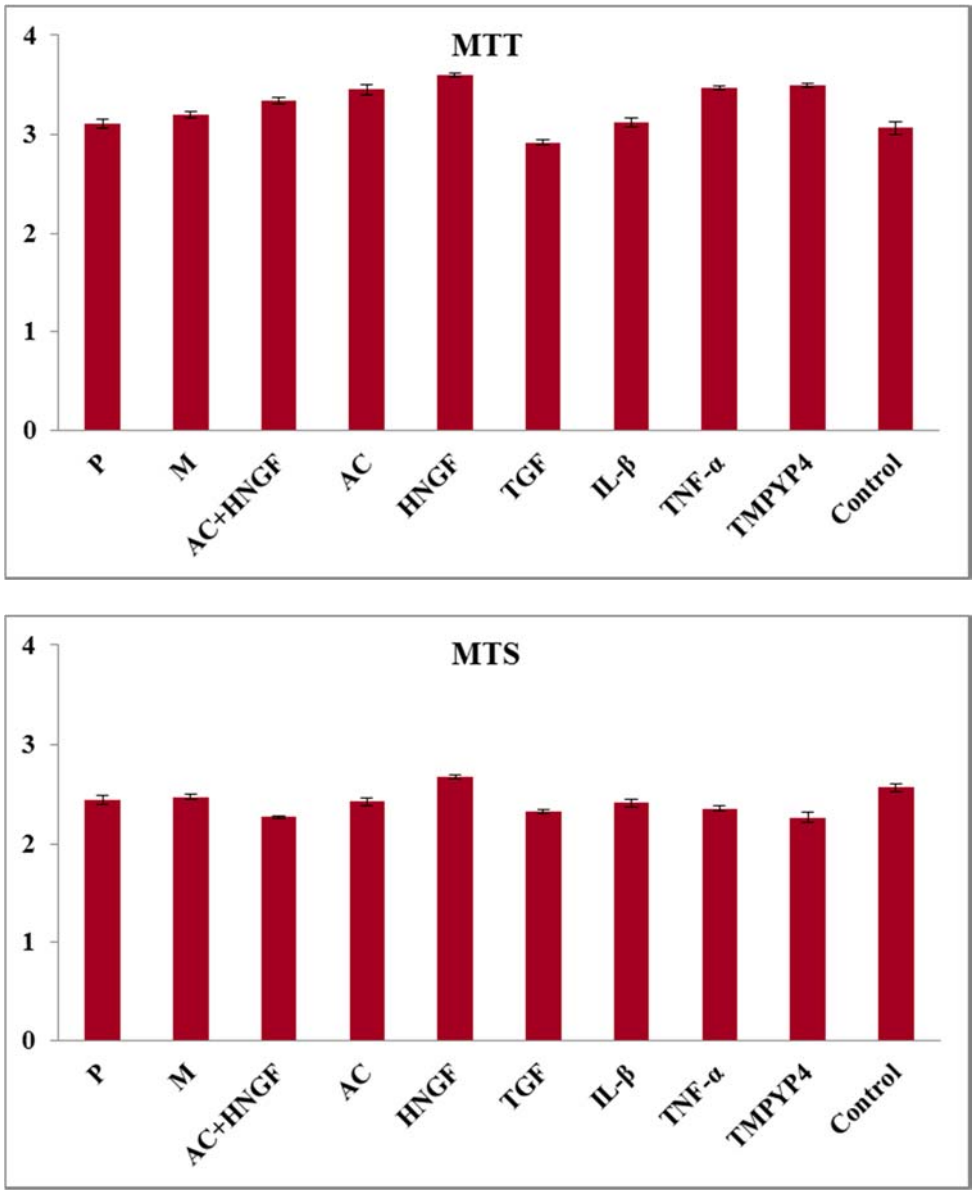


Figure 21. The Colorimetric MTT and MTS Proliferation Assays were used to Test the Viability of Transfected Cells Exposed to Varied Environmental Conditions. IMR-32 transfected cells (with GFP vectors, modified and wild type) were incubated for further 24 hours at the absence (control) or presence of following additives: AC, (AC+HNGF), HNGF, TGF, TNF-α, IL-β, TMPYP4. M condition is cells expose to normal complete medium without any additives. Control and P are Un-transfected and transfected cells with parent plasmid (pAcGFP1-N3) used as negative and positive controls, respectively. The resulted colored solution was detected at 490 and 570 nm using a spectrometer. The absorbance values represent the mean of two separate experiments of four readings and the corresponding standard errors of the means (s.e.m.).

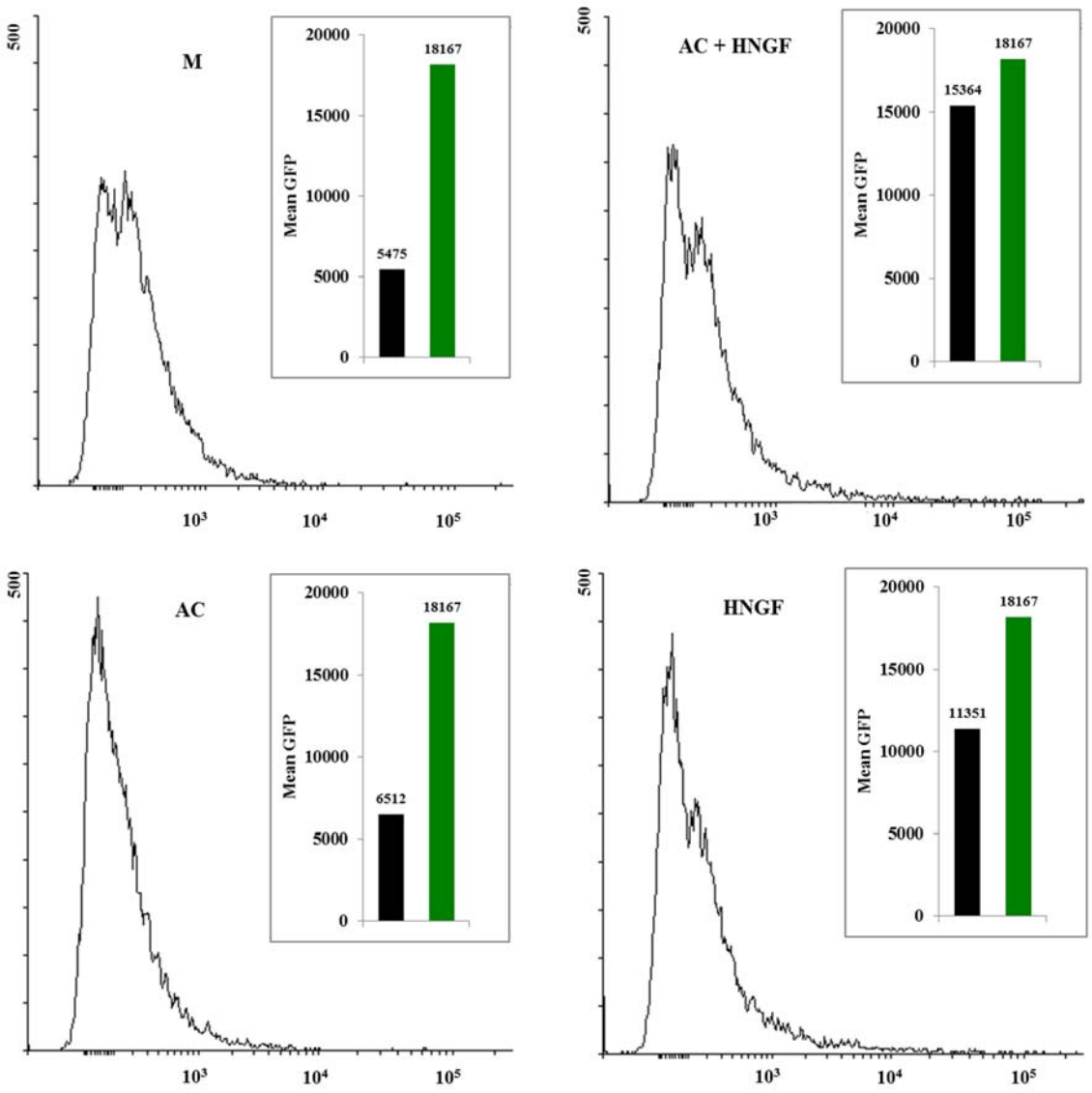


Figure 22. Flow Cytometry Analysis of the Human Active Promoter Sequence, part 1. IMR-32 cells were transfected with pAcGFP1-N- modified vector, followed by exposing to various environmental conditions (presence of Aconitine, HNGF, AConitine + HNGF, and normal complete media: M) for 24 h incubation. In each panel, histogram of fluorescence value of individual experiment is reported in black line. The horizontal bar indicates the cells reaching these fluorescence values out of total cells. The inset histograms represent the mean fluorescence of cells in each experiment in compare to positive control (GFP wild type). GFP expression in cells exposed to AC + HNGF is approximately similar to positive control.

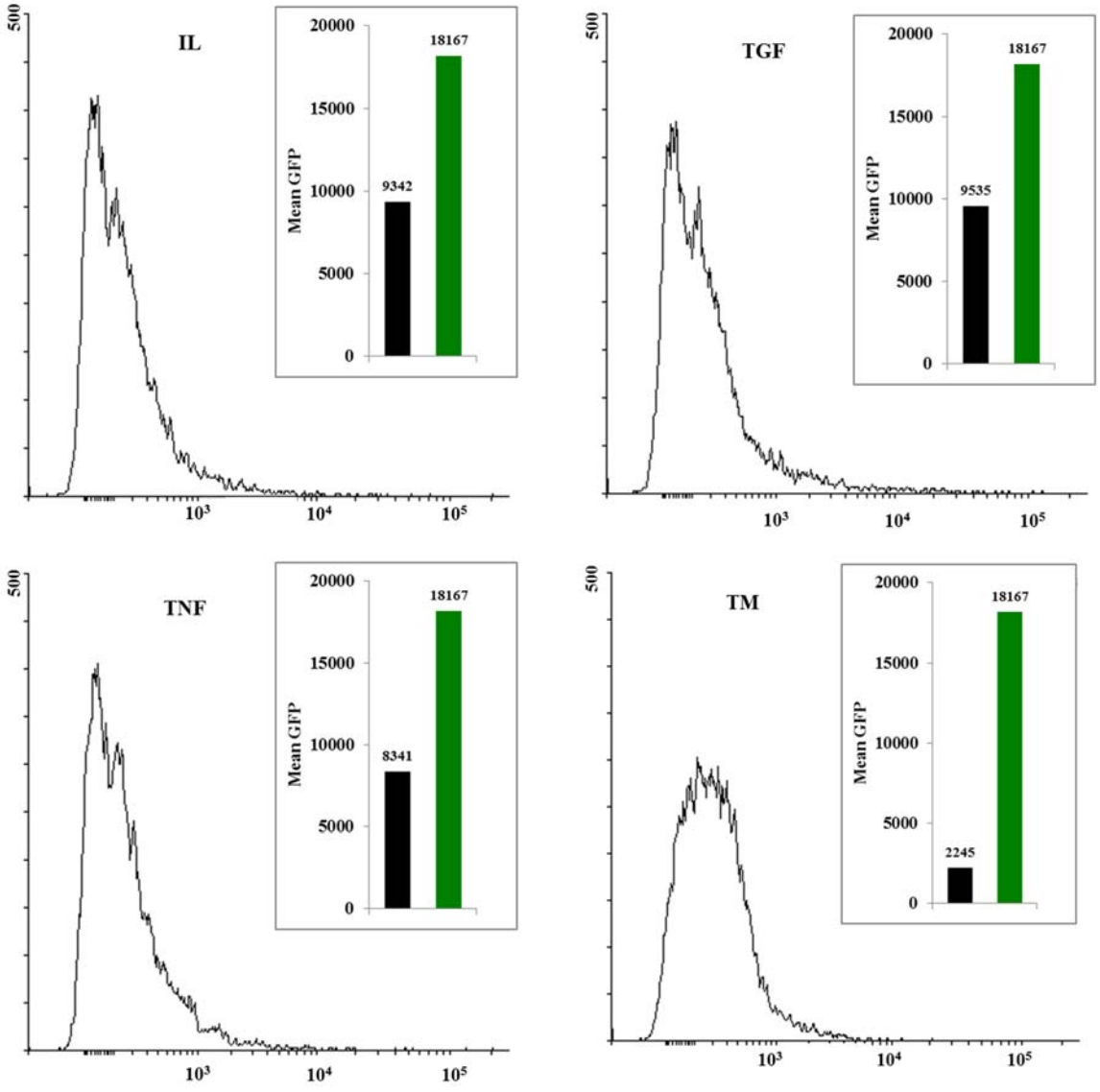


Figure 23. Flow Cytometry Analysis of the Human Active Promoter Sequence, part 2. IMR-32 cells were transfected with pAcGFP1-N- modified vector, followed by exposing to various environmental conditions (presence of IL-β, TGF, TNF-α, and TMPYP4) for 24 h incubation. In each panel, histogram of fluorescence value of individual experiment is reported in black line. The horizontal bar indicates the cells reaching these fluorescence values out of total cells. The inset histograms represent the mean fluorescence of cells in each experiment in compare to positive control (GFP wild type).GFP expression of cells exposed to TMPYP4, G-QPX stabilizer, is significantly lower than positive control.

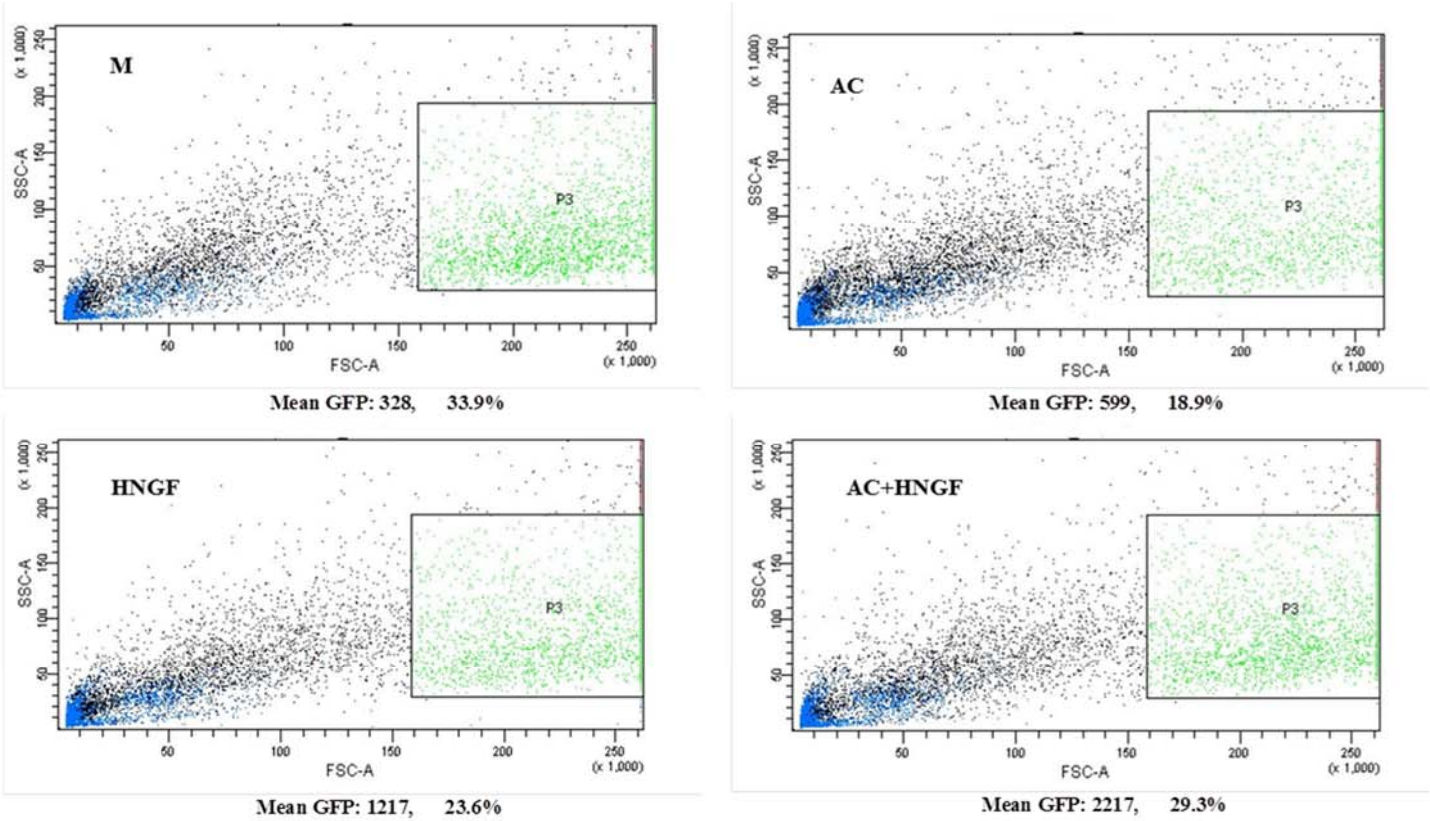


Figure 24. Quantitation of FITC-A using Flow Cytometry for P3 Population, part 1. IMR-32 cells were transfected with pAcGFP1-N- modified vector, followed by exposing to various environmental conditions for 24 h incubation. In each panel, plot dot of fluorescence modified GFP is reported. The P3 gate is represent in green dots. The inset lines represent the mean of GFP in the P3 population in each experiment relative to positive control. The percentage value indicates the cells reaching these fluorescence values out of total cells in P3 population. P3 population of cells showed the maximum GFP expression (2217) in presence of aconitine with existence of HNGF.

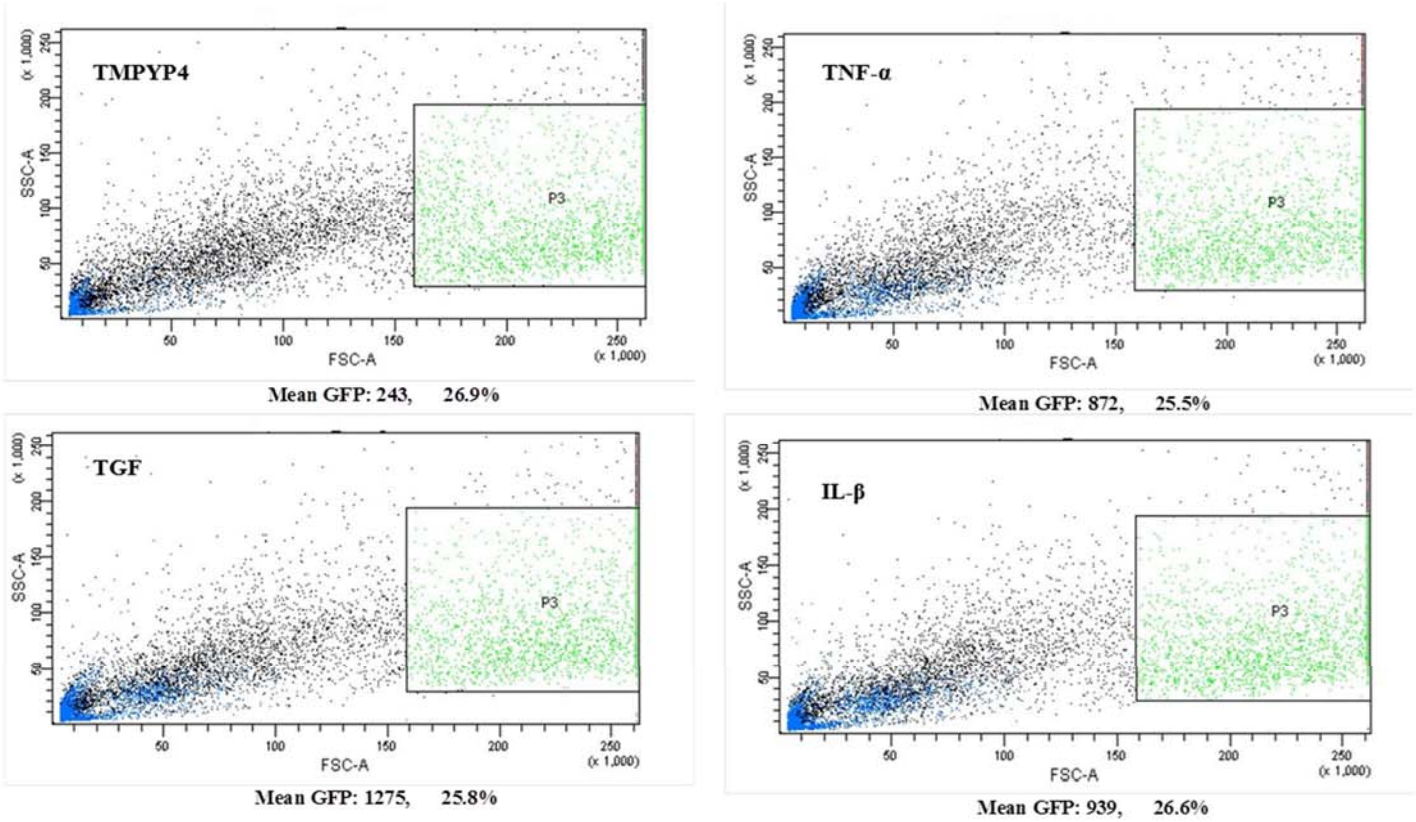


Figure 25. Quantitation of FITC-A using Flow Cytometry of P3 Population, part 2. IMR-32 cells were transfected with pAcGFP1-N- modified vector, followed by exposing to various environmental for 24 h incubation. In each panel, plot dot of fluorescence modified GFP is reported. The P3 gate is represent in green dots. The inset lines represent the mean of GFP in the P3 population in each experiment relative to positive control. The percentage value indicates the cells reaching these fluorescence values out of total cells in P3 population. P3 population of cells showed the minimum GFP expression (243) in presence of TMPYP4.

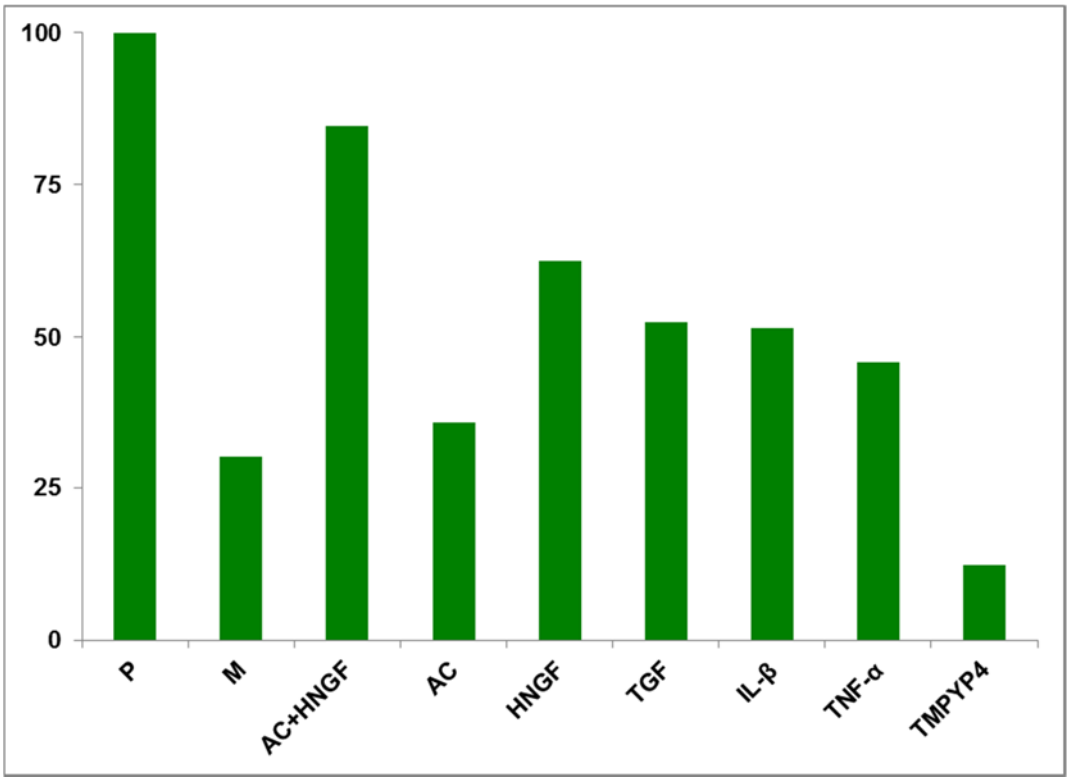


Figure 26. The Percentage of Mean GFP Fluorescence Intensity using Flow Cytometerty. IMR-32 transfected cells (with GFP vectors, modified and wild type) were incubated for further 24 hours in the presence of following additives: AC, (AC+HNGF), HNGF, TGF, TNF-α, IL-β, KCl:10mM, TMPYP4. M condition is cells expose to normal complete medium without any additives. P is represents cells transfected with parent plasmid (pAcGFP1-N3) used as positive controls. The bar chart shows that the percentage of GFP expression in cells exposed to AC+HNGF is close to positive control while the TMPYP4 shows the lowest mean GFP intensity percentage.

3.4.6 Quantitating the ChAT mRNA expression using RT-PCR

mRNA expression of the human ChAT gene in neuroblastoma (IMR-32) cell line was measured using q RT-PCR after 4hrs of treatment. The half-life of ChAT mRNA is 7 hr so to get the maximum amount of ChAT mRNA, extraction and purification was done after 4 hr.

To obtain $\Delta\Delta C_t$ values for all the samples, the data were analyzed using user defined threshold. The threshold definition is one of the most important steps in qPCR analysis; therefore, extra care was taken in order to use a suitable threshold value. One of the challenges in defining threshold came from the huge difference in amplification plot of TMPYP4 compared with the rest of samples. Figure 27 shows such plot for one AC sample and one TMPYP4 sample where no significant change in reaction can be detected for TMPYP4 (the TMPYP4 specimen line does not cross the threshold line).

Similar expression (mRNA amplification) rates were observed (Fig. 28 - Top), demonstrating that mRNA expression of the ChAT gene significantly increase when cells exposed to medium containing aconitine with the presence of HNGF and a dramatic knockdown was observed in mRNA expression of cells treated with TMPYP4. Therefore, it could be interpreted that TMPYP4 should have notable impact on the human ChAT gene expression, verifying the stabilizing function of TMPYP4 on G-QPX formations in promoter regions. Indeed, Na^+ as a destabilizing ions melt the G-QPX structures in the promoter through movement of the Na^+ into the cells (ion channels opening), resulting in elevation of ChAT gene mRNA production.

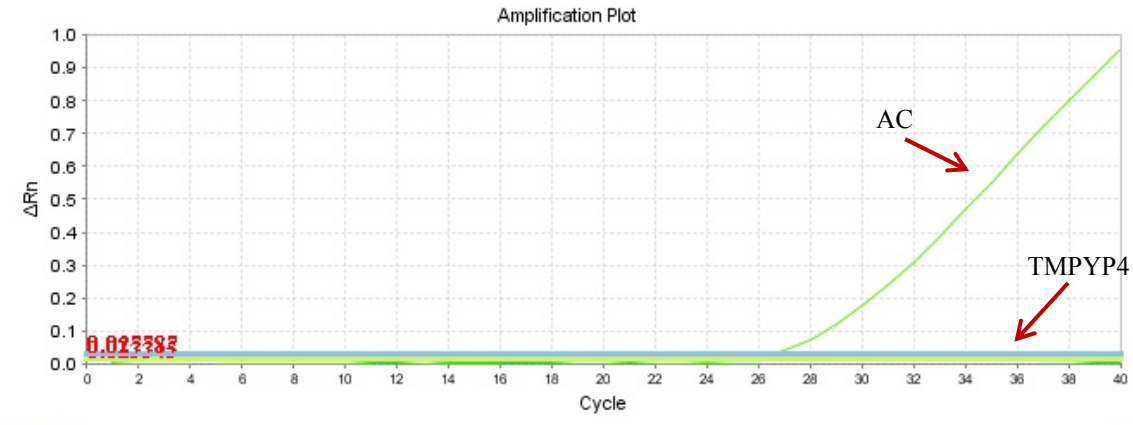


Figure 27. Amplification Plot of RT-qPCR for a Typical AC and TMPYP4 Specimens. Blue line indicates the threshold line and the two green line indicate the AC (top) and TMPYP4 (bottom) specimens.

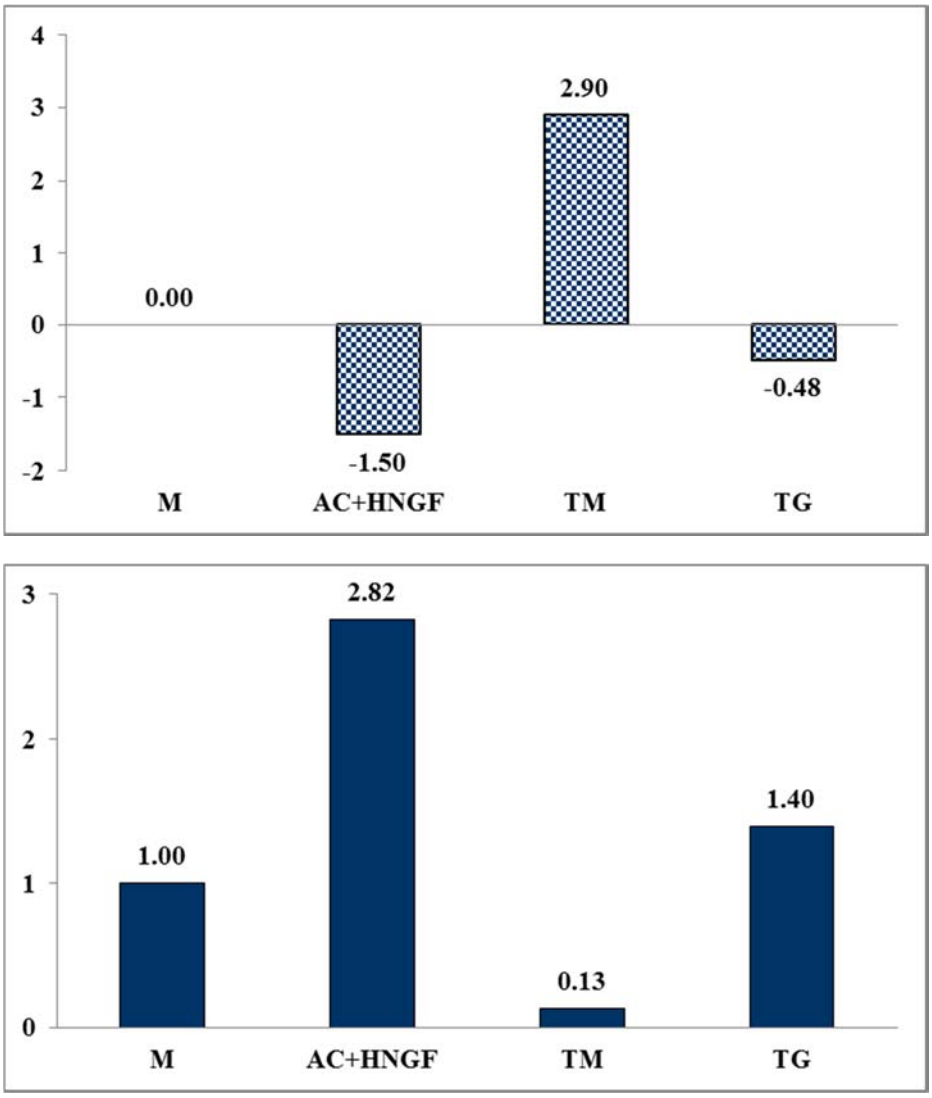


Figure 28. Quantitative Comparison of ChAT Gene Expression using Real Time RT-PCR. Human neuroblastoma cell line (IMR-32) exposed to the complete medium in presence of different additives: Aconitine in combination with HNGF, TGF, and TMPYP4. After 4 hr exposing to different environmental conditions, ChAT mRNA extracted, purified for qRT-PCR. Cells exposed to normal complete medium used as a control. USB1 gene has been used as a reference gene for Q-PCR analysis. Top) $\Delta\Delta C_t$ results, Bottom) Folding measurement. Compared to untreated cells, an approximate 87% ChAT gene knockdown was observed in the sample treated with TMPYP4.

3.5 Conclusion and Discussion

Computational studies have suggested that promoter of human genes have a great potential to form G-QPX motifs. However, a few in vitro studies have shown the functional role of G-QPX structures. For instance, G-QPX formation upstream the P1 promoter of C-MYC has shown acting as a repressor [15].The G-QPX interaction with TFBS leads to proposing the regulatory function of G-QPX cellular activity. Generally, most of TFBS bind to promoter sequence upstream of the TSS up - regulate the transcription activity so formation of G-QPX in the TFBS recognition sites will lead to down-regulation the transcription activity. However, there are some TFBS that exhibited the binding tendency to G-QPX structures in promoter [83] which in this scenario, they could be important factors for specific TFBS binding. Under this circumstance, down-regulating transcriptional activity required inhabitation of G-QPX formations. To conclude, since G-QPX structures are behaving differently in biological systems due to diversity in their sizes, topology, loop size and importantly their interactions with other molecules or essential factors, it is important that study these unique structures individually and not draw a precise conclusion about their formation, structures, stability and their cellular functions.

Here, we develop a novel construct (pAcGFP1-N- modified), derivative from wild type pAcGFP1-N, which the whole promoter regions (approximately 600 bp) substituted with the active promoter region of human ChAT gene (174 bp). The modified GFP vector contains the G-rich regions in its promoter with great potential to form G-QPX structures. Formation of nano-switch G-QPX structure in promoter have been suggested regulates

transcription machinery. Therefore, to validate formation, stability and functional effects of these unique structures we did in vitro reporter gene measurement using neuroblastoma cell line (IMR-32). IMR-32 cell line is a type of cell line that they are not easy to transfect with foreign DNA unlike other type of cell lines such as HEK, A549, and PC12. To get the high efficient transfection, transfast reagent kit has been used to effectively transfect the cells with Modified GFP vector. HNGF, TGF, IL-β and TNF-α has been individually used as essential factors that facilitate binding the TFBS, in particular Sp1 and AP2-α, to the active promoter and up-regulate the GFP expression. In vitro fluorescence GFP intensity measurement showed enhancement compare to the basic level of expression (medium exposure without any additives). Aconitine as a Na^+ - channel dependent drug in the presence of HNGF led to significant increase in GFP expression similar to GFP expression of wild type plasmid. This result supports the fact that monovalent cations have obvious impact on stability / instability of G-QPX motifs. The mechanism of drug – dependent channel is that aconitine opens the Na^+ channels which lead to influx of Na^+ ions into the cell, and subsequently melt the G-QPX structures, followed by induction of GFP reporter genre. More importantly, it confirm our novel hypothesis, claiming that neural action - potential mechanism regulate G-QPX formation/ deformation due to movement of monovalent cations across the membrane, which also consist with gene silencing and expression during the neuron resting and firing. Flow cytometry analysis and fluorescence microscopic imaging have been used and both analysis further confirm the results (Fig. 20, 24 and 25).

TMPYP4 as G-QPX stabilizer compound has been used in literature in most of cases while the G-QPX structural and sequential diversities have found other small selective ligands molecules as stabilizer. Therefore, to get a better understanding of our identified multiple G-QPX, a separate transfection experiment followed by treatment with TMPyP4 was conducted. The result clearly showed that the G-QPX formation significantly diminish the fluorescence GFP expression in IMR-32 cells.

Similarly, the quantitative real time RT-PCR assay was operated, and a significant inhabitation of the human ChAT gene was observed, 87% ChAT gene knockdown when cells exposed to TMPYP4 and significant increase in ChAT gene expression when cells treated with aconitine in presence of HNGF. Figure 28 (bottom) summarizes the analyzed data where 2.8 fold increase in expression was observed for AC while TG samples demonstrated only 1.4 fold increase.

The result demonstrates that formation of G-QPX affect binding of TFBS (AP2-αand Sp1) to their binding site which makes an influence on transcriptional activity of ChAT gene. No cytotoxicity observed after exposing cells to various conditions.

CHAPTER IV

PHYSICAL STUDIES OF G-QPX MOTIFs FORMATION AND THEIR STABILITIES

4.1 Introduction

Nucleic acids are polymorphic molecules which can adopt different types of conformations classified as B-form and non B-form structures. The non B-form arrangements are usually formed on specific sequence motifs which associate with various molecular interactions. To date, variety of non-B nucleic acids conformations have been identified such as B-DNA, Z-DNA, H-DNA, holiday junction and i-motifs. Among these secondary structures, intensive studies focused on G-QPX formations and their properties in various genomic functional regions. In gene promoters, DNA sequence analysis showed the existence of guanine-rich regions [59, 84].

G-QPX structure can form in guanine-rich sequence consists of planer arrangement of four guanine residues via Hoogsteen pairing [85]. It is guanines ability to generate self-associate hydrogen-bounds in aqueous environment, the G-quartets. The G-quartets stacking creates a core by holding intervene sequences together and arrange single-strand loops on exterior of the core [86]. These sequences usually consist of at least four guanine residues with variable length. Several biophysical methods have been used to characterize G-QPX structures such as X- ray diffraction (XRD), circular

dichroism (CD), and nuclear magnetic resonance (NMR). XRD and NMR are two major methods to reveal atomistic details of macromolecules structures. However, these superior techniques have some limitations. For instance, samples used for these analyses must contain high concentration of nucleic acid, and short or single dominant DNA fragments in order to determine the atomic structure [84]. The third common technique is CD spectroscopy, which works based on the differential absorption of left-handed circularly polarized light (L-CPL) and right-handed circularly polarized light (R-CPL) and it occurs when a molecule contains one or more chiral chromophores (light-absorbing groups).

$$\text{Circular dichroism} = \Delta A(\lambda) = A(\lambda)LCPL - A(\lambda)RCPL, \text{ where } \lambda \text{ is the wavelength}$$

Although CD does not provide information regarding atomic resolution structure of macromolecules but it has distinctive advantages. First, it's a fast, simple and extremely sensitive technique which requires small amount of nucleic acids concentrations. Second, both long and short DNA molecules can be effectively evaluated under various environmental conditions such as presence of various salts, different temperature and pH. Change in environmental condition has a critical impact on conformational transition of DNA molecule [87]. DNA conformational properties are important as basics in gene expression regulation [88]. CD spectra demonstrates the conformations of DNA but for analyzing the data, one must consider the whole spectrum since the presence and intensity of a single peak can be influenced by changes in the guanine tetrad staking interactions as result of different coordination of monovalent

cations [89]. G-QPX structures can be stabilized by monovalent cations (e.g., Na^+, K^+), resulting in formation of a parallel or antiparallel orientation of four DNA strands. In fact, occupation of the central cavity by monovalent cations neutralizes the electrostatic repulsion of guanine oxygens and forming stable structure [90]. It has been suggested that potassium ions stabilize the G-QPXs better than sodium ions since they are placed in the middle of tetrad planes. Although the CD spectra is a powerful tool to study macromolecules conformational properties but interpretation of spectra for structural conclusion needs to be obtained with caution. As mentioned before, it is important to note that an accurate interpretation only can be achieved by tracing the whole spectrum region, and not following a single band (eg. the band at 260 or 290nm) [84]. In most be noted, while the interaction between G-tetrad stacks, and their coordination with cations and/or ligands (as stabilizing / destabilizing molecules) have been previously studied using biophysical methods (XRD, CD, NMR); interaction of G-QPX formation at the single - molecule level has not been investigated.

Atomic force microscopy (AFM) has been used extensively to visualize DNA at the single-molecule level with only minimal sample preparation [91, 92]. In fact, minimal sample preparation involved in AFM imaging allows visualizing G-QPXs in the context of a more physiologically relevant DNA molecule (ex: visualizing ability of G-rich sequence in a single strand to form a condense structure and the loops can be symmetric or asymmetric) [93]. One of the key steps in imaging biological molecules using AFM is sample immobilization due to the inherent negative charge of nucleic acids (DNA or RNA) [94]. A number of surface modification techniques has been developed to

form a positively charge substrate to electrostatically hold the nucleic acids. Mica is the most widely used substrate for deposition of various biological objects (nucleic acids in particular) [95, 96] followed by gold films [97, 98] and glass [99]. Mica (primarily muscovite mica) is a layered mineral with the smooth surface which makes it appealing substrate for biological AFM imaging [100]. Moreover, in the process of preparing sample for AFM imaging, the pretreatment of specimen with Mg^{2+} cations is another important step [101-103].

Moreover, small compound molecules such as pyridostatin [104], TAP1 [105], TMPYP4 [21], PIPER [24], BRACO-19 [23], have been used to bind to G-QPXs with high affinity and affecting stabilizing or destabilizing. Here we used TMPYP4 as a common compound to confirm the formation of G-QPX structures in an active promoter region of ChAT gene. The structure of this compound is shown in figure 29.

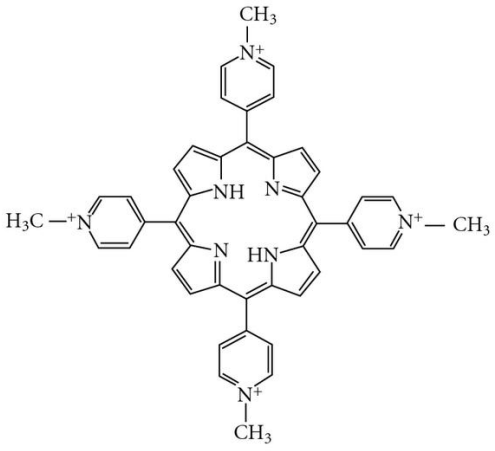


Figure 29. Structure of TMPYP4.

4.2 Materials and Methods

4.2.1 DNA sample

DNA oligonucleotides were HPLC-purified purchased from IDT (Table 9), which were stored at −20 °C. DNA concentration was measured by nano-drop ND-2000c spectrophotometer (Thermo Fisher Scientific Inc.,Waltham,MA).

4.2.2 DNA substrate for AFM imaging

The recombinant plasmid used in AFM experiments, named as pAcGFP1-N3-modified (bp), designed and provided by our lab (Appendix C, Fig.1), is derivative of pAcGFP1-N3 (4722 bp). The pAcGFP1-N3- modified plasmid (4189 bp) has a 147 bp inserted sequence of the active promoter region of ChAT gene, which contains multiple QPXs structures just upstream of TSS (Appendix C, Fig.2). To get a better G-QPX visualization, modified plasmid digested using restriction enzyme ApaLI and AFIII (Appendix C, Table1). The digested DNA fragment from 4350 bp to 1153 bp of modified plasmid (cut- pAcGFP1-N3- modified plasmid, 1508bp), containing active promoter region of ChAT gene, was amplified using PCR. The total volume of PCR reaction mixture contained 2 μL PCR buffer (10X), 500 μmol/L dNTP mix, 0.15 μM each primer (shown in Table 1(AF-ChAT S & AF-ChAT A), 1 μL DNA template, 1 U Taq DNA polymerase, and nuclease-free water. The protocol used for PCR includes a denaturing cycle of 2 min at 95 °C, 35 cycles of PCR (95 °C for 1min, 58 °C for 1min, 72 °C for 2 min), and then 72°C for 10 min, followed by holding at 4 °C. The PCR then loaded in

1% agarose gel. After electrophoresis, the gel was imaged using UV-shadowing (Appendix C, Fig.3).

4.2.3 Circular dichroism (CD)

CD spectra were recorded at 20 °C on Olis DSM 17 spectropolarimeter; figure 24 (Jasco, Easton, MD) using a quartz cell of 1 mm path length. The instrument scanning speed was set at 100 nm/min with a response time of 1 s and bandwidth of 2 nm, over a wavelength range of 220 - 320 nm. The DNA oligonucleotides were prepared in 5 μM solution in buffer containing Tris–HCl (10mM, pH7.5) and different concentrations of KCl and/or NaCl. The DNA samples were heated at 95 °C for 5 min, and then slowly cooled to room temperature overnight. Data were baseline-corrected for signal contributions due to the buffer. Figure 30 shows the schematic of CD instrument.

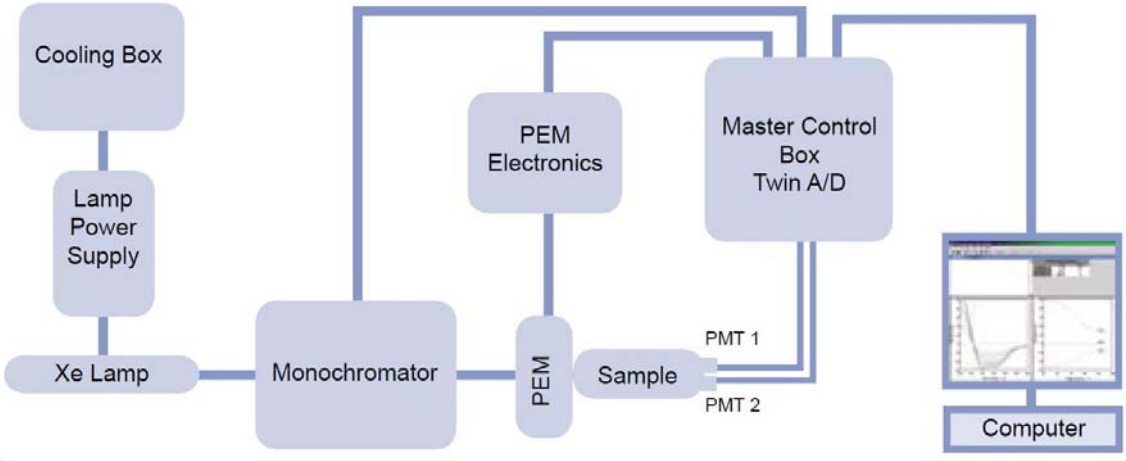


Figure 30. Schematic of an Olis CD Instrument.

Table 9. Oligonucleotide Sequences have been used in Present Study.

Name	Sequence (5`→3`)
ChAT G17	GAGGCTGGTGTGGGGCG
ChAT G17-2	AAGGGCGGGGCCTGGCG
ChAT G30	CTGGAGCGGGTCCTCAGGCTCCCCCGGGCG
ChAT G29	GGGGATGCCGCCCGGGGGAGCCTGAGGAC
ChAT G21 S	CCCGGGGGAGCCTGAGGACCC
ChAT G21 A	GGGTCCTCAGGCTCCCCCGGG
AF- ChAT S	TTGGAGCGAACGACCTACA
AF- ChAT A	CGCTCACTTGTACAGCTCAT

4.2.3 AFM imaging

Mica disks (Ted Pella, CA, USA) glued to 13-mm steel pucks were used as substrates for AFM imaging. Prior to DNA deposition, the top layer of the mica cleaved using Scotch tape to reveal an atomically flat surface. Purified DNA was diluted to a concentration of ~0.5×10^{-14} mol/μL in buffer containing [10 mM Tris–HCl (pH 7.5)] in the presence of KCl, 100 mM or TMPYP4, 500mM, and was heated to 95°C for 5 min followed by slow overnight cool down to the room temperature (in addition to 95°C, 37°C for 24hr and 37°C for one week is also used). Immediately prior to deposition, 10 mM $MgCl_2$ was added to the imaging buffer containing DNA. Finally, 25 μL of DNA solution was deposited onto freshly cleaved mica disks and incubated at room temperature for 5 min. Excess solution was gently rinsed off with MilliQ water (Millipore System, MA, USA) and the water wicked from the surface using tissue paper. The disks were then dried under a stream of nitrogen gas.

AFM imaging was performed with a large stage 5600LS AFM (Keysight Technologies) in tapping mode with TAP-300 silicon cantilever tips (Budget Sensors, 40 N/m nominal spring constant, 300 kHz nominal frequency). Images were processed and analyzed with Gwyddion and ImageJ open-source software. Figure 31 shows the schematic of bimodal AFM instrument.

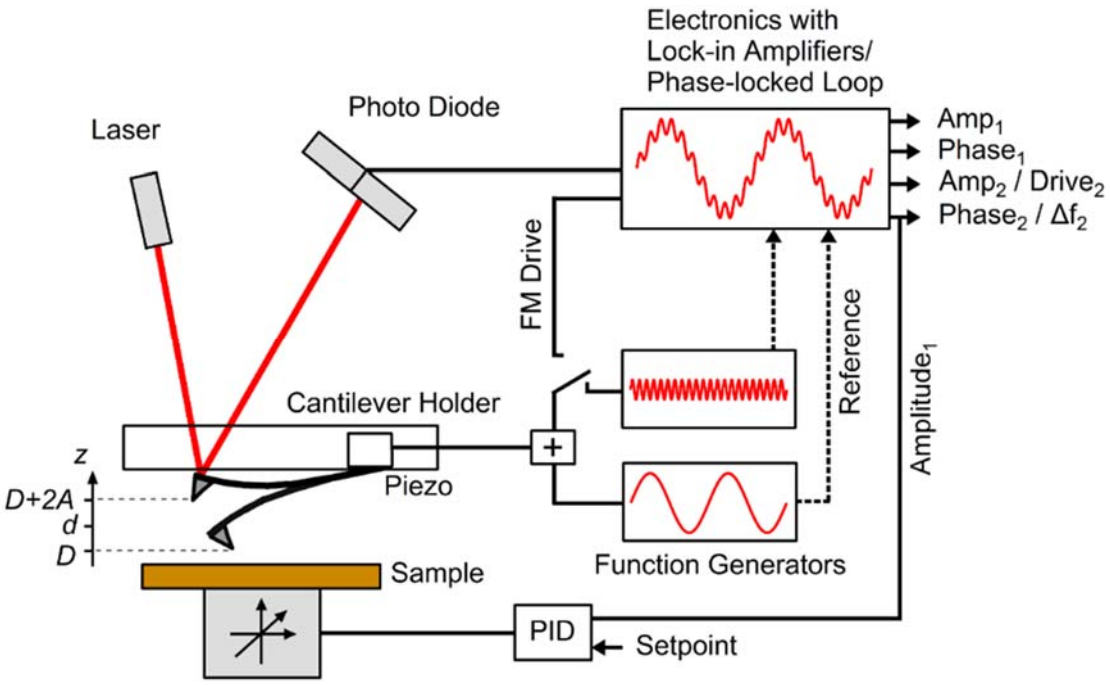


Figure 31. Schematic Setup of Bimodal AFM.

4.3 Results

4.3.1 Circular dichrosim

Difference in G-QPX sequences and their loop configuration lead into different conformational properties which in terns results in changes in position and amplitude of their CD bands. Figures 32 & 33 shows the CD spectra of ChAT G17-2 after 24h in solution (Tris-HCl: 10mM, pH: 7.5) containing various monovalent cations. Figure 34 showed the stability formation of ChAT G17 G-QPX in KCl solution. ChAT G 17 nucleic acid did not form reliable spectra in NaCl and LiCl solutions (data are not shown). Figures (35 & 36) and (37 & 38) show the ChAT G30 and ChAT G29 G-QPX structures, respectively, in different concentration of monovalent cations with Tris-HCl: 10 mM, pH:7.5 solution.

Figures 39 and 40 demonstrate that the intermolecular G-QPX could be formed within two complementary DNA double strands. This structure was observed upon formation of higher order structure within ChAT G-21S & ChAT G21A (called ChAT G21-ds) in Tris –HCl: 10mM, pH:7.5 solution containing different cations.

Generally, it is observed that unique G-QPX structures of the active ChAT promoter region show higher stability at K^+ solution.

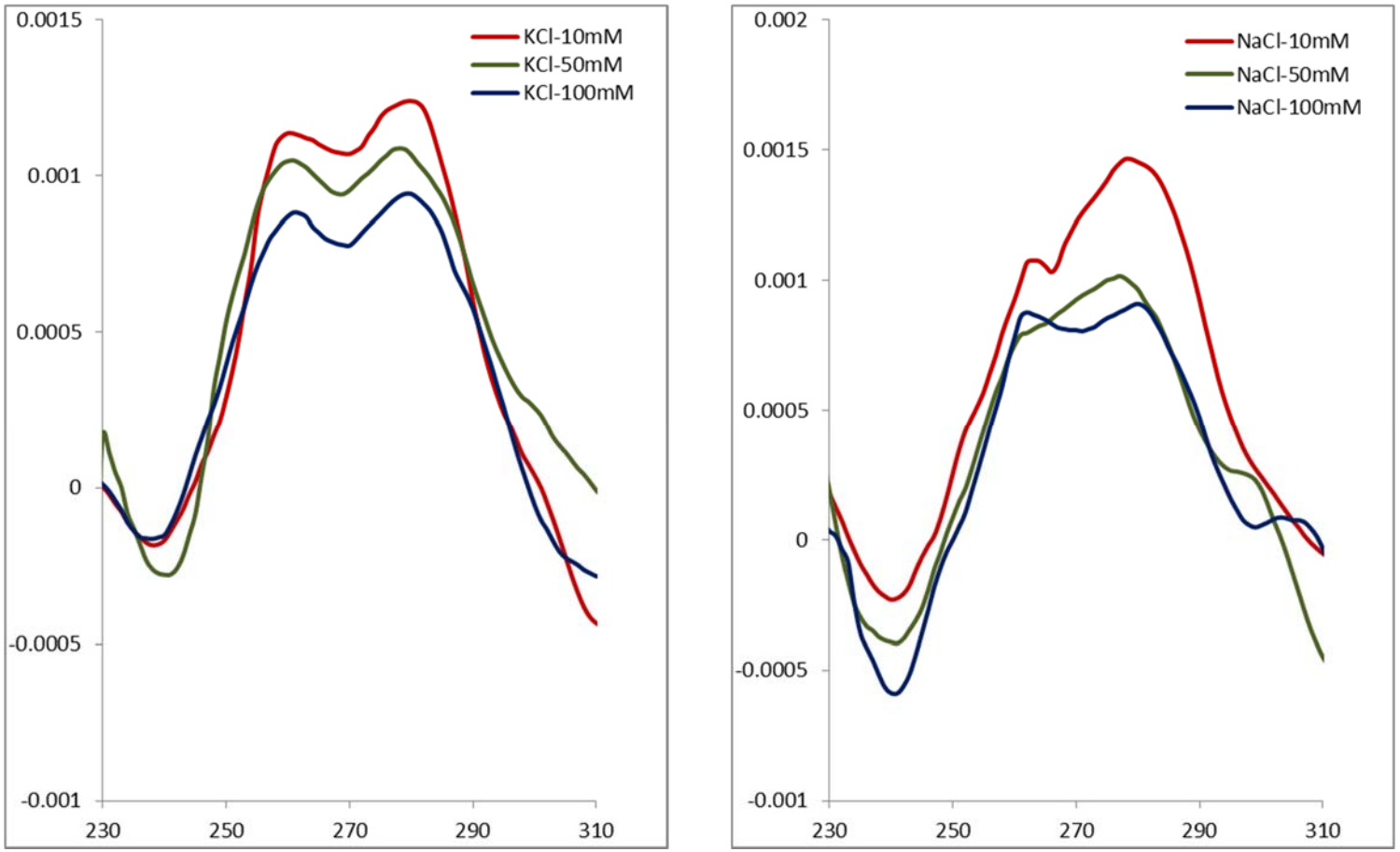


Figure 32. The CD Spectra of the ChAT G17-2 G-QPX. The spectra of 5 µM ChAT G17-2 in 10mMTris–HCl buffer solution (pH 7.5) with KCl (10, 50 & 100mM) (left) and with NaCl (10, 50 & 100mM) (Right). Data were baseline-corrected for signal contributions due to the buffer. The horizontal line represents the molar explicitly and vertical line represents wavelength.

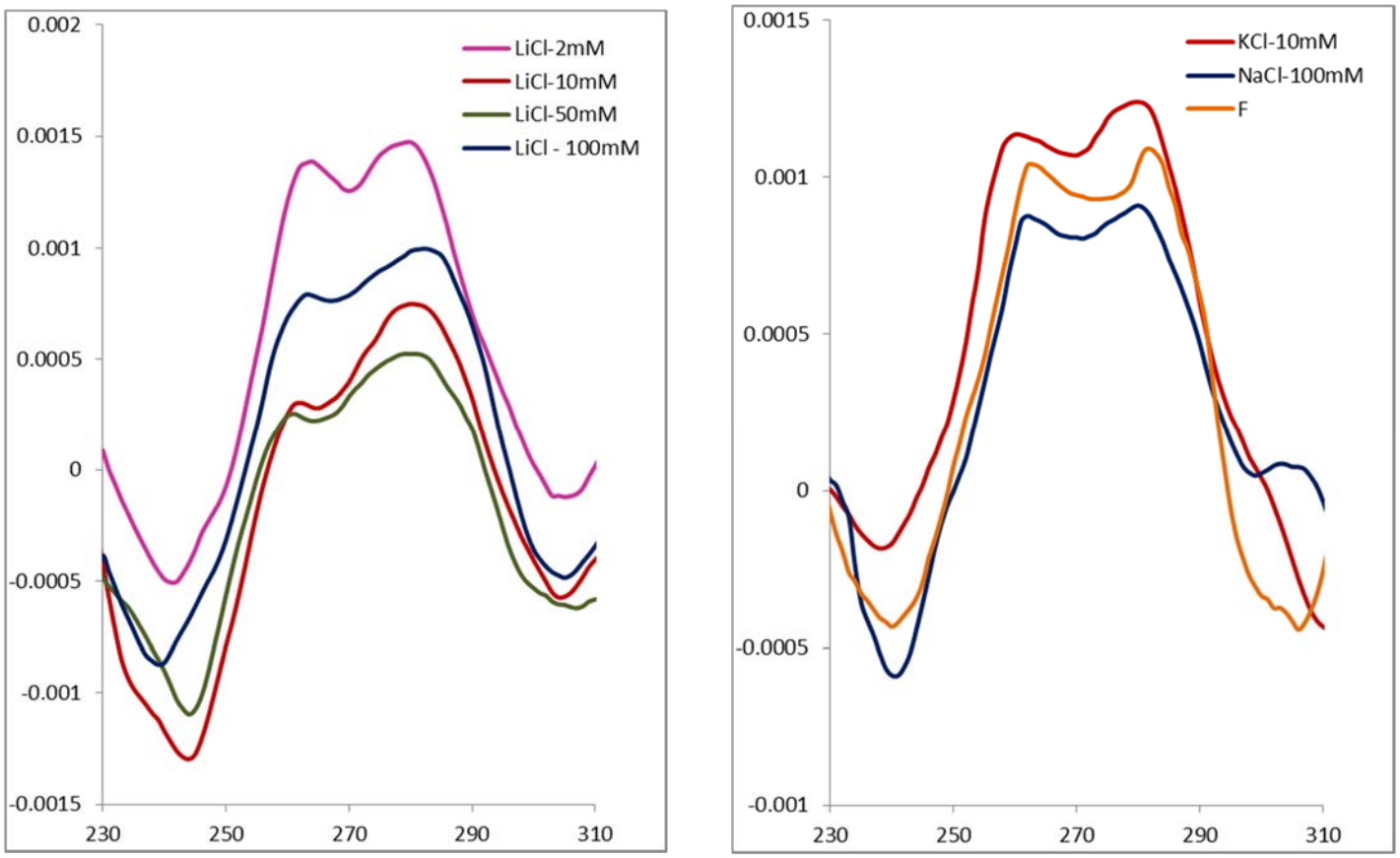


Figure 33. The CD Spectra of the ChAT G17-2 G-QPX. The spectra of 5 µM ChAT G17-2 in 10mMTris–HCl buffer solution (pH 7.5) with LiCl (2, 10, 50 & 100mM) (left) and with nerve firing condition (KCl : 10mM & NaCl: 100mM) (Right). Data were baseline-corrected for signal contributions due to the buffer. The horizontal line represents the molar explicitly and vertical line represents wavelength.

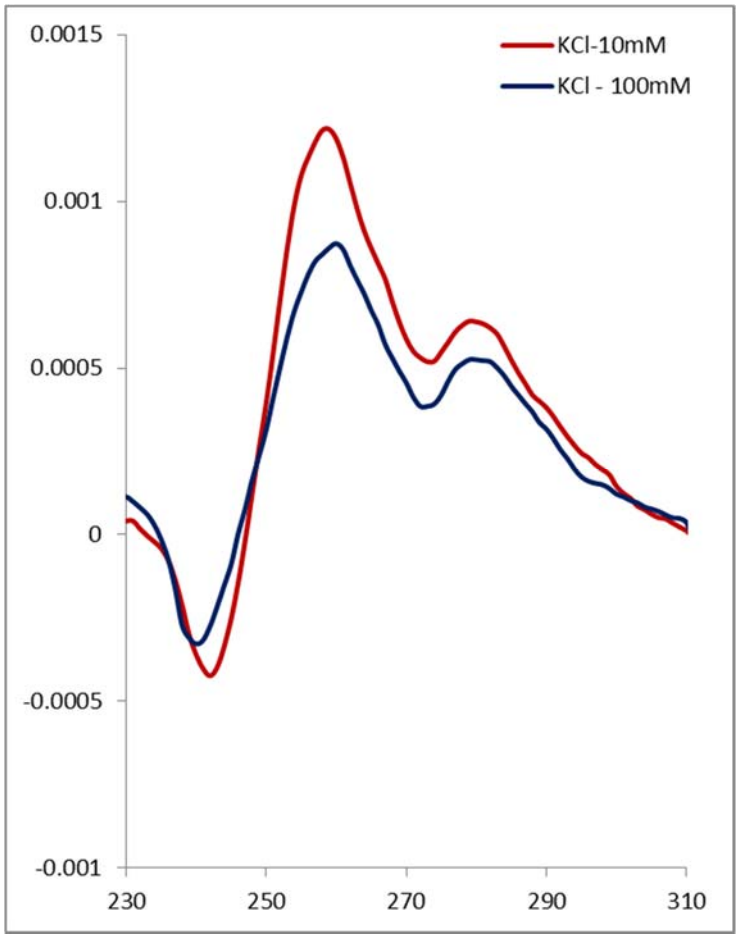


Figure 34. The CD Spectra of the ChAT G17 G-QPX. The spectra of 5 μM ChAT G17 in 10mMTris–HCl buffer solution (pH 7.5) with KCl (10 & 100mM). Data were baseline-corrected for signal contributions due to the buffer. The horizontal line represents the molar explicitly and vertical line represents wavelength.

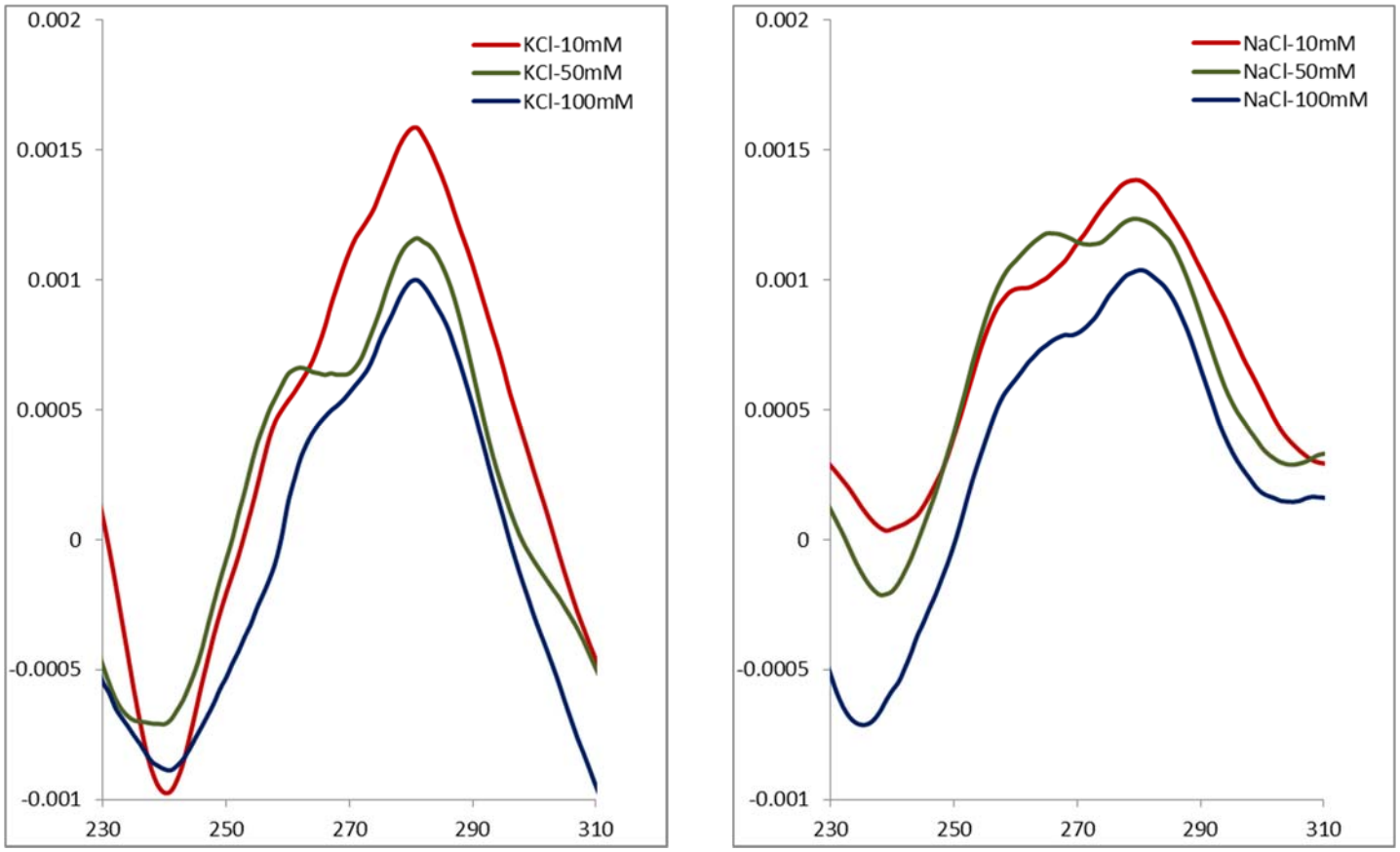


Figure 35. The CD Spectra of the ChAT G30 G-QPX. The spectra of 5 μM ChAT G30 in 10mMTris–HCl buffer solution (pH 7.5) with KCl (10, 50 & 100mM) (left) and with NaCl (10, 50 & 100mM) (Right). All spectra were measured after 24 h in buffer solution. Data were baseline-corrected for signal contributions due to the buffer. The horizontal line represents the molar explicitly and vertical line represents wavelength.

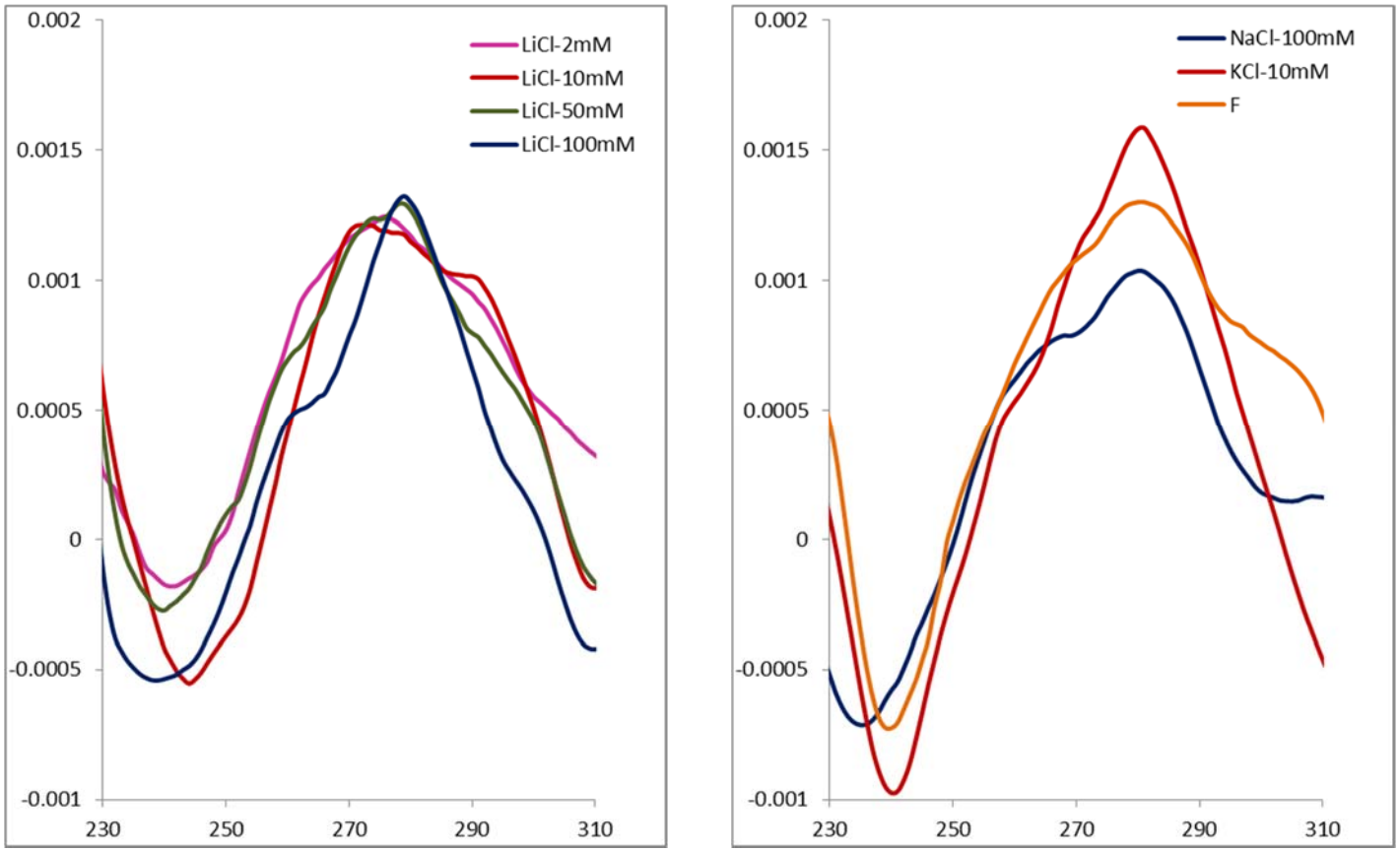


Figure 36. The CD Spectra of the ChAT G30 G-QPX. The spectra of 5 µM ChAT G30 in 10mMTris–HCl buffer solution (pH 7.5) with LiCl (2, 10, 50 & 100mM) (left) and with nerve firing condition (KCl : 10mM & NaCl: 100mM) (Right). All spectra were measured after 24 h in buffer solution. Data were baseline-corrected for signal contributions due to the buffer. The horizontal line represents the molar explicitly and vertical line represents wavelength.

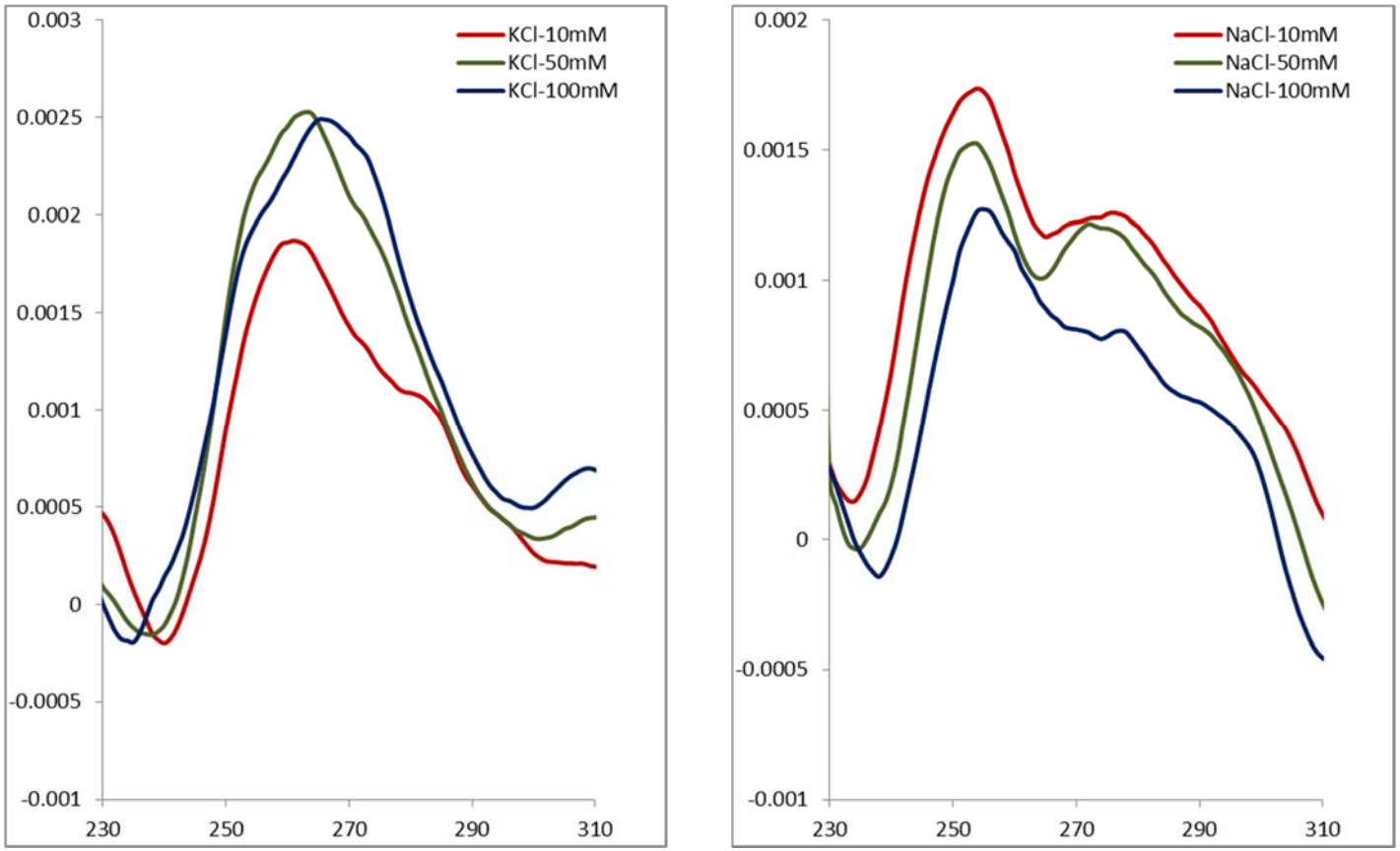


Figure 37. The CD Spectra of the ChAT G29 G-QPX. The spectra of 5 µM ChAT G29 in 10mMTris–HCl buffer solution (pH 7.5) with KCl (10, 50 & 100mM) (left) and with NaCl (10, 50 & 100mM) (Right). All spectra were measured after 24 h in buffer solution. Data were baseline-corrected for signal contributions due to the buffer. The horizontal line represents the molar explicitly and vertical line represents wavelength.

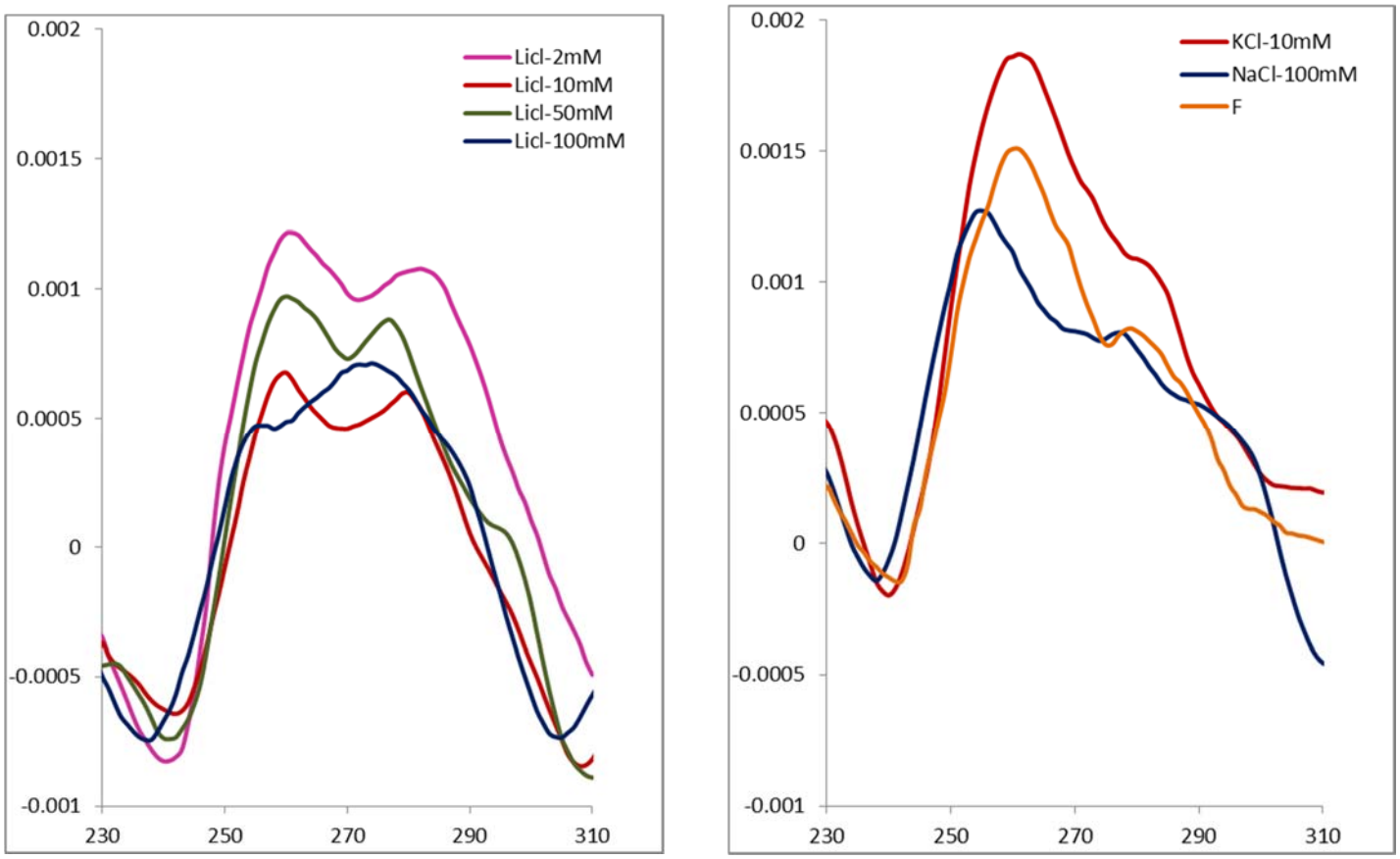


Figure 38. The CD Spectra of the ChAT G29 G-QPX. The spectra of 5 µM ChAT G29 in 10mMTris–HCl buffer solution (pH 7.5) with LiCl (2, 10, 50 & 100mM) (left) and with nerve firing condition (KCl : 10mM & NaCl: 100mM) (Right). All spectra were measured after 24 h in buffer solution. Data were baseline-corrected for signal contributions due to the buffer. The horizontal line represents the molar explicitly and vertical line represents wavelength.

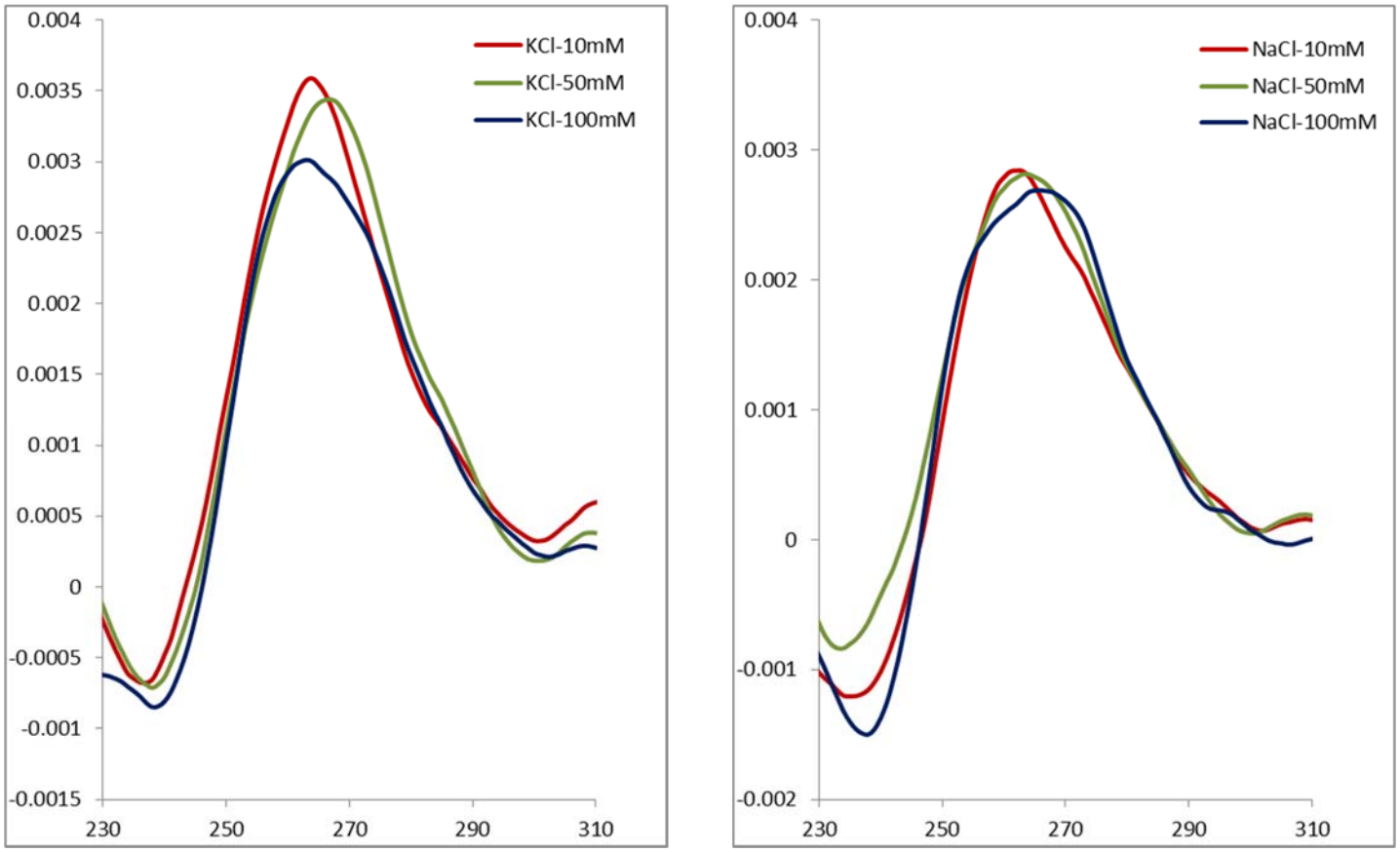


Figure 39. The CD Spectra of the ChAT G21- ds G-QPX. The spectra of 5 µM ChAT G21-ds in 10mMTris–HCl buffer solution (pH 7.5) with KCl (10, 50 & 100mM) (left) and with NaCl (10, 50 & 100mM) (Right). All spectra were measured after 24 h in buffer solution. Data were baseline-corrected for signal contributions due to the buffer. The horizontal line represents the molar explicitly and vertical line represents wavelength.

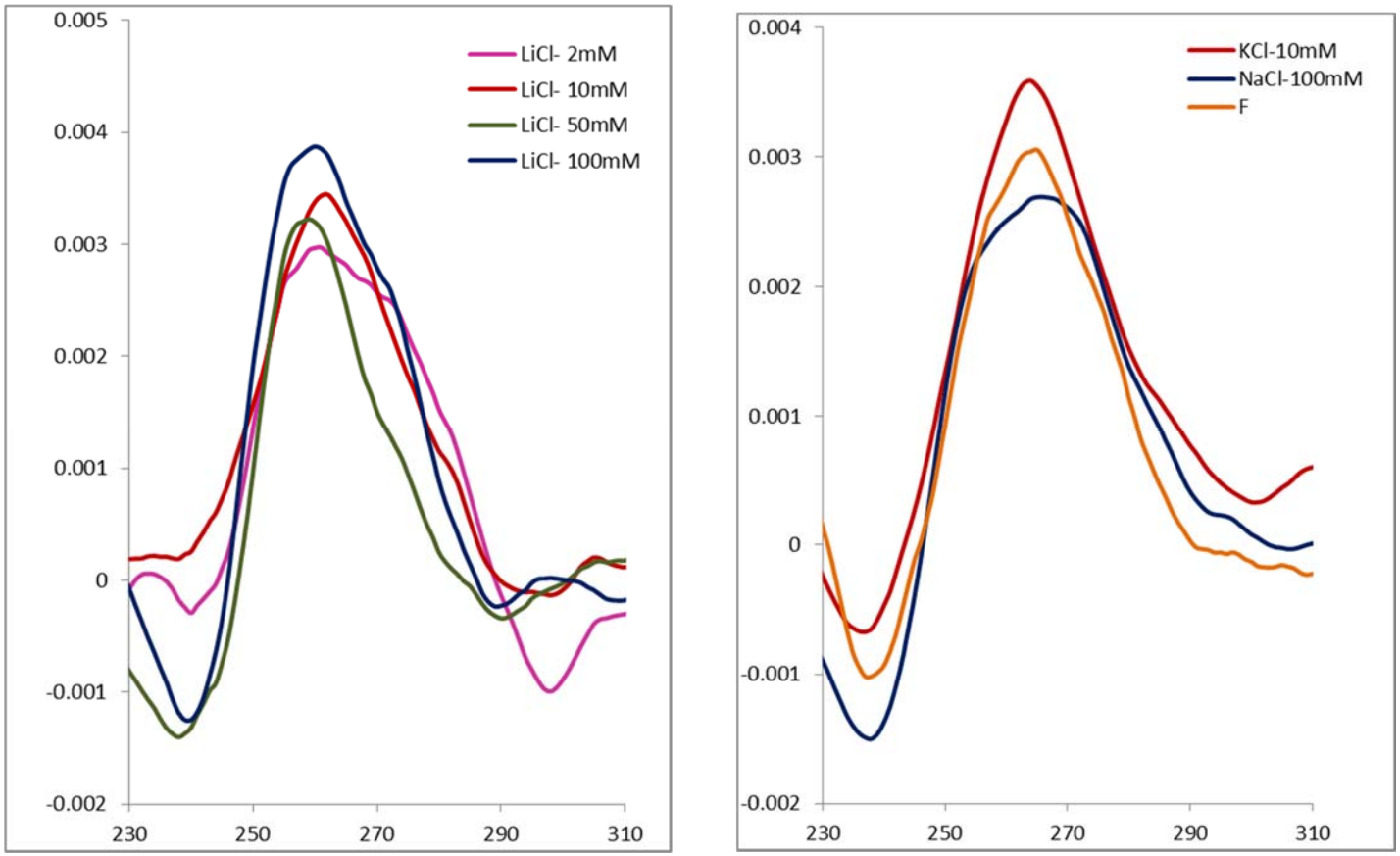


Figure 40. The CD Spectra of the ChAT G21-ds G-QPX. The spectra of 5 μM ChAT G21-ds in 10mMTris–HCl buffer solution (pH 7.5) with LiCl (2, 10, 50 & 100mM) (left) and with nerve firing condition (KCl : 10mM & NaCl: 100mM) Right). All spectra were measured after 24 h in buffer solution. Data were baseline-corrected for signal contributions due to the buffer. The horizontal line represents the molar explicitly and vertical line represents wavelength.

4.3.2 AFM imaging

The preparation for AFM imaging was performed with 5 min incubation on the mica surface followed by two times gently wash. The attempt to study cut- pAcGFP1-N3- modified plasmid, 1508 bp DNA fragment (cut-plasmid) in the solution containing K^+ ions did not lead into conclusive AFM image (Fig. 41-A). On the other hand, in presence of only stabilizing compound (TMPYP4, 500 mM) in the solution, stabilized G-QPX structures were observed (Fig. 41-B, Fig. 42). Figure 41-C shows the frequency distribution of cut-plasmid length in solution containing K^+. The processing of the image was done using image analysis software (ImageJ) where average length of 502 nm was measured. Similar effort for measuring the average length of cut-plasmid in solution containing TMPYP did not lead into fruitful data due to the fact that the structures in Fig. 41-B and 42 are clustered strands and it is not possible to differentiate an individual strand, even in ImageJ software.

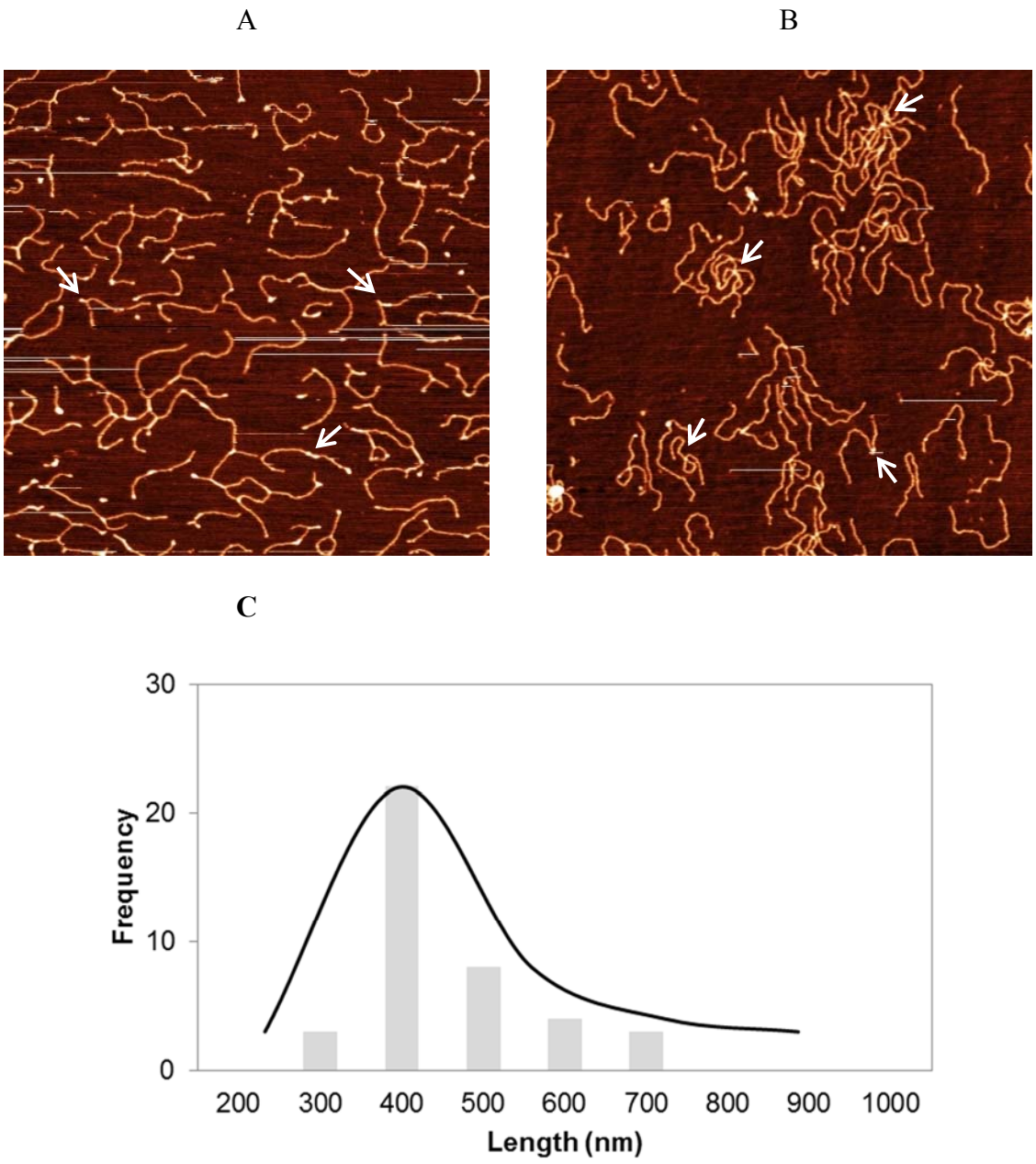


Figure 41. AFM Images of G-wires. A) G-wires on cut- pAcGFP1-N3- modified plasmid (1508 bp DNA fragment) incubated in buffer containing 100mM Kcl. B) G-wires on cut- pAcGFP1-N3- modified plasmid (1508 bp DNA fragment) incubated in buffer containing 500 mM TMPYP4. The white line arrows indicated the position of G-wires. The image of G- wires is 2x2 μm^2 at 1.2-nm height scale. C) The curve indicates the fitted Gaussian function.

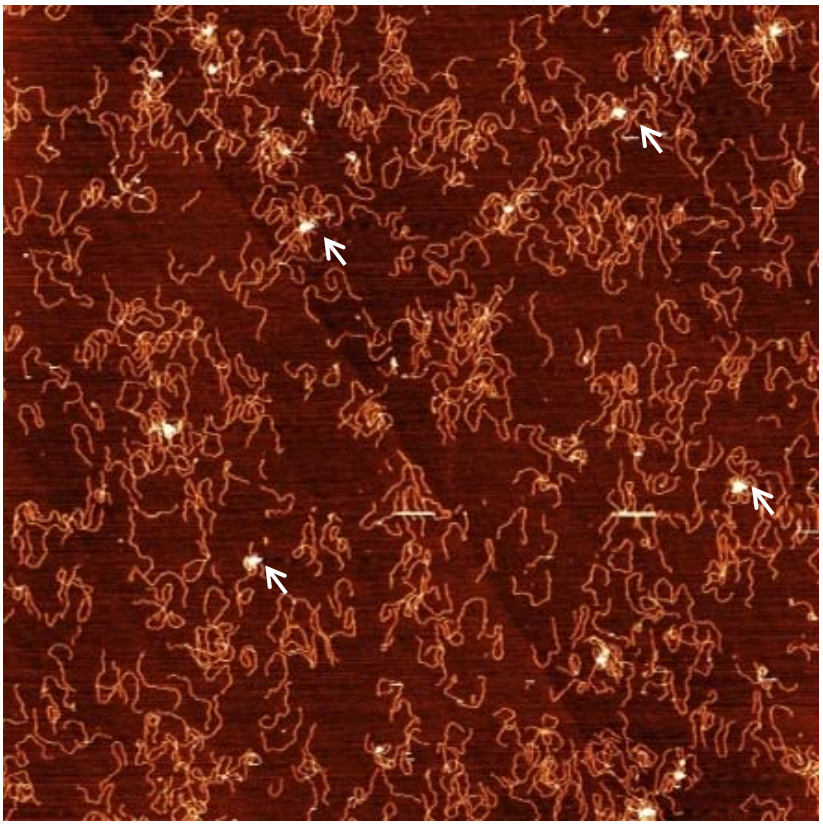


Figure 42. AFM Images of G-wires on Cut- pAcGFP1-N3- Modified Plasmid (1508 bp DNA Fragment). DNA substrate were incubated in a buffer containing 500mM TMPYP4 and deposited at 5 nM concentration (see under "Materials and Methods"). The white line arrows indicated the position of G-wires. The image of G- wires is 5x5 μm^2 at 1.2-nm height scale.

4.4 Conclusion and Discussion

Circular dichroism has been used regularly as a powerful tool to estimate the G-QPXs formation in macromolecules (such as DNA, RNA and proteins) and monitor their conformational [106]. Increased investigation of G-QPX sequences within a promoter region of oncogenes has led to proposing these sequences as regulatory regions and potent targets for anticancer drug delivery [3, 16]. For instance, several studies have reported G-QPX structure in a promoter region of c-MYC gene serving as silencer elements, and its binding to the small stabilizing molecules, TMPYP4, resulting in repression of c-MYC gene [15, 81, 107].

The overall picture emerging from CD study is that G-QPX structures (intramolecular and / or intermolecular) are able to interact with monovalent cations (individual or mixed like nerve firing condition). G-QPX structures are categorized as parallel, antiparallel and mixed – type or hybrid (parallel /antiparallel mixed topology) according to their sequence and loop configurations. Parallel topology exhibits the minimum ellipticity at a range of 238-244 nm and maximum at a range of 260-266 nm, whereas antiparallel structure shows the minimum ellipticity at a range of 280-295 nm and maximum at a range of 260-266 nm. For instance, Kange et.al, has reported a parallel G-QPX structure with a minimum of 240 and maximum of 263 nm ellipticity [37], whereas Bugaut *et.al* reported an antiparallel structure showing the minimum ellipticity at 295 nm and maximum at 260 nm [108]. Structures that have both conformations are called hybrid [105], which have the minimum ellipticity at 240 (typically range of 238-

244) and two maximum ellipticity at 263 and 295nm (typically at a range of 260-266 and 280-295, respectively).

Nucleic acid sequences (ChAT G17 and ChAT G17-2) showed an increase in absorbance at 260 nm and 280nm, upon formation of higher order structures, G-QPX, at lower K^+ concentration (Fig. 32 & 34 , both left image). As noted in previous publications, a stable hybrid structure appears with two positive spectral at 260 and >280 nm in the presence of KCl. Similarly, for ChAT G30, high order stable hybrid structure was observed in KCl solution in concentration – dependent fashion meaning increase stability at lower K^+ concentration. As presented in figure 35 - left the CD spectra showed a slit shoulder at 264 nm and a clear positive pick at 280 nm. ChAT G29 nucleic acids showed a decrease in absorbance at 260 and 280 nm, due to formation of G-QPX structure at lower K^+ concentration compared to higher concentrations (Fig. 37 – left). The folding topology of G-QPX structures could be changed in different salt solutions. Generally, it was demonstrated by CD spectroscopy that all of intamolecular G - QPX structures (ChAT G17-2, ChAT G17, ChAT G30, & ChAT G 29) are more stable in potassium environments than sodium environments and the stability enhanced in dose – dependence fashion except for ChAT G 30. Formation of G-QPX in Li^+ solution showed the highest stability in 2 mM concentrations for both ChAT G17-2 and ChAT G29 (Fig. 33 & 38 - both left image). As it has been suggested in previous publications, the maximum Li^+ dosage in living cells without any toxic effects is 2mM.

ChAT G21-ds nucleic acid sequence have potential G - runs in double stranded DNA within small loop distance, which could form intermolecular G-QPX between two complementary strands. CD results showed that a highly stable structure at low concentrations of KCl solution (10 & 50 mM) in compare to NaCl solutions (Fig. 39). In Li^+ solution the structure showed the best stability at 100 mM concentration. Although the CD spectra in various monavalants cations demonstrated a parallel topology configuration of nucleic acids but it cannot be claimed with certainty that these spectral features form due to folding of intermolecular G-QPX or DNA double strands. Therefore, further analytical studies required to confirm out hypothesis.

Moreover, the folding topology of unique G-QPX structures could be changed in the presence of solution with different salt concentrations. Figure 37 – left showed that ChAT G29 forms a parallel structure in solution containing KCl 50 and 100 mM while it formed a stable hybrid structure in KCl 10mM solution. Kuryavyi. *et.al* conducted NMR study of G-QPX in the promoter of C-Kit2, and demonstrated that G-QPX forms an intramolecular monomeric and dimeric parallel structure in the solution containing KCl 20mM and 100mM, respectively [50]. Therefore, for the future work, NMR analysis is important to get a better understanding of the actual folding topology of our G-QPX structures in various monovalent solutions (individual or mixed).

Firing condition which is mixed ionic environment (K^+:10mM and Na^+:100mM) was further support the CD results of G-QPX stability formation in compare to individual monovalent cations. Figures. 33, 36, 38, and 40 – all right images, showed that the high

order G - QPX stability formation was more and less dominant in presence of K^+ and Na^+ ions, respectively in compare to mixed ionic environment (nerve firing). Therefore, a new hypothesis could be concluded that neural action/ potential ionic condition which is nerve firing and resting may control G-QPX formation / deformation.

Importantly, the above results showed that G-QPX binding affinity in solution containing KCl was higher than other cationic environment (individual or mixed) but the thermal stability of the binding complex needs to be studied. For instance, it has been reported that ligands (TMPYP4) binding affinity to G-QPX in buffer containing NaCl was stronger than in buffer containing KCl at low temperature, but the thermal stability of binding complex were showing opposite behavior [109]. Therefore, to further analysis the predicted ChAT G-QPXs, it is important to investigate ligands binding to these high order secondary structures in solution with and without monovalent cations, studying their binding affinity and their thermal stability of the complex.

In conclusion, it should be considered that the interpretation of CD results requires caution. Although the CD spectra of G-QPXs showed monavalaent cations could stabilized the high order unique structures formed by multiple ChAT G-runs but their stabilities are a little different which needs to be further studied. Moreover, to get a better image of folding topology, the whole spectrum should be taken into account and not relying only on the presence of a single band (eg., 260 or 280). This study shows that CD spectroscopy is a powerful method in study of G-QPXs but it is important to use

complementary methods such as NMR and XRD to get a better view of these novel structures as a novel future anti-cancer therapy approach.

AFM as a single molecule approach allows direct visualization of nucleic acids structures and their distribution has been used to further support CD results. Previous studies have demonstrated the effectiveness of AFM in visualization of single G-QPX units [58-60]. It is shown that G-QPX structures could be categorized into three major groups (depending on their shapes): loop, blob and spur (Fig. 43) [93, 110]. Moreover, these groups can have different characteristics; for instance, G-loops could be symmetric or asymmetric with short or long length. Blobs could be single or multiple and form by aggregation of small number of G-runs with tight space (Fig. 43 A and D). When the loophole is invisible, spur structures could be formed due to bonding of both sides of loop to the mica surface. Recent advances in this area include demonstration of G-QPX in response to ionic conditions, where the G-QPX formation is allowed to be stabilized in K^+ environment. From biology point of view, it is suggested that formation of G-QPXs in DNA mostly works as a transcriptional off-switch and abolish transcription activity [110]. These structures could be stabilized by the presence of; hybridized mRNA/ aptamers on the opposing strand and/or diverse environmental conditions.

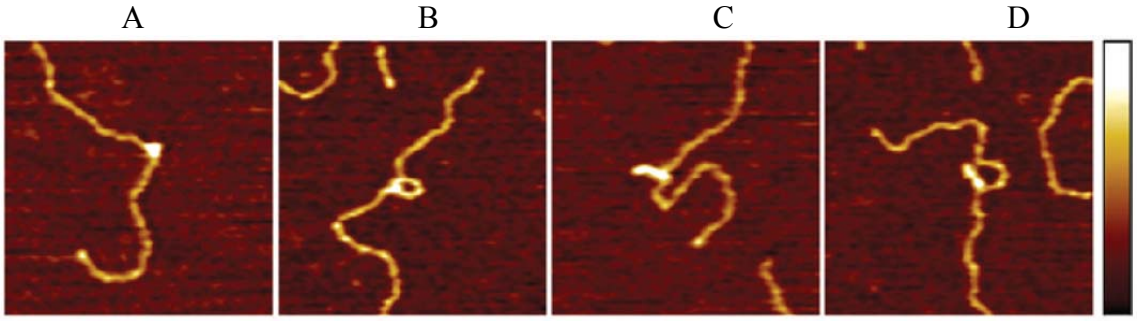


Figure 43. Different Shapes of G-QPX Structures in Zoomed AFM Image. 300x300 nm^2 areas of samples of transcribed plasmid showing (A), a blob (B), a loop (C), a spur and (D), double blob (2X). Shaded bar shows the height scale from 0–1 nm [110].

In the human ChAT promoter model (home lab - designed) G-rich sequences have the potential to form multiple G-QPXs within significantly close space (with distance less than 10 nucleotides). Condensation of consecutive G-QPXs is the main challenge blocking any attempt to separate and visualize an individual G-QPX structure.

Fig.41-A shows the AFM image of 1508 plasmid fragment in presence of K^+. Compared with the image of fragment in solution with no additive salt (only Tris - Hcl, pH=7)(data not shown), number of bright spots are identified in this image which can be caused by interaction of salt with either G-QPXs or ds DNA. Therefore, it can be concluded that the addition of K^+ to the solution, triggers some interaction. The DNA assemblies appear to be stable because even after the solution was allowed to sit at room temperature for a week no significant difference was observed between 24hr incubated sample and 1 week incubated one (data not shown).

Significant increase in amount of bright spots identified in the AFM image shown in Fig. 42 as well as the increased clustering of DNA strands indicates the presence of TMPYP4 acting as stabilizing component for the consecutive G-QPXs in the active promoter region of the human ChAT gene. Interestingly, the G-QPXs shown in Fig. 42 demonstrate formation of star shape structures (clusters of strands) as many G-QPXs aggregate into large, individual clusters that bind together multiple strands. Unfortunately, it is almost impossible to identify individual G-QPX in such structures due to its complexity. This type of structures (Fig. 42) can be categorized as multiple blobs (similar to Fig. 43-D) as suggested by Mela et al [110].

CHAPTER V

CONCLUSIONS AND FUTURE STUDIES

5.1 Conclusions

The identification, formation and function of G-QPX nanoswitch structures upstream of the transcription start side (TSS) in the active promoter region of an enzyme involved in neurotransmitter synthesis, specifically choline acetyltransferase (ChAT), have been investigated for the first time using computational and experimental techniques. A G-rich region with high potential for the formation of G-QPX has been identified as the active promoter region of the human ChAT gene, extending out 1000 bp upstream of the TSS. Within this region, there is a potential for four distinct G-QPXsto form in both DNA complementary strands, identified using several independent G-QPX prediction tools (bioinformatics analysis). Three consecutive intramolecular G-QPX structures in the negative strand and one of intramolecular G-QPX structure in the positive strand were identified (relative to the sense of the encoded ChAT gene). The results also suggest the possibility that nearby G-runs in opposed DNA strands that are too short to form an intrastrand G-QPX may be able to form a stable intermolecular G-QPX (i.e., by interstrand H-bonding within a separated region of a DNA duplex).

The active promoter region of ChAT is highly CG-rich (74%) and thus has high potnetila to be aG-QPX forming region. This relatively short region has demonstrated remarkable effects on the control of gene expression [72]. While exploring literature no TFBS mapping information showing the physical binding of transcription factors to the active promoter region of human ChAT was found. Without the presence of TFBS mapping information, the assessment of importance of potential transcription factor binding sites on formation of novel G-QPXs is difficult task. As a result, performing a TFBS mapping analysis was a necessity initial step. The TFBS analysis presented in this work revealed that the active promoter (-1007 to -925) contains multiple consensus AP-2α and Sp1 binding sites and consensus sites for other TF including multiple sites for GR-alpha, Pax-5, p53 and GC box. Disrupting the TFBS within this region could help us to get a better estimation of the silencing potential of identified G-QPX structures, where dramatic decrease in promoter activity due to formation of G-QPX structures leading to loss of the binding affinity of TF to their recognition sites is expected. Considering the possibility of G-QPX formation along with the observed effect of the QPX stabilizing drug TMPYP4 using our modified GFP plasmid with the ChAT promoter region showed that the likelihood of a silencing effect of these unique structures is significantly high. This result was confirmed in the transfected neuronal cells in which influx of Na^+ ions was achieved via use of the Na^+ ion-channel opening drug aconitine. Aconitine potentiated the action of the transcriptional activator NGF, suggesting that the effect of sodium is contrary to that of TMPYP4, i.e., that an increase in promoter activity may be due to instability of G-QPX structures in a high Na^+ environment, which results in

melting these structures, enabling dsDNA formation required for the binding transcription factors to their recognition sites for initiation of transcription. The obtained results were confirmed in several independent sets of experiments, using GFP reporter gene measurement by plate reader and by flow cytometry. Also, further verification was reported by visualizing the fluorescence intensity using EVOS-microscopy.

As final validation of the results obtained using the synthetic ChAT promoter-reporter gene construct, quantitative RT-PCR was conducted to examine the regulation of the human ChAT gene expression in a neuroblastoma cell line by the same factors used in the reporter gene experimentss. The outcome indicated that the presence of TMPYP4 lead to a significant knockdown in ChAT mRNA expression (87%), suggesting that G-QPX stabilization inhibits promoter activity as expected. Results also showed that aconitine-mediated influx of Na^+ ions has an opposite effect, presumably by inhibiting the formation of stable G-QPX structures, resulting in an increase in ChAT mRNA expression.

Physical studies including CD spectroscopy and AFM imaging were conducted to analyze the folding topology and stability of identified consecutive G-QPX structures. CD results showed G-QPX structures with hybrid folding topology and high stability in K^+ solutions. The hybrid configuration of ChAT G17-2 and ChAT G17 were in line with their arrangements of G-runs, suggesting the presence of a zero loop for both unique structures. The three intramolecular G-QPX structures (ChAT G17-2, ChAT G17, and ChAT G29) demonstrated enhanced stability inas potassium concentration was increased; G-QPX ChAT G 30 was an exception. Also, a proposed unique intermolecular G-QPX

that could form between the separated strands of the DNA duplex showed a highly stable parallel G-QPX formation in high K^+ solution.

In Li^+ solution (2mM), ChAT G17-2 and ChAT G29 showed high stability, which opens new possibilities for the pharmacological mechanism of action of lithium salts as effective therapeutics for CNS diseases such as biopolar disorders. Lithium is known to somehow stabilize neurotransmitter levels and functions, so the possibility that this is mediated in part via the action of Li^+ ions on G-QPX structures in the promoters of genes for neutrotransmitter synthesis merits further investigation.

More importantly, the CD spectra of G-QPX motifs support the proposed novel hypothesis, regulating G-QPX formation and deformation through the influx and efflux of Na^+ and K^+ ions during the neuronal action–potential mechanism. Specifically, in the light of the entire body of results presented here, it can explain the gene silencing during nerve resting (stability of G-QPX due to high K^+ inside the cell) and gene expression during prolonged nerve firing (melting the G-QPX through sustained influx of Na^+ ions allowing Na^+ accumulation in the nucleus).

AFM imaging demonstrated star-shaped structures (involving clusters of DNA strands) due to stabilization of G-QPX structures with TMPYP4, resulting in their aggregation into large cluster of binding multiple strands.

As mentioned before, the AFM images suggest that upon incubation with TMPYP4, a greater number of these G-rich sequences have converted to G-QPX structures. Unfortunately, it is not possible to determine how many molecules of

TMPYP4 are bound within the loops. For instance, it has been reported that a parallel/antiparallel G-QPX hybrid structure binds externally with two TMPYP4 molecules, interacting with two external loops [111]. Also, the binding of small molecules to G-QPXs may not change their stability and conformation but it may change the molecular recognition and binding of TFBS upstream of TSS which subsequently alter the gene transcription activity [112].

Overall, this body of work contributes to identification of G-rich sequence in the active promoter region of the human ChAT gene with great potential to form multiple G-QPXs, stabilization of which results in down-regulation of ChAT gene transcription. The selected runs of guanine are close to binding sites for Sp1 and AP2-α transcription factors; their conversion to high order non-B form secondary structures (G-QPXs) would result in preventing access of Sp1 and AP2-α to their binding sites. The binding of Sp1 and AP2-α is believed to up-regulate human ChAT gene transcription. Binding of TMPYP4 (meso-tetra (N-methyl-4-pyridyl)porphine) as a stabilizing compound to these G-QPXs result in down regulation of the human ChAT gene.

The G-rich region of the active promoter of the human ChAT gene represents a new molecular target in development of small therapeutic molecules. It will be important to obtain more precise information about the actual secondary and 3D structures of naturally occurring G-QPXs and their structural complexes with drugs. An improved understanding of the structural details of these G-QPXs, and how they are influenced by

changes in the concentrations of alkali metal ions, will be required for this potential to be realized.

5.2 Future Studies

This study improves our current understanding of the mechanisms involved in the regulation of neurotransmitter synthesis during neural signaling. In a larger picture, it contributes to increasing our knowledge of human brain functionality and provides a blueprint for collecting data and computational model design of nerve system for building artificial structures, i.e., synthetic biology approaches. This could include the development of improved probes and devices (higher order structures) for the regulation of genes (neurogenesis, proto-oncogenes, and oncogenes) as nanostructured ON/OFF switches, which could result in activating or inhibiting transcriptional machinery, or other systems of interest. This work highlights the need to perform both experimental and computational modeling studies to confirm the obtained results and improve understanding of the G-QPX structures, stabilities, formation and importantly their biological function as novel small targets for small drug-like molecules. This can be achieved by conducting NMR analysis which is critical to understanding the folding topology of G-QPXs in different ionic solutions (individual or mixed ionic solutions). Additionally, although we have inferred effects upon the binding affinity of G-QPX motifs as a way to explain the observed effects of TMPYP4 and altered Na+ and K^+ environments on our model system (ChAT gene expression), a more precise understanding of the thermodynamic stability of these structures and their interactions

with small molecule ligands and/or metal ions binding between the tetrad layers will require further in depth investigation by more direct physicochemical methods.

REFERENCES

1. Mohanty, A.C.B.a.J., *Targeting G-quadruplex structures with extrinsic fluorogenic dyes: promising fluorescence sensors.* The Royal Society of Chemistry- Chem Commun, 2015. **51**: p. 7581-7597.
2. Giulia Biffi, D.T., John McCafferty & Shankar Balasubramanian, *Quantitative visualization of DNA G-quadruplex structures in human cells.* Nature Chemistry, 2013. **5**: p. 182-186.
3. Balasubramanian, S., L.H. Hurley, and S. Neidle, *Targeting G-quadruplexes in gene promoters: a novel anticancer strategy?* Nat Rev Drug Discov, 2011. **10**(4): p. 261-75.
4. Morgan, R.K., et al., *Identification and characterization of a new G-quadruplex forming region within the kRAS promoter as a transcriptional regulator.* Biochim Biophys Acta, 2016. **1859**(2): p. 235-45.
5. Mohanty, A.C.B.a.J., *Targeting G-quadruplex structures with extrinsic fluorogenic dyes: promising fluorescence sensors.* Chem. Commun., 2015. **51**: p. 7581-7591.
6. Xiao, S., et al., *Bioinformatic analysis reveals an evolutional selection for DNA:RNA hybrid G-quadruplex structures as putative transcription regulatory elements in warm-blooded animals.* Nucleic Acids Res, 2013. **41**(22): p. 10379-90.
7. Waller, Z.A., et al., *Identification and characterisation of a G-quadruplex forming sequence in the promoter region of nuclear factor (erythroid-derived 2)-like 2 (Nrf2).* Biochem Biophys Res Commun, 2014. **447**(1): p. 128-32.
8. JL., H., *Structure, location and interactions of G-quadruplexes.* FEBS J. , 2010. **277**(17): p. 3452-3458.
9. Brooks, T.A., S. Kendrick, and L. Hurley, *Making sense of G-quadruplex and i-motif functions in oncogene promoters.* FEBS J, 2010. **277**(17): p. 3459-69.
10. Lipps HJ, R.D., *G-quadruplex structures: in vivo evidence and function.* Trends Cell Biol., 2009. **19**(8): p. 414-422.
11. Bugaut, A. and S. Balasubramanian, *5'-UTR RNA G-quadruplexes: translation regulation and targeting.* Nucleic Acids Res, 2012. **40**(11): p. 4727-41.
12. Hewitt, G., et al., *Telomeres are favoured targets of a persistent DNA damage response in ageing and stress-induced senescence.* Nat Commun, 2012. **3**: p. 708.
13. Beaudoin, J.D. and J.P. Perreault, *Exploring mRNA 3'-UTR G-quadruplexes: evidence of roles in both alternative polyadenylation and mRNA shortening.* Nucleic Acids Res, 2013. **41**(11): p. 5898-911.
14. Rhodes, D. and H.J. Lipps, *G-quadruplexes and their regulatory roles in biology.* Nucleic Acids Res, 2015. **43**(18): p. 8627-37.
15. Siddiqui-Jain, A., et al., *Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription.* Proc Natl Acad Sci U S A, 2002. **99**(18): p. 11593-8.
16. Huppert, J.L. and S. Balasubramanian, *G-quadruplexes in promoters throughout the human genome.* Nucleic Acids Res, 2007. **35**(2): p. 406-13.

17. Hidenobu Yaku, T.F., Takashi Murashima,Daisuke Miyoshi and Naoki Sugimotobc *Phthalocyanines: a new class of G-quadruplex-ligands with many potential applications.* Chem. Commun, 2012. **48**: p. 6203-6216.
18. He, Q., et al., *G-quadruplex-mediated regulation of telomere binding protein POT1 gene expression.* Biochim Biophys Acta, 2014. **1840**(7): p. 2222-33.
19. Yan, J., et al., *An intramolecular G-quadruplex structure formed in the human MET promoter region and its biological relevance.* Mol Carcinog, 2016. **55**(5): p. 897-909.
20. Phatak, P., et al., *Telomere uncapping by the G-quadruplex ligand RHPS4 inhibits clonogenic tumour cell growth in vitro and in vivo consistent with a cancer stem cell targeting mechanism.* Br J Cancer, 2007. **96**(8): p. 1223-33.
21. Gary N. Parkinson, R.G., and Stephen Neidle, *Structural Basis for Binding of Porphyrin to Human Telomeres.* Biochemistry 2007. **46**: p. 2390-2397.
22. Parkinson, G.N., F. Cuenca, and S. Neidle, *Topology conservation and loop flexibility in quadruplex-drug recognition: crystal structures of inter- and intramolecular telomeric DNA quadruplex-drug complexes.* J Mol Biol, 2008. **381**(5): p. 1145-56.
23. Nancy H. Campbell, G.N.P., Anthony P. Reszka, and Stephen Neidle, *Structural Basis of DNA Quadruplex Recognition by an Acridine Drug.* J. AM. CHEM. SOC, 2008. **130**: p. 6722–6724.
24. Casagrande, V., et al., *N-cyclic bay-substituted perylene G-quadruplex ligands have selective antiproliferative effects on cancer cells and induce telomere damage.* J Med Chem, 2011. **54**(5): p. 1140-56.
25. Xiong, Y.-X., et al., *Blocking the binding of WT1 to bcl-2 promoter by G-quadruplex ligand SYUIQ-FM05.* Biochemistry and Biophysics Reports, 2016. **5**: p. 346-352.
26. Collie, G.W. and G.N. Parkinson, *The application of DNA and RNA G-quadruplexes to therapeutic medicines.* Chem Soc Rev, 2011. **40**(12): p. 5867-92.
27. Perrone, R., et al., *Formation of a unique cluster of G-quadruplex structures in the HIV-1 Nef coding region: implications for antiviral activity.* PLoS One, 2013. **8**(8): p. e73121.
28. Risitano, A. and K.R. Fox, *Inosine substitutions demonstrate that intramolecular DNA quadruplexes adopt different conformations in the presence of sodium and potassium.* Bioorg Med Chem Lett, 2005. **15**(8): p. 2047-50.
29. Webba da Silva, M., *Geometric formalism for DNA quadruplex folding.* Chemistry, 2007. **13**(35): p. 9738-45.
30. Shalaby, T., et al., *G-quadruplexes as potential therapeutic targets for embryonal tumors.* Molecules, 2013. **18**(10): p. 12500-37.
31. Patel, D.J., A.T. Phan, and V. Kuryavyi, *Human telomere, oncogenic promoter and 5'-UTR G-quadruplexes: diverse higher order DNA and RNA targets for cancer therapeutics.* Nucleic Acids Res, 2007. **35**(22): p. 7429-55.
32. TERRI L. DAVIS, A.B.F., AND ALAN J. KINNIBURGH, *Ribonucleoprotein and protein factors bind to an H-DNA-forming c-myc DNA element: Possible regulators of the c-myc gene.* Biochemistry, 1989. **86**: p. 9682-9686.
33. Ou, T.M., et al., *Inhibition of cell proliferation by quindoline derivative (SYUIQ-05) through its preferential interaction with c-myc promoter G-quadruplex.* J Med Chem, 2011. **54**(16): p. 5671-9.

34. Kostadinov, R., et al., *GRSDB: a database of quadruplex forming G-rich sequences in alternatively processed mammalian pre-mRNA sequences.* Nucleic Acids Res, 2006. **34**(Database issue): p. D119-24.
35. Margarita I. Zarudnaya, I.M.K., Andriy L. Potyahaylo and and D.M. Hovorun, *Downstream elements of mammalian pre-mRNA polyadenylation signals: primary, secondary and higher-order structures.* Nucleic Acids Research, 2003. **31**(5): p. 1375-1386.
36. Kerry J. Woodford, R.M.H., and Karen Usdin, *A Novel K-dependent and Synthesis Arrest Site in a Commonly Occurring Sequence Motif in Eukaryotes.* THE JOURNOFA BLI OLOGICCAHLE MISTRY, 1994. **269**(43): p. 27029-27035.
37. Kang, H.J. and H.J. Park, *In silico identification of novel ligands for G-quadruplex in the c-MYC promoter.* J Comput Aided Mol Des, 2015. **29**(4): p. 339-48.
38. Brown, R.V., et al., *Demonstration that drug-targeted down-regulation of MYC in non-Hodgkins lymphoma is directly mediated through the promoter G-quadruplex.* J Biol Chem, 2011. **286**(47): p. 41018-27.
39. Sun, D., et al., *The proximal promoter region of the human vascular endothelial growth factor gene has a G-quadruplex structure that can be targeted by G-quadruplex-interactive agents.* Mol Cancer Ther, 2008. **7**(4): p. 880-9.
40. Agrawal, P., et al., *Solution structure of the major G-quadruplex formed in the human VEGF promoter in K+: insights into loop interactions of the parallel G-quadruplexes.* Nucleic Acids Res, 2013. **41**(22): p. 10584-92.
41. Richard De Armond, S.W., Daekyu Sun, Laurence H. Hurley, and Scot W. Ebbinghaus, *Evidence for the Presence of a Guanine Quadruplex Forming Region within a Polypurine Tract of the Hypoxia Inducible Factor 1R Promoter.* Biochemistry, 2005. **44**: p. 16341-16350.
42. Guo, K., et al., *Formation of pseudosymmetrical G-quadruplex and i-motif structures in the proximal promoter region of the RET oncogene.* J Am Chem Soc, 2007. **129**(33): p. 10220-8.
43. Cogoi, S. and L.E. Xodo, *G-quadruplex formation within the promoter of the KRAS proto-oncogene and its effect on transcription.* Nucleic Acids Res, 2006. **34**(9): p. 2536-49.
44. Cogoi, S., A.E. Shchekotikhin, and L.E. Xodo, *HRAS is silenced by two neighboring G-quadruplexes and activated by MAZ, a zinc-finger transcription factor with DNA unfolding property.* Nucleic Acids Res, 2014. **42**(13): p. 8379-88.
45. Kendrick, S., et al., *The dynamic character of the BCL2 promoter i-motif provides a mechanism for modulation of gene expression by compounds that bind selectively to the alternative DNA hairpin structure.* J Am Chem Soc, 2014. **136**(11): p. 4161-71.
46. Agrawal, P., et al., *The major G-quadruplex formed in the human BCL-2 proximal promoter adopts a parallel structure with a 13-nt loop in K+ solution.* J Am Chem Soc, 2014. **136**(5): p. 1750-3.
47. Anh Tuaˆn Phan, V.K., Sarah Burge, Stephen Neidle, and and D.J. Patel, *Structure of an Unprecedented G-Quadruplex Scaffold in the Human c-kit Promoter.* J. AM. CHEM. SOC., 2007. **129**: p. 4386-4392.
48. Pravin S. Shirude, B.O., Liming Ying, Taekjip Ha, and Shankar Balasubramanian., *Single-Molecule Conformational Analysis of G-Quadruplex Formation in the Promoter DNA Duplex of the Proto-Oncogene C-Kit.* J. AM. CHEM. SOC., 2007. **129**: p. 7484-7485.

49. Todd, A.K., et al., *Sequence occurrence and structural uniqueness of a G-quadruplex in the human c-kit promoter.* Nucleic Acids Res, 2007. **35**(17): p. 5799-808.
50. Kuryavyi, V., A.T. Phan, and D.J. Patel, *Solution structures of all parallel-stranded monomeric and dimeric G-quadruplex scaffolds of the human c-kit2 promoter.* Nucleic Acids Res, 2010. **38**(19): p. 6757-73.
51. Qin, Y., et al., *Characterization of the G-quadruplexes in the duplex nuclease hypersensitive element of the PDGF-A promoter and modulation of PDGF-A promoter activity by TMPyP4.* Nucleic Acids Res, 2007. **35**(22): p. 7698-713.
52. Palumbo, S.L., et al., *A novel G-quadruplex-forming GGA repeat region in the c-myb promoter is a critical regulator of promoter activity.* Nucleic Acids Res, 2008. **36**(6): p. 1755-69.
53. SunMi L. Palumbo, S.W.E., and Laurence H. Hurley, *Formation of a Unique End-to-End Stacked Pair of G-Quadruplexes in the hTERT Core Promoter with Implications for Inhibition of Telomerase by G-Quadruplex-Interactive Ligands.* J. AM. CHEM. SOC., 2009. **131**: p. 10878–10891.
54. Kah Wai Lim, L.L., Doris Jia En Yue, Joefina Kim Cheow Lim, Jocelyn Mei Wen Lim, and Anh Tuaˆn Phan, *Coexistence of Two Distinct G-Quadruplex Conformations in the hTERT Promoter.* J. AM. CHEM. SOC., 2010. **132**: p. 12331–12342.
55. Xu, Y. and H. Sugiyama, *Formation of the G-quadruplex and i-motif structures in retinoblastoma susceptibility genes (Rb).* Nucleic Acids Res, 2006. **34**(3): p. 949-54.
56. Amrane, S., et al., *Topology of a DNA G-quadruplex structure formed in the HIV-1 promoter: a potential target for anti-HIV drug development.* J Am Chem Soc, 2014. **136**(14): p. 5249-52.
57. Nancy H. Campbell, G.N.P., Anthony P. Reszka, and Stephen Neidle, *Structural Basis of DNA Quadruplex Recognition by an Acridine Drug.* J. AM. CHEM. SOC., 2008. **130**: p. 6722–6724.
58. Todd, A.K., M. Johnston, and S. Neidle, *Highly prevalent putative quadruplex sequence motifs in human DNA.* Nucleic Acids Res, 2005. **33**(9): p. 2901-7.
59. Huppert, J.L. and S. Balasubramanian, *Prevalence of quadruplexes in the human genome.* Nucleic Acids Res, 2005. **33**(9): p. 2908-16.
60. Bagga, L.D.A.a.P., *Computational Methods for Predicting Intramolecular G -quadruplexes in Nucleotide Sequences.* IEEE Computational Systems Bioinformatics Conference, 2004.
61. Huppert, J.L., *Hunting G-quadruplexes.* Biochimie, 2008. **90**(8): p. 1140-8.
62. Todd, A.K., *Bioinformatics approaches to quadruplex sequence location.* Methods, 2007. **43**(4): p. 246-51.
63. Kikin, O., L. D'Antonio, and P.S. Bagga, *QGRS Mapper: a web-based server for predicting G-quadruplexes in nucleotide sequences.* Nucleic Acids Res, 2006. **34**(Web Server issue): p. W676-82.
64. Scaria, V., et al., *Quadfinder: server for identification and analysis of quadruplex-forming motifs in nucleotide sequences.* Nucleic Acids Res, 2006. **34**(Web Server issue): p. W683-5.
65. Yadav, V.K., et al., *QuadBase: genome-wide database of G4 DNA--occurrence and conservation in human, chimpanzee, mouse and rat promoters and 146 microbes.* Nucleic Acids Res, 2008. **36**(Database issue): p. D381-5.

66. Stegle, O., et al., *Predicting and understanding the stability of G-quadruplexes.* Bioinformatics, 2009. **25**(12): p. i374-82.
67. Eddy, J. and N. Maizels, *Gene function correlates with potential for G4 DNA formation in the human genome.* Nucleic Acids Res, 2006. **34**(14): p. 3887-96.
68. Eddy, J. and N. Maizels, *Conserved elements with potential to form polymorphic G-quadruplex structures in the first intron of human genes.* Nucleic Acids Res, 2008. **36**(4): p. 1321-33.
69. Scott Frees1, C.M., Matt Crum and Paramjeet S Bagga, *QGRS-Conserve: a computational method for discovering evolutionarily conserved G-quadruplex motifs.* Human Genomics, 2014. **8**.
70. Wong, H.M., et al., *A toolbox for predicting g-quadruplex formation and stability.* J Nucleic Acids, 2010. **2010**.
71. Lam, E.Y., et al., *G-quadruplex structures are stable and detectable in human genomic DNA.* Nat Commun, 2013. **4**: p. 1796.
72. Louis B. Hersh, C.F.K., Craig Sampson, Gabriele Mues, Yi-Ping Li, Ann Fisher, Dana Hilt, and E. Edward Baetge, *Comparison of the Promoter Region of the Human and Porcine Choline Acetyltransferase Genes: Localization of an Important Enhancer Region.* Journal of Nanochemistry., 1992. **61**(1): p. 306-314.
73. Brown, M.S.S.a.D.D., *The Transcriptional Regulation of Xenopus 5s RNA Genes in Chromatin: The Roles of Active Stable Transcription Complexes and Histone H 1.* cell, 1984. **37**: p. 903-913.
74. Bejanin. S, H.E., Berrard. S, Edwards. JB, Loeffler. JP, Mallet. J., *Promoter elements of the rat choline acetyltransferase gene allowing nerve growth factor inducibility in transfected primary cultured cells.* J Neurochem, 1992. **58**(4): p. 1580-1583.
75. Persson, C.F.I.a., *Localization of Sequences Determining Cell Type Specificity and NGF Responsiveness in the Promoter Region of the Rat Choline Acetyltransferase Gene.* European journal of neuroscience, 1991. **3**(12): p. 1309 –1315.
76. Tjian, T.W.a.R., *Analysis of the DNA-binding and activation properties of the human transcription factor AP-2.* GENES & DEVELOPMENT, 1991. **5**: p. 670-682.
77. Jens Gille, R.A.S.a.S.W.C., *Transforming growth factor-α-induced transcriptional activation of the vascular permeability factor (VPF/VEGF) gene requires AP-2 dependent DNA binding and transactivation.* The EMBO Journal, 1997. **16**(4): p. 750–759.
78. Jordano J, P.M., *Initial characterization of a potential transcriptional enhancer for the human c-K-ras gene.* Oncogene, 1988. **2**(4): p. 359-366.
79. Yamamoto F, P.M., *Characterization of the human c-K-ras gene promoter.* Oncogene Res, 1988. **3**(2): p. 125-130.
80. Onel, B., et al., *A New G-Quadruplex with Hairpin Loop Immediately Upstream of the Human BCL2 P1 Promoter Modulates Transcription.* J Am Chem Soc, 2016. **138**(8): p. 2563-70.
81. Laurence H. Hurley, D.D.V.H., Adam Siddiqui-Jain, Danzhou Yang., *Drug Targeting of the c-MYC Promoter to Repress Gene Expression via a G-Quadruplex Silencer Element.* Seminars in Oncology, 2006. **33**(4): p. 498–512.
82. Zhi-Qiang Hu, H.-Y.H., and Chuan-Feng Chen, *Phenanthroline Dicarboxamide-Based Helical Foldamers: Stable Helical Structures in Methanol.* J. Org. Chem, 2006. **71**(3): p. 1131-1138.

83. Kumar, P., et al., *Zinc-finger transcription factors are associated with guanine quadruplex motifs in human, chimpanzee, mouse and rat promoters genome-wide.* Nucleic Acids Res, 2011. **39**(18): p. 8005-16.
84. Vorlickova, M., et al., *Circular dichroism and guanine quadruplexes.* Methods, 2012. **57**(1): p. 64-75.
85. Henderson, A., et al., *Detection of G-quadruplex DNA in mammalian cells.* Nucleic Acids Res, 2014. **42**(2): p. 860-9.
86. Bochman, M.L., K. Paeschke, and V.A. Zakian, *DNA secondary structures: stability and function of G-quadruplex structures.* Nat Rev Genet, 2012. **13**(11): p. 770-80.
87. Kypr, J., et al., *Circular dichroism and conformational polymorphism of DNA.* Nucleic Acids Res, 2009. **37**(6): p. 1713-25.
88. Wells, R.D., *Non-B DNA conformations, mutagenesis and disease.* Trends Biochem Sci, 2007. **32**(6): p. 271-8.
89. Vorlekova, M., *Cooperative Changes in the Chiroptical Properties of DNA Induced by Methanol.* Biopolymers, 1984. **23**.
90. Balasubramanian, S. and S. Neidle, *G-quadruplex nucleic acids as therapeutic targets.* Curr Opin Chem Biol, 2009. **13**(3): p. 345-53.
91. Neal Crampton, S.R., David TF Dryden, Desirazu N Rao,J Michael Edwardson and Robert M Henderson., *DNA looping and translocation provide an optimal cleavage mechanism for the type III restriction enzymes.* The EMBO Journal, 2007. **26**: p. 3815–3825.
92. Crampton, N., et al., *Fast-scan atomic force microscopy reveals that the type III restriction enzyme EcoP15I is capable of DNA translocation and looping.* Proc Natl Acad Sci U S A, 2007. **104**(31): p. 12755-60.
93. Neaves, K.J., et al., *Direct visualization of G-quadruplexes in DNA using atomic force microscopy.* Nucleic Acids Res, 2009. **37**(18): p. 6269-75.
94. Y. L. Lyubchenko, A.A.G., L. S. Shlyakhtenk., *Atomic force microscopy of DNA and protein-DNA complexes using functionalized mica substrates.* Methods in molecular biology 2001. **148**: p. 569-578.
95. Carlos Bustamante, C.R.a.D.J.K., *Scanning force microscopy under aqueous solutions.* Biophysical methods, 1997: p. 709-715.
96. Shlyakhtenko, L.S., et al., *Silatrane-based surface chemistry for immobilization of DNA, protein-DNA complexes and other biological materials.* Ultramicroscopy, 2003. **97**(1-4): p. 279-287.
97. Medalia O, E.J., Guckenberger R, Sperling J., *AFM imaging in solution of protein-DNA complexes formed on DNA anchored to a gold surface.* Ultramicroscopy, 2001. **90**(2): p. 103-112.
98. Song Y, L.Z., Liu Z, Wei G, Wang L, Sun L., *Immobilization of DNA on 11-mercaptoundecanoic acid-modified gold (111) surface for atomic force microscopy imaging.* Microsc Res Tech, 2005. **68**(2): p. 59-64.
99. Wang, T.W., et al., *Application of highly sensitive, modified glass substrate-based immuno-PCR on the early detection of nasopharyngeal carcinoma.* Biomaterials, 2008. **29**(33): p. 4447-54.
100. Liu, Z., et al., *Observation of the mica surface by atomic force microscopy.* Micron, 2005. **36**(6): p. 525-31.

101. Vesenka J, G.M., Tang CL, Keller D, Delaine E, Bustamante C., *Substrate preparation for reliable imaging of DNA molecules with the scanning force microscope.* Ultramicroscopy, 1992. **42**(44): p. 1243-1249.
102. Bustamante. C, V.J., Tang. CL, Rees. W, Guthold. M, Keller. R., *Circular DNA molecules imaged in air by scanning force microscopy.* Biochemistry., 1992. **31**(1): p. 22-26.
103. Lyubchenko, Y.L., L.S. Shlyakhtenko, and T. Ando, *Imaging of nucleic acids with atomic force microscopy.* Methods, 2011. **54**(2): p. 274-83.
104. Muller, S., et al., *Small-molecule-mediated G-quadruplex isolation from human cells.* Nat Chem, 2010. **2**(12): p. 1095-8.
105. Waller, Z.A., et al., *A small molecule that disrupts G-quadruplex DNA structure and enhances gene expression.* J Am Chem Soc, 2009. **131**(35): p. 12628-33.
106. Greenfield, N.J., *Using circular dichroism spectra to estimate protein secondary structure.* Nat Protoc, 2006. **1**(6): p. 2876-90.
107. Cory L. Grand, H.H., Rube´n M. Munoz, Steve Weitman, Daniel D. Von Hoff, Laurence H. Hurley, and David J. Bearss, *The Cationic Porphyrin TMPyP4 Down-Regulates c-MYC and Human Telomerase Reverse Transcriptase Expression and Inhibits Tumor Growth in Vivo.* Molecular Cancer Therapeutics, 2002: p. 565-573.
108. Bugaut, A. and S. Balasubramanian, *A sequence-independent study of the influence of short loop lengths on the stability and topology of intramolecular DNA G-quadruplexes.* Biochemistry, 2008. **47**(2): p. 689-97.
109. Dong-Fang Shi, R.T.W., Daekyu Sun, and Laurence H. Hurley, *Quadruplex-Interactive Agents as Telomerase Inhibitors: Synthesis of Porphyrins and Structure-Activity Relationship for the Inhibition of Telomerase.* J. Med. Chem., 2001. **4**: p. 4509-4523.
110. Mela, I., et al., *Demonstration of ligand decoration, and ligand-induced perturbation, of G-quadruplexes in a plasmid using atomic force microscopy.* Biochemistry, 2012. **51**(2): p. 578-85.
111. Jeyaprakashnarayanan Seenisamy, E.M.R., Tiffanie J. Powell, Denise Tye, Vijay Gokhale, Chandana Sharma Joshi, Adam Siddiqui-Jain, and Laurence H. Hurley, *The Dynamic Character of the G-Quadruplex Element in the c-MYC Promoter and Modification by TMPyP4.* J. AM. CHEM. SOC., 2004. **126**: p. 8702-8709.
112. LH., H., *Secondary DNA structures as molecular targets for cancer therapeutics.* Biochem Soc Trans., 2001. **29**: p. 692-696.

APPENDIX A

CELL CULTURE AND CYTOTOXICITY OF TRANSFAST TRANSFECTION

IMR-32 Cell culture and passaging

Human neuroblastoma, IMR-32 cells were purchased from American Type Cell Culture ATCC (Catalog No. CCL-127). IMR-32 cells are fibroblast or neuroblast type of cells which obtained from human brain tissue or derived from human metastatic site: abdominal mass (13 monts male). The cultural property of IMR-32 cells as adherent cell is that they are suitable for transfection, so they should be useful for expression measurement with exogenous genes. However, there has been in consist reports about their transfection efficiency, some literature studies reported that these cells are not easy to transfect and exhibit a high transfection efficiency while others suggesting that with using proper transfection reagent they can be easily transfect.

IMR-32 cells were grown in Minimum Essential Medium (MEM), 1X (low glucose, and 2.2 g/L sodium bicarbonate with L-glutamine & phenol red, no sodium pyruvate) (catalogue no. 11095080, Thermofisher), the media was mixed with 10% Fetal Bovine Serum to make the complete medium. Cell culture was maintained without the use of any antibiotics, incubation at 37 °C and 5% CO_2 in a humidified environment.

For passaging, Cells routinely were passaged when they reach 75% confluent, sub - cultured every 3-4 days at a seeding density of 2 x 10^4 in 25 cm^2 vent cap T- flasks (Corning) in a total volume of 5 mL media per flask. Cells were split at a ratio of 1:50 due to their fast rate growth. Incubation condition for cells was 37 °C and 5% CO_2 in a humidified environment. Cell counting was performed using TC10 automated cell counter (Bio-Rad) and hemocytometer, for both techniques cell dilution was 1:1 in 0. 04% Trypan Blue.

The medium was removed and cells were washed with 2-4 ml of pre-warmed 37 °C complete medium (which was placed in the incubator for at least 15 minutes) to remove the residual old medium. For passaging, Cells were then incubated with pre-warmed Trypsin EDTA solution (Catalog no. NC0043665)at 37 °C for 5-7 minutes. After detaching almost all of the cells, 2 ml of pre-warmed complete medium was added to stop trypsinization, pipetting was conducted to detach all cells. Next, cell suspension was transferred to 15 ml falcon tube (Falcon™ 15mL conical centrifuge tube) and centrifuged for 7 minutes. The supernatant discarded and the pellet was suspended with 1 ml of pre-warmed complete medium. Cells were diluted, counted and seeded as described previously.

Different cell batches of the IMR-32 cell line were frozen and kept in liquid nitrogen. For cryopreservation; cells were diluted with a cold mixture of 90% compete medium (MEM

+ FBS) and 10% of DMSO to get a concentration of 1.0 x 10^6 cells/. Cells were kept in freezing container and transferred to refrigerator for 2-3 hours followed by -80 °C incubation overnight and eventually stored in liquid nitrogen.

Recovery of frozen cells was conducted through an immediate placing them in 37 °C water bath to rapidly thaw the cells (2 minutes). Followed by spraying with 70% ethanol and wiping with Kim wipe paper. Next, the vial content was transferred to 15 ml falcon tube (Falcon™ 15mL conical centrifuge tube) containing 9ml of 37 °C complete medium (which was placed in the incubator for at least 15 minutes) and centrifuged at 2000 g (125 xg) for 7 minutes. The supernatant was discarded and the pellet was suspended in 1 ml of pre-warmed 37 °C complete medium (Pipetting was done to mix the cells well with fresh media and separation). Next, cells were diluted and transferred to a 25 cm^2 flask, and the total volume is brought to 5 ml by adding pre-warmed 37 °C complete medium. Finally, cells were observed under the microscope, followed by incubation at 37 °C and 5% CO_2.

Cytotoxicity Assays (MTT and MTS)

MTT and MTS are tetrazoles and are used for proliferation assay; the reduction of tetrazolium salt indicates cells metabolic activity. MTT (3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide, is a common proliferation assay and is used for measuring the reactivity of oxi-reductase enzymes (water soluble yellow tetrazole). MTS a novel tetrazolium compound [3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium] in the presence of electron coupling reagent, phenazine methosulfate (PMS) is the same as MTT assay but it is a one-step assay (using one solution instead of two). It does not need the solubilization step which is required for MTT, and can produce the soluble formazan through the reduction with specific intercellular enzymes. They can be detected at absorbance arrange of 490-570 nm.

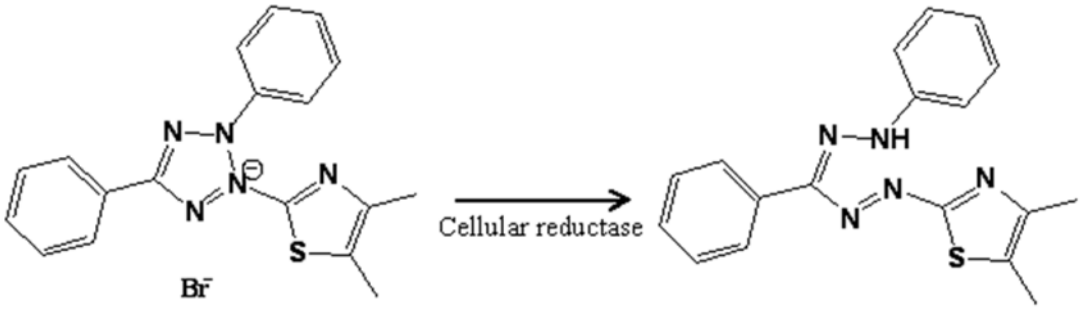


Figure 44. MTT Contains a Comment Tetrazole [the diphenyl-tertrazolium bromide (Water Soluble Yellow Tetrazole)] which upon the Reduction Mechanism, it turns into Insoluble Purple Formazan.

APPENDIX B

PROTOCOLS OF PLASMID TRANSFECTION AND qRT-PCR

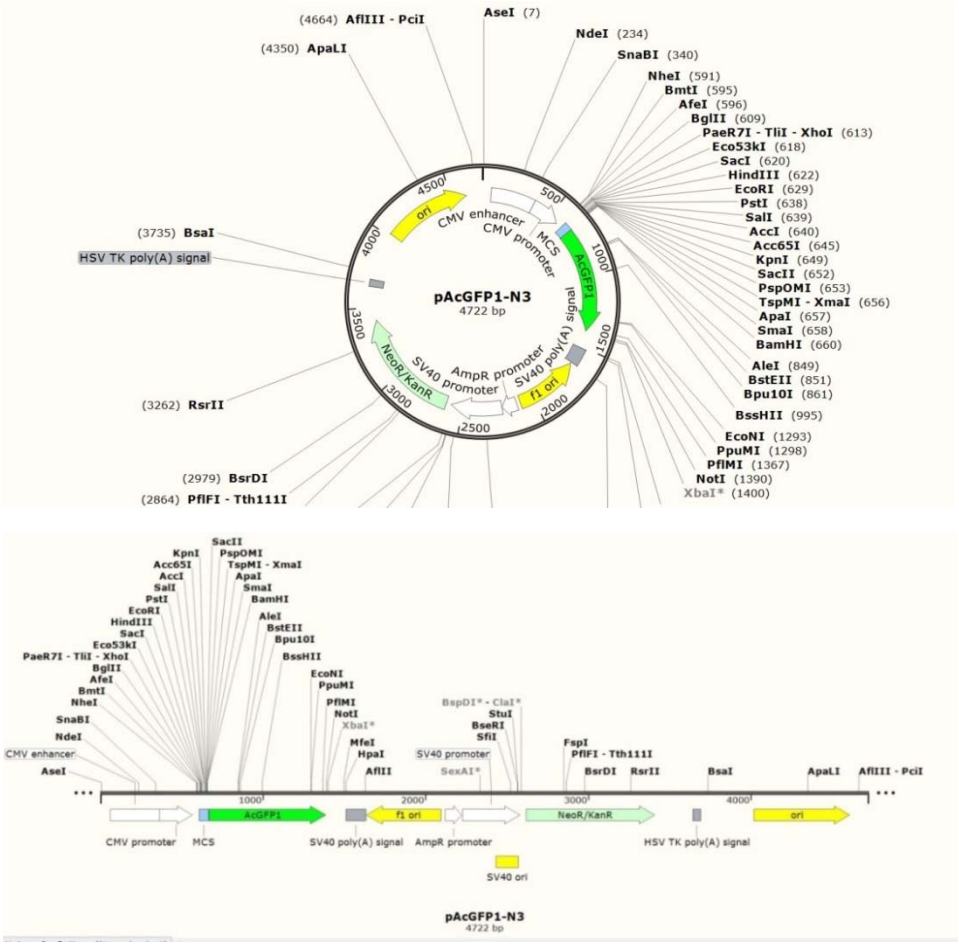


Figure 45. The pACGFP1-N3 Wild Type Plasmid Manipulation. It contains the CMV and CMV enhancer to highly express in mammalian cells. It was digested with HindIII and PciI restriction enzymes, the whole promoter regions (around 600 bp) replaced with active promoter region of thehuman ChAT gene (174 bp). The plasmid was extracted from bacterial cells (JM109) using PureYield™ Plasmid Midiprep System from Promega (Catalog #: A2492).

4600 - 4700 Sequence: (PciI- 4662) 5'ACATGT3'

```
4601 TATGGAAAAA CGCCAGCAAC GCGGCCTTTT TACGGTTCCT GGCCTTTTGC 4650

4651 TGGCCTTTTG CTCACATGTT CTTTCCTGCG TTATCCCCTG ATTCTGTGGA 4700
```

600-750 sequence: (HindIII-622) 5'AAGCTT3'

```
601  CCGGACTCAG ATCTCGAGCT CAAGCTTCGA ATTCTGCAGT CGACGGTACC 650
651  GCGGGCCCGG GATCCATCAT GGTGAGCAAG GGCGCCGAGC TGTTCACCGG 700
701  CATCGTGCCC ATCCTGATCG AGCTGAATGG CGATGTGAAT GGCCACAAGT 750
```

Active human ChAT promoter sequence + RE + overhang:

5'GCTTGTCGACTCAGAGCTCTGAGGCTGGTGTGGGGCGTGTCCTCAAAGGGCGGGGC
CTGGCGTGCCTAGCTGGAGCGGGTCCTCAGGCTCCCCCGGGCGGCATCCCCAGGTCT
ATATAAGCAATGCT AAGCTTGACC3'-137 (137+10) =147 10= PciI and A overhang

Primers:

F: 5'- AAA AAC ATG TGCTT G TCG ACT CAG AGC TCT GAG GCT G -3'

R: 5'- GGT C AA GCT T AG CAT TGC TTA TAT AGA CCT GGG GAT G -3'

4600F primer, PciI (AC ATG T), Insert, Hind III (AAGCTTGACC),plus Actual GFP sequence:

CTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTT
TTGCTC**ACATG**TGCTT**GTCGAC**TCAGAGCTCTGAGGCTGGTGTGGGGCGTGTCCTCA
AAGGGCGGGGCCTGGCGTGCCTAGCTGGAGCGGGTCCTCAGGCTCCCCCGGGCGGC
ATCCCC**AGGTCTATATAAGCAATGCTAAGCTTGACC**CGAATTCTGCAGTCGACGGT
ACCGCGGGCCCGGGATCCATCATGGTGAGCAAGGGCGCCGAGCTGTTCACCGGCATC
GTGCCCATCCTGATCGAGCTGAATGGCGATGTGAATGGCCACAAGT

qRT-PCR

Primers designed for the human ChAT gene mRNA:

Ref#: NM_001142933.1

>gi|218931223:510-2528 Homo sapiens choline O-acetyltransferase (CHAT), transcript variant S, mRNA

GTCCACACCTCTGCATCCCTGCACCAGGACTCACCAAGACGCCCATCCTGGAAAAGGTCCCCCGTAAGAT
GGCAGCAAAAACTCCCAGCAGTGAGGAGTCTGGGCTGCCCAAACTGCCCGTGCCCCCGCTGCAGCAGACC
CTGGCCACGTACCTGCAGTGCATGCGACACTTGGTGTCTGAGGAGCAGTTCAGGAAGAGCCAGGCCATTG
TGCAGCAGTTTGGGGCCCCTGGTGGCCTCGGCGAGACCCTGCAGCAGAAACTCCTGGAGCGGCAGGAGAA
GACAGCCAACTGGGTGTCTGAGTACTGGCTGAATGACATGTATCTCAACAACCGCCTGGCCCTGCCTGTC
AACTCCAGCCCTGCCGTGATCTTTGCTCGGCAGCACTTCCCTGGCACCGATGACCAGCTGAGGTTTGCAG
CCAGCCTCATCTCTGGTGTACTCAGCTACAAGGCCCTGCTGGACAGCCACTCCATTCCCACTGACTGTGC
CAAAGGCCAGCTGTCAGGGCAGCCCCTTTGCATGAAGCAATACTATGGGCTCTTCTCCTCCTACCGGCTC
CCCGGCCATACCCAGGACACGCTGGTGGCTCAGAACAGCAGCATCATGCCGGAGCCTGAGCACGTCATCG
TAGCCTGCTGCAATCAGTTCTTTGTCTTGGATGTTGTCATTAATTTCCGCCGTCTCAGTGAGGGGGATCT
GTTCACTCAGTTGAGAAAGATAGTCAAAATGGCTTCCAACGAGGACGAGCGTTTGCCTCCAATTGGCCTG
CTGACGTCTGACGGGAGGAGCGAGTGGGCCGAGGCCAGGACGGTCCTCGTGAAAGACTCCACCAACCGGG
ACTCGCTGGACATGATTGAGCGCTGCATCTGCCTTGTATGCCTGGACGCGCCAGGAGGCGTGGAGCTCAG
CGACACCCACAGGGCACTCCAGCTCCTTCACGGCGGAGGCTACAGCAAGAACGGGGCCAATCGCTGGTAC
GACAAGTCCCTGCAGTTTGTGGTGGGCCGAGACGGCACCTGCGGTGTGGTGTGCGAACACTCCCCATTCG
ATGGCATCGTCCTGGTGCAGTGCACTGAGCATCTGCTCAAGCACATGACGCAGAGCAGCAGGAAGCTGAT
CCGAGCAGACTCCGTCAGCGAGCTCCCCGCCCCCCGGAGGCTGCGGTGGAAATGCTCCCCGGAAATTCAA
GGCCACTTAGCCTCCTCGGCAGAAAAACTTCAACGAATAGTAAAGAACCTTGACTTCATTGTCTATAAGT
TTGACAACTATGGGAAAACATTCATTAAGAAGCAGAAATGCAGCCCTGATGCCTTCATCCAGGTGGCCCT
CCAGCTGGCCTTCTACAGGCTCCATCGAAGACTGGTGCCCACCTACGAGAGCGCGTCCATCCGCCGATTC
CAGGAGGGACGCGTGGACAACATCAGATCGGCCACTCCAGAGGCACTGGCTTTTGTGAGAGCCGTGACTG
ACCACAAGGCTGCTGTGCCAGCTTCTGAGAAGCTTCTGCTCCTGAAGGATGCCATCCGTGCCCAGACTGC
ATACACAGTCATGGCCATAACAGGGATGGCCATTGACAACCACCTGCTGGCACTGCGGGAGCTGGCCCGG
GCCATGTGCAAGGAGCTGCCCGAGATGTTCATGGATGAAACCTACCTGATGAGCAACCGGTTTGTCCTCT
CCACTAGCCAGGTGCCCACAACCACGGAGATGTTCTGCTGCTATGGTCCTGTGGTCCCAAATGGGTATGG
TGCCTGCTACAACCCCCAGCCAGAGACCATCCTTTTCTGCATCTCTAGCTTTCACAGCTGCAAAGAGACT
TCTTCTAGCAAGTTTGCAAAAGCTGTGGAAGAAAGCCTCATTGACATGAGAGACCTCTGCAGTCTGCTGC
CGCCTACTGAGAGCAAGCCATTGGCAACAAAGGAAAAAGCCACGAGGCCCAGCCAGGGACACCAACCTTG
ACTCCTGCCACTAGGTTTCACCTCCCAAACCCAGCCTCTAGAACAGCCAGACCCTGCAG

Primer design

Amplicon length: 83

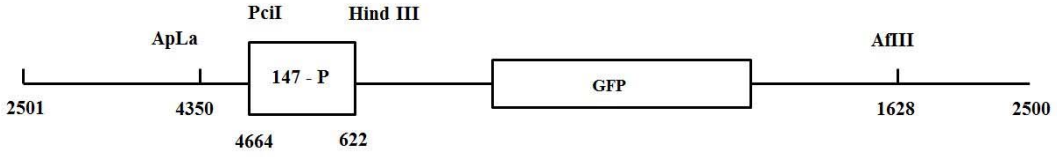
F: **TGTGGTCCCAAATGGGTATG** (sense) Length: 20 TM: 62 GC%: 50

R: **TGCAGCTGTGAAAGCTAGAG** (antisense) Length: 20 TM: 62 GC%: 50

APPENDIX C

AFM DNA SAMPLE PREPERATION AND INITIAL IMAGING

pAcGFP1-N3- modified plasmid



Cut- pAcGFP1-N3- modified plasmid

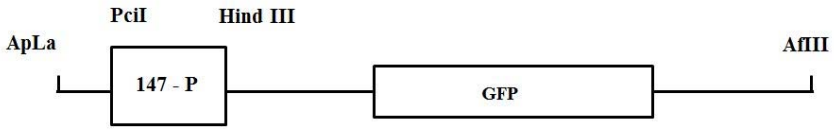


Figure 46. AFM Plasmid Constuction. A) pAcGFP1-N3- modified plasmid (4189 bp)with active promoter region of the human ChAt gene sequence, 147bp (wild type CMV promoter cut out). B) Cut- pAcGFP1-N3- modified plasmid (1508 bp), double digested fragment between ApLa and AflII restriction enzymes.

A B

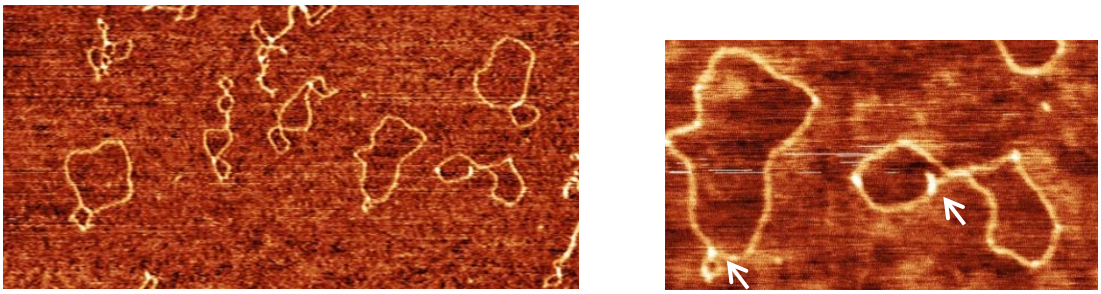


Figure 47. AFM Images of G-wires on pAcGFP1-N3- Modified Plasmid (1508 bp DNA Fragment). The image of G- wires is 5x5 μm^2 at 2-nm height scale (A), and 600x600 nm^2 at 2-nm height scale. DNA substrate were incubated in a buffer containing 100mM KCl and deposited at 500 nM concentration. The white line arrows indicated the position of G-wires.

Table 10. Single - Temprature Digestion Protocol**.** Single and double digestion of pAcGFP1-N3- modified plasmid reaction mixture using ApLa and AfIII restriction enzymes, incubation at 37 °C for 1-2 h.

Reaction mixture	Double	Single	single
DNA template	5µL	2.5µL	2.5µL
NEB 2.1 butter (10x)	5µL	2.5µL	2.5µL
ApLa	1µL	1µL	-
AfIII	1µL	-	1µL
dH_2O	Up to 50µL	Up to 25µL	Up to 25µL

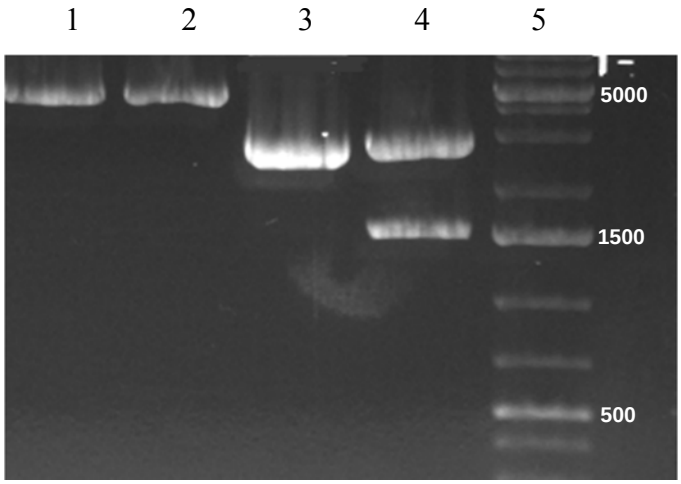


Figure 48. Digestion Image of pAcGFP1-N3- Modified Plasmid. Both single and double digestions performed at 37 °C incubation for 1-2 hr, then samples run on a 1% agarose gel and imaged by UV-shadowing. Lanes 1 & 2 are for single digestion (ApLa & AfIII, respectively). Laneis for un-cut plasmid an dlanes 3 &4 are for double digestion (ApLa & AFIII) & Marker, respectively.