Molecular Approaches for Intervening Biological Functions of DNA Secondary Structures

Thesis Submitted for the Degree

of

Doctor of Philosophy (Science)

In Life Science and Biotechnology

By

Khushnood Fatma

Jadavpur University, Kolkata, India

2023

CERTIFICATE FROM THE SUPERVISOR

This is to certify that the thesis entitled **"Molecular Approaches for Intervening Biological Functions of DNA Secondary Structures"** submitted by **Smt. Khushnood Fatma**, who got her name registered on 20.09.2018 for the award of Ph. D. (Science) Degree of Jadavpur University, is absolutely based upon her own work under the supervision of **Prof. Jyotirmayee Dash** and that neither this thesis nor any part of it has been submitted for any degree/diploma or any other academic award anywhere before.

Dash 6.10.23

Prof. Jyotirmayee Dash

[Signature of the Supervisor(s) & date with official seal]

Prof. Jyotirmayee Dash
Senior Professor
School of Chemical Sciences
IACS, Jadavpur
Kolkata - 700032

Prof. Jyotirmayee Dash
Senior Professor
School of Chemical Sciences
IACS, Jadavpur
Kolkata - 700032

DECLARATION

The research work embodied in this thesis entitled **"Molecular Approaches for Intervening Biological Functions of DNA Secondary Structures"** being submitted to Jadavpur University, Kolkata, has been carried out at the Indian Association for the Cultivation of Science, Jadavpur, under the supervision of Prof. Jyotirmayee Dash, Senior Professor, School of Chemical Sciences, Indian Association for the Cultivation of Science. This work is original and has not been submitted in part or in full, for any degree or diploma to this or any other university.

KHUSHNOOD FATMA

Index No. **187/18/Life Sc./26**

Dedicated to

my family

Acknowledgements

The last few years of my life have been single-mindedly driven with one goal and one goal only---to give my best effort towards doing my research work and seeing the fruits of my labor. As now I have arrived near the culmination point of this journey of mine that for me is my Ph.D., I realize that this journey of mine hasn't been mine alone. I may have been in the driver's seat throughout this journey but it is because of innumerable people in my life that my journey has been as fruitful as I envisioned when I first embarked upon this unknown path that now I have traversed through. The last few years have been an incredible journey akin to a rollercoaster ride with its fair share of lows and highs. And above all it has been a period of immense growth both personally and professionally.

The people that I feel are the recipients for my utmost gratitude for their unwavering support during my Ph.D. journey are innumerable. So I would take a little time to thank every last one of them, one at a time so that every person is paid their due respect and are appreciated.

First and foremost, my deepest gratitude is for my Ph.D. supervisor, Prof. Jyotirmayee Dash. It has been her expertise, guidance, support and patience that has allowed me to grow as a researcher. I have learnt many skills from her throughout my Ph.D. tenure - be it the ability to design and execute experiments or just having the curiosity to explore to see what results are found in doing more confirmatory experiments. The scientific writing skills is something that my supervisor excels at and under her tutelage I have attempted to learn this skill as well. The skills to work independently with the right guidance cannot be undermined and that is what I have received in this lab group which has allowed me to hone my scientific aptitude.

I am also very thankful to my senior, Puja di, from whom I have learnt a lot in the short time that I received her guidance during my early Ph.D. years. I thank my lab mates, Tania di, Manish da, Panda da, Rakesh da, Seemantee di, Joyshree di, Sayantan da, Dipayan da, Sameer da, Subhro da, Tanima di, Tridev da, Gargi di, Tirtha da, Debasish da, Snehashish da, Joydev da ,Ritapa, Prasanth, Raj, Shilpi, Debasmita, Neha, Somnath, Nabin, Binayak, Soumya, Pratap.

My heartfelt gratitude and a big shout-out goes to my friends that I have been fortunate to make in my lab during my Ph.D years. Ritapa,has been an incredible friend, labmate and co-

worker. She has been the ear to listen to all my woes, tirades of rants and happiness equally. She has been my friend not only within the boundaries of the lab but even beyond it. Prasanth for being a great human being, a friend and an incredible co-worker. It has been a privilege to work with him.

My childhood friends are special to me because they have known me for the longest period and have always given me their support. Fatima and Srabani have been my constants in life while growing up and as I continue to traverse through life, they remain extremely special. They have always been there to check up on me and help me in times of crises and shared my fortunes and misfortunes alike. They are simply irreplaceable in both my heart and my life.

My beloved friend Priya, who has been my constant in my life since the days of my M.Sc course, when I first dared to dream to one day be a Ph. D. degree holder. She is also someone who has never lost touch and has always been there for me in my good times and bad times.

My longtime school friends, Oindrilla, Ayesha, Fauzia, Ifrah, Nazia, Shinjini and college friends Rini, Samin. As the years pass and we go about our separate ways in terms of career, these friends have always kept in touch and together we continue to revisit the nostalgia of the past while creating new memories along the way.

I would also like to thank all the academic and non-academic personnel of IACS for making my journey at IACS smoother and easier. My time at IACS would have been much more difficult without their guidance in official matters.

I would also like to thank the eminent institution of IACS for giving me the platform and the means to conduct my research work. And I hope to be a proud alumni of this institution.

In this journey my family gave me the strength to pursue it all — my father, Syed Mahmood Raza, my mother, Shabana Khanam and my brother, Syed Shabbar Raza. They never held me back from anything. They gave me the independence that I required to grow into a dependable person. They gave me the freedom that I sought to live my life as I deem fit and taught me that making mistakes is a part of life and that I should take my successes and failures in stride.

My family is the pillar that forever stands in my support. The time and tide is inconsequential when it comes to the support that I have received from them and continue to get each and

every day. This is the support that I cannot bear to be bereft from for they are not just my support system but also my strength and driving force. They are the ones that inspire me and gave me the strength to dare to dream. They let me know that the ability to dream is the starting point of making that dream come true. They are both my dream and my reality for now and for always.

I have always been sincere and hardworking to be able to reach this goal. Sometimes, I have even taken longer to complete projects solely because I wasn't satisfied with the results and wanted to go the extra mile no matter the additional effort. Also, for me doing anything half-baked and half-heartedly just never has been an option. This aspect of mine has been nothing short of a double edged sword. It has led to uncountable disappointments and extra hours spent working tirelessly but has also brought the confidence and contentment that can only be achieved by putting my best effort into something that I believe with all my heart. In retrospect, at the end of this journey, it has also been highly rewarding and a constant source of self-motivation as I have gained more than I ever thought possible. So now I believe that my Ph.D. tenure hasn't only been a journey where I learnt and gained new insights but it has been an experience like no other— Ph.D. for me is inexplicable!

Finally I would like to thank "the Almighty" for the fortunes and opportunities that I have been blessed with to carve out my own destiny.

Thank you everyone.

Khushnood Fatma

Senior Research Fellow,

Indian Association for the Cultivation of Science,

Kolkata-700032

Table of Contents

List of abbreviations

° C	Degree Celsius
aq.	aqueous
CD	Circular dichroism
DC_{50}	Half-maximal dissociation concentration
DMS	Dimethylsulphate
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
ds	Double stranded
Eq.	Equivalent
FACS	Fluorescence-activated cell sorting
FAM	Fluorescein amidites
FBS	Fetal bovine serum
FID	Fluorescent indicator displacement
FRET	Förster Resonance Energy Transfer
G	Guanine
G4	G-quadruplex
h	Hour
H_2O	Water
H-Bonding	Hydrogen bonding
IC_{50}	Half-maximal inhibitory concentration
ITC	Isothermal titration calorimetry
K	Potassium
K_d	Dissociation constant
min	Minute
Na	Sodium
NMR	Nuclear magnetic resonance
PDB	Protein data bank
PI	Propidium iodide
RNA	Ribonucleic acid

List of Abbreviations

RT	Room temperature
TAMRA	Tetramethylrhodamine
T_m	Melting temperature
UTR	Untranslated region
UV	Ultraviolet
Vis	Visible
XTT	(2,3-Bis-(2-Methoxy-4-Nitro-5-Sulfophenyl)-2H-Tetrazolium-5-Carboxanilide)
ΔG	Gibbs Free energy

Molecular Approaches for Intervening Biological Functions of DNA Secondary Structures

Abstract

Nucleic acids exist as deoxyribonucleic acid (DNA) and ribonucleic acid (RNA). Apart from the well-known canonical double helix, nucleic acids can also fold into various non-canonical secondary structures. The non-canonical DNA secondary structures include G-quadruplex, i-motif, Z DNA, hairpin, cruciform, triplex, etc. G-quadruplex structures are widely abundant in the eukaryotic genomes at regions of biological significance, like the telomeres (DNA and RNA G-quadruplexes), promoters of oncogenes (*c-MYC*, *BCL-2*, *c-KIT*, *k-RAS*, etc), untranslated regions of genes (*n-RAS*, *VEGF*, etc), as well as hotspots of recombination. These G-quadruplex secondary structures can act as molecular switches having roles in gene regulation, DNA damage and repair, genomic instability and are involved in diseases like cancer, diabetes, and neurological disorders.

In recent times, G-quadruplexes have evolved from mere structural curiosity to highly integral therapeutic targets evidenced from their functions in regulation of essential cellular processes like transcription, replication, genome integrity and epigenetic expression. Thus, therapeutic intervention of these G-quadruplex structures using small molecules are considered as alternative strategies for anticancer discovery.

Chapter 1 provides a general introduction of the G-quadruplex secondary structure. It also details the function and occurrence of both DNA and RNA G-quadruplexes in the geneome. Moreover, the therapeutic strategy to target G-quadruplexes have been discussed. The role of computational and biological tools to search and explore novel G-quadruplex targets have also been emphasized.

Chapter 2 demonstrates the inhibition of the most explored G-quadruplex found in the telomeric region of eukaryotic chromosomes. Stabilization of telomeric G-quadruplexes has been reported as an effective method of inhibiting telomerase mediated telomere lengthening. Indoloquinoxaline containing small molecules were screened and it was found that **IQ2** with carboxamide side chain exhibited selective binding interaction with the human telomeric G-quadruplex DNA and inhibited telomerase activity. The ligand also

triggered DNA damage, induced cell cycle arrest in S and G2/M phases and activated caspases-3/7 that ultimately led to cancer cell death through apoptotic pathway.

Chapter 3 illustrates the downregulation of an oncogenic promoter gene by finding a selective G-quadruplex ligand for the *c-KIT* G-quadruplex which is implicated in different cancer types including gastrointestinal stromal tumors, melanomas, mastocytosis, and acute myeloid leukemia. Different biophysical and biological assays were used to study the interaction of a library of small molecules with different side chains and different heteroaromatic scaffolds to ascertain the most effective G4 binder for *c-KIT* G4. A carbazole derivative, **MC-4** was found to have the highest capability for selective recognition and stabilization of the parallel *c-KIT 1* G4. Detailed investigations revealed that the **MC-4** facilitates the visualization of stabilized cellular G4s. In addition, the ligand induced *c-KIT* gene downregulation, cell cycle arrest and apoptosis in leukemia cells.

Chapter 4 emphasized the need to find novel G4 structures in the DNA and RNA sequences of a critically important gene, mammalian target of rapamycin (*mTOR*) which is implicated in various cancers. Bioinformatics tools were employed to explore the presence of potential G-quadruplex forming sequences in the mTOR gene. It was found that a G-rich sequence was capable of forming the G-quadruplex structure both in the DNA (P2) as well as the 5'UTR of the mRNA (rP2) of *mTOR* gene. The presence of the novel *mTOR* G-quadruplex was confirmed using various techniques including circular dichroism, DMS footprinting, nuclear magnetic resonance (NMR) and its biological implications were affirmed using gene expression studies. This is the first report of a novel G-quadruplex in the mRNA of the *mTOR* gene with a role in the gene regulatory mechanism of *mTOR*.

Dash 6/10/2023

Signature of Supervisor

Signature of Candidate

Prof. Jyotirmayee Dash
Senior Professor
School of Chemical Sciences
IACS, Jadavpur
Kolkata - 700032

Chapter 1

Introduction

1.1.Nucleic Acid secondary structures

Nucleic acids play indispensable roles in the preservation and transmission of vital genetic information in all living organisms. Deoxyribonucleic acid (DNA) serves as the repository of instructions for cellular proteins, encompassing the genetic code for replication and transcription to produce messenger RNA (mRNA), and ultimately leading to protein synthesis. The study and elucidation of the DNA structure and function unravels fundamental insights of the storage and inheritance of genetic data, thereby ensuring the transfer of genetic data to the next generation.[1]

Double-stranded DNA consists of two helical strands held together by Watson-Crick bonds, predominantly adopting the canonical right-handed double helical structure known as B-DNA. However, certain environments and sequence motifs of DNA tend to favor alternative DNA conformations, known as non-canonical secondary structures. These secondary structures are also formed in ribonucleic acid (RNA).

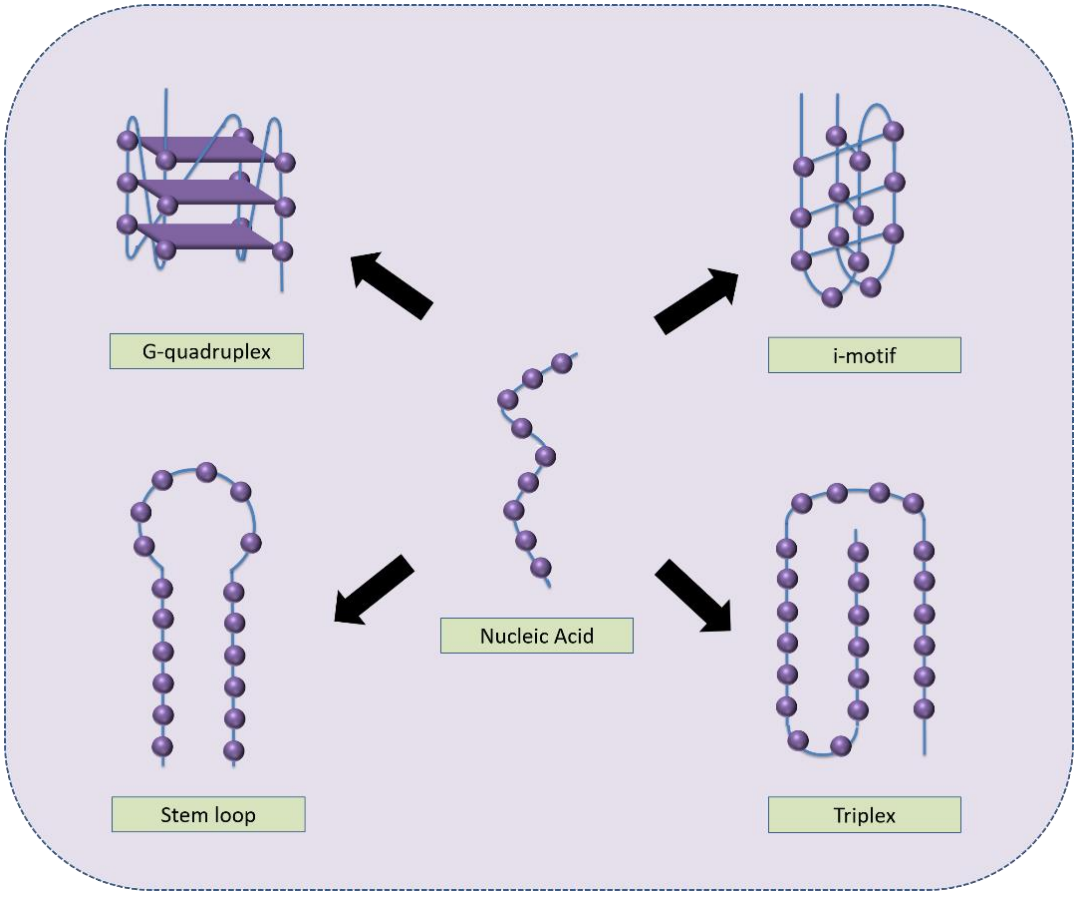


Figure 1.1. Nucleic acids including DNA and RNA can adopt different secondary structures depending on sequence and physiological conditions.

DNA and RNA sequences can form a range of secondary structures including hairpins, cruciforms, R-loops, G-quadruplexes, i-motifs, triplexes, Z-DNA, Z-RNA, etc (Figure 1.1).[2] Sequences that can adopt a secondary structure are often enriched at regulatory regions of the genome, including open chromatin regions, promoters, 5'UTRs and 3'UTRs. Importantly, they are over-represented in regions governing key cellular processes, such as transcription start and transcription end sites, splice junctions and translation initiation.[3, 4]

Hairpin structures can arise in regions containing inverted repeats of sequences consisting of two adjacent copies of the same sequence. One of these copies is oriented in a reverse complement direction, facilitating the formation of a hairpin structure through hydrogen bonding between the two complementary sequences, while the intervening spacer remains single-stranded. A closely related secondary structure, known as a cruciform, is composed of two hairpin structures converging at a 4-way junction.[5]

Another secondary structure is formed by the assembly of nucleic acid where one strand folds back to join the double-stranded DNA, and is stabilized by reverse-Hoogsteen or Hoogsteen bonds. The other strand remains single, and the result a triplex consisting of a triple helical structure is formed. Various triplex structures have been identified in interactions between homopyrimidine-homopurine duplexes and single-stranded oligonucleotides.[6]

In addition to these structures, G-quadruplexes and i-motifs are tetraplex structures formed in G-rich and C-rich regions of the genome respectively (Figure 1.2).[7, 8]

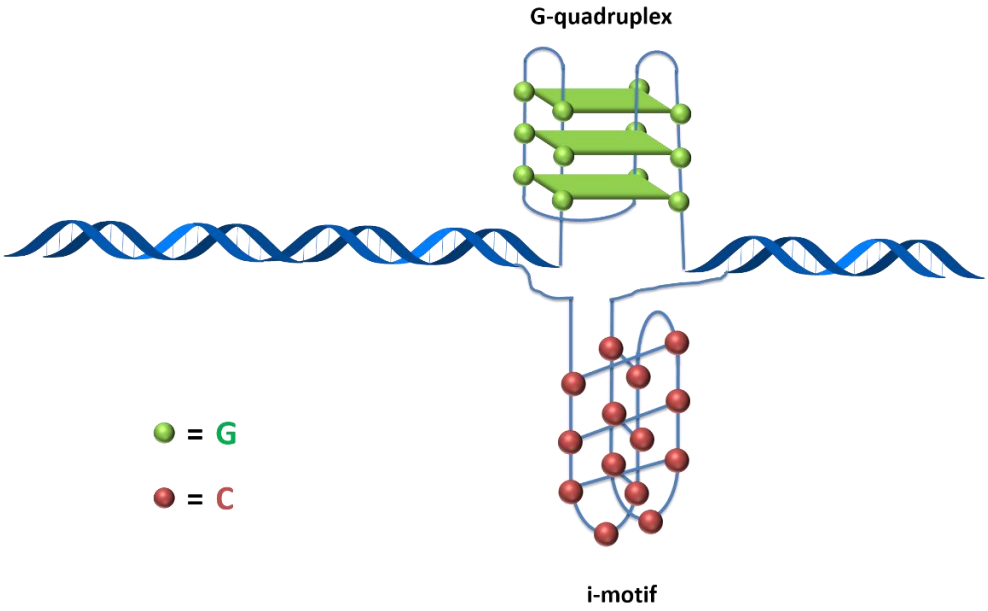


Figure 1.2. The G-rich and the C-rich regions of the DNA form G-quadruplex and the i-motif respectively

1.2. G-quadruplexes

While the peculiar nature of guanine-rich (G-rich) DNA solutions to form gel-like aggregates was initially observed as far back as 1910, the precise character of these aggregates remained elusive until 1960s.[9, 10] In the early 1960s, observations pertaining to the self-assembly of guanylic acid resulted in the identification of the guanine tetrad-forming sequence motif. In this motif, guanine molecules interconnected through Hoogsteen hydrogen base-pairing, creating a planar arrangement (Figure 1.3) .[11] This structure is further stabilized by interactions between positively charged ions and the O-6 lone-pair electrons of each guanine. The initial evidence for the formation of four-stranded G-quadruplex (G4) structures from natural sequences came from *in vitro* experiments where higher-order secondary structures formed from oligonucleotides resembling G-rich sequences found in immunoglobulin switch regions.[12] Subsequent investigations using biophysical and structural biology techniques provided substantial physical proof for the

formation of both intermolecular and intramolecular G4 structures from DNA and RNA *in vitro*. This research also established a framework for identifying sequences likely to fold into G4 structures and led to subsequent research in this field.

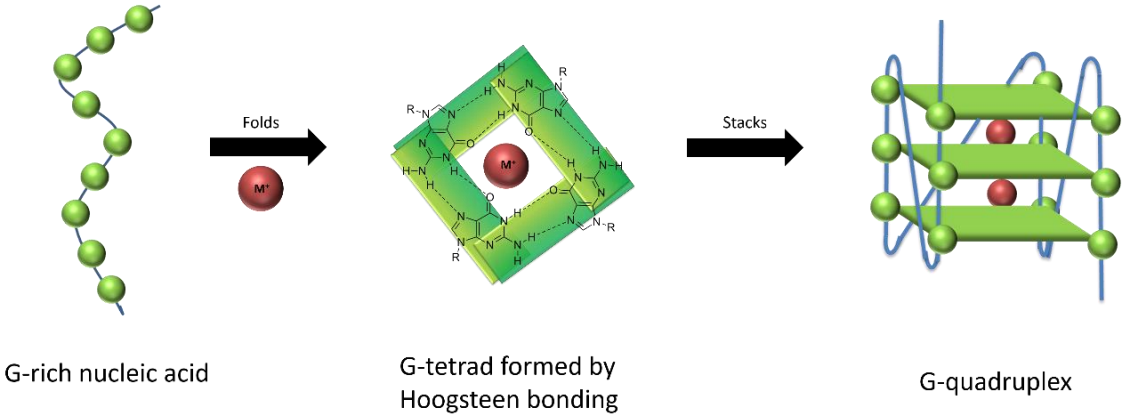


Figure 1.3. The Gs in the G-rich strand of nucleic acid associate *via* Hoogsteen bonding to form a square planar G-tetrad. The G-tetrads stack upon each other to adopt the folded G-quadruplex secondary structure.

G-quadruplex DNA structures commonly referred to as G4 DNA have attracted considerable interest in recent years. G4 DNA structures characteristically deviate from the canonical B-form Watson–Crick double helix by their Hoogsteen hydrogen bonding. G4s are evolutionarily conserved four stranded DNA structures that occur in guanine (G) rich regions of the genome such as telomeres, ribosomal DNA, immunoglobulin class switch regions within the IgG locus and G-rich micro and minisatellites. Typically, guanine-rich consensus sequences ($G_3N_{1–7}G_3N_{1–7}G_3N_{1–7}G_3$) tend to fold into a G4 structure.[13] Conventionally, four guanine nucleotides associate *via* Hoogsteen hydrogen bonding to form a G-tetrad which is a square planar structure. G4s are stabilized by a monovalent metal ion in its center. Stacking of these planar structures leads to G4 formation.

Thus, the G4 structure consists of the core G-tetrad components stacked on top of each other without interaction with the connecting loops of variable composition. The G-rich core typically consists of two or more stacked G-tetrads with a right-handed helical

twist.[14] The stacks are joined together by the normal sugar–phosphate backbone. The binding energy responsible for the G4 folded structure is mainly due to three factors: hydrogen-bonding between the guanines in a plane, π–π interactions between the guanines in adjacent planes and charge–charge interactions between the partially negative O6 of the guanines and cations in the octahedral position between the stacks.[15] Monovalent cations like K^+, Na^+, NH^+, etc, especially K^+, are stabilizing factors resulting to stable G4 formations. In addition, varying any of the nucleotide bases to non-G bases results in a highly destabilizing effect which may disrupt the secondary structure or render the G formation unlikely *in vivo*.

The intervening variable-length sequences composed of bases other than G, hold together the G-quartets and form the loops that are arranged on the exterior of the G4 core. The loops are one of the major elements that define structural variability in G-quadruplexes, and are analogous to side-chains in amino acids.

1.3. G-quadruplex –Structure and topology

Inspite of all quadruplexes containing the basic motif of the the G-tetrad, comprised of four guanine bases held in plane by Hoogsteen hydrogen bonding, a range of polymorphism is observed in G-quadruplexes resulting in a diverse family of structures that can adopt various topologies. The different topologies of G-quadruplexes are determined by the orientation of interconnecting loops and the arrangement of strand polarities.[15] When all the Gs participating in the formation of the G-quadruplex arise on the same nucleic acid strand it gives rise to the intramolecular G-quadruplex. Furthemore, when two or more strands are involved in the G-tetrad formation, it gives rise to intermolecular G-quadruplex (Figure 1.4).[16]

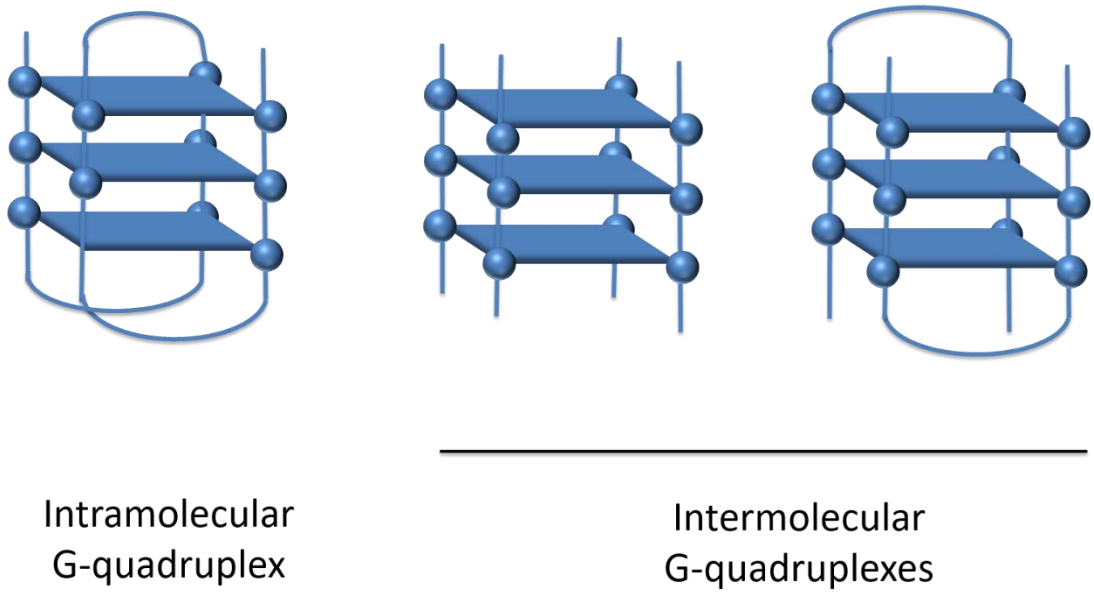


Figure 1.4. G-quadruplexes adopt intramolecular or intermolecular topologies depending on the nature and number of the nucleic acid strand involved in the gormation of the G-quadruplex.

Another distinguishing feature to determine the G-quadruplex topology is the stand orientation. The nature of the loops is related to the directionality of the four G-rich tracts making up the core of the G4 structure. Based on the nature of the strand orientation, G4s may be – (i) parallel (containing double chain reversal loops); (ii) antiparallel (with two strands running in each direction and either lateral or diagonal loops); and (iii) mixed or hybrid structure with a mixture of loop types (Figure 1.5). Which of these structures is preferably adopted by the G-rich strand depends on the sequence and length of the loops, where shorter loops tend to favour parallel conformations.

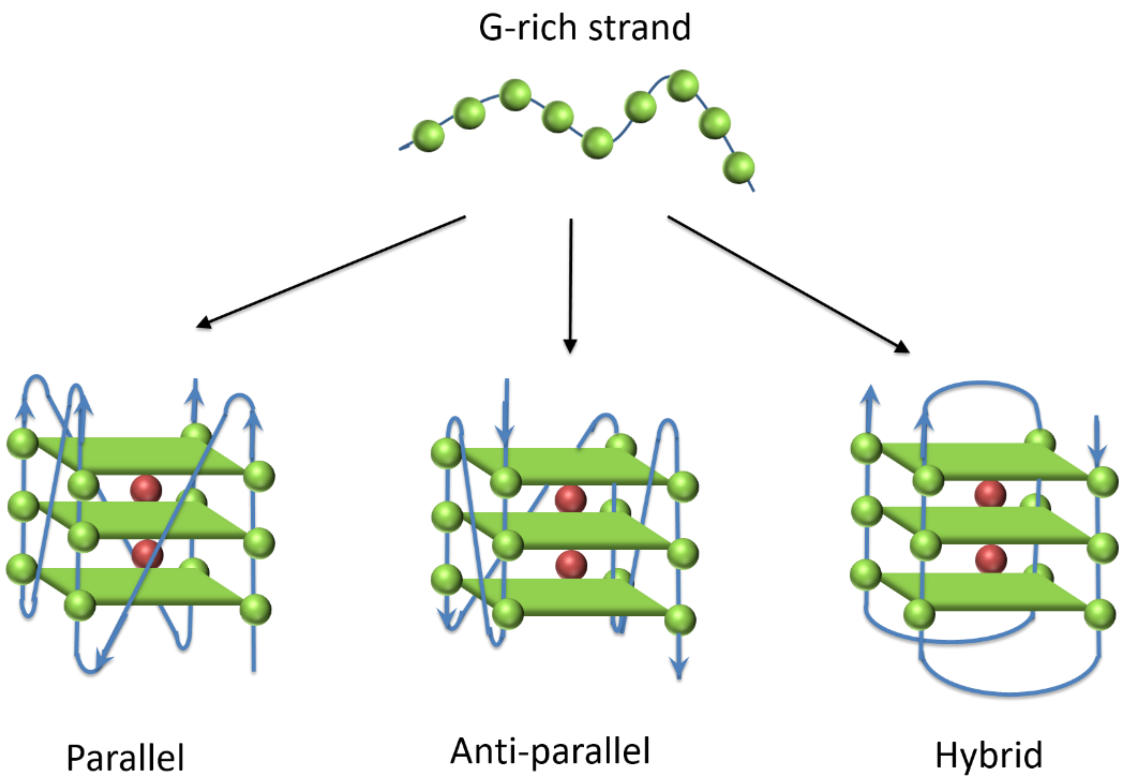


Figure1.5. The G-rich nucleic acid sequence can fold into mainly three types of G-quadruplexes including parallel, antiparallel and mixed topologies based on the direction of the nucleic acid strand.

1.4. Factors affecting formation of G-quadruplexes

The unique negative charge channels of the G-quadruplex promotes interaction with cations, including K^+, Na^+, and Li^+ through electrostatic forces.[17] Monovalent cations such as K^+ and Na^+ stabilize the G-quadruplex structures by coordinating with eight carbonyl oxygen atoms present between stacked tetrads. Among the cations, K^+ has the strongest stabilizing effect on G4s, and high intracellular levels of K^+ are highly conducive to forming G4s.[18] Thus, K^+ is routinely used *in vitro* to induce G4 formation.

1.5.Prevalence of G-quadruplexes

i. DNA G-quadruplexes

Recent genome-wide sequence analyses employing G-quadruplex signature sequences have provided significant insights into the presence and distribution of sequences with the potential to form stable G-quadruplexes. These G4s are abundant in both prokaryotic and eukaryotic genomes, spanning from yeast to humans. Remarkably, the human genome alone contains over 3,00,000 putative G4 sequences.[19, 20] While the actual occurrence of G4s in genomes is about a third less than expected by random chance, their occurrence is not random and appears to correlate with functional genomic regions.

Predictably, G4s have been found to be most prevalent in repetitive DNA segments. These segments include the G-rich 3'overhang of telomeres. In addition, several investigations have unveiled the widespread occurrence of G-quadruplex (G4) structures within the promoter regions of genes, suggesting their potential regulatory roles in gene transcription (Figure 1.6).[21] The initial discovery of a G4 structure in a promoter region was made within the nuclease hypersensitivity element III1

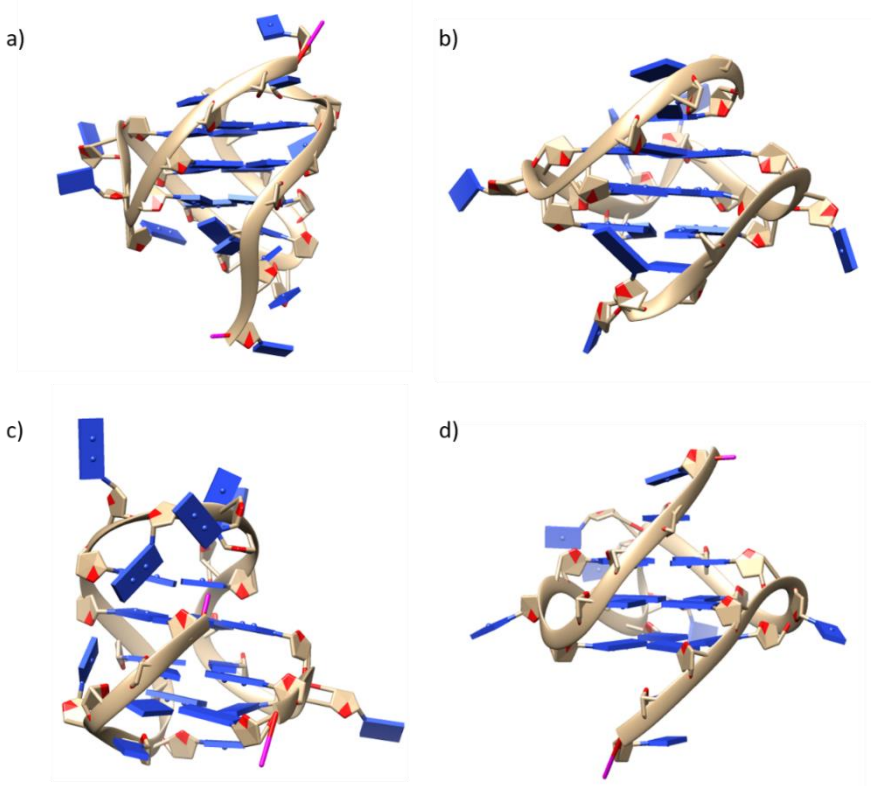


Figure 1.6. The PDB structures of DNA G-quadruplexes a) *h-TELO* (PDB ID: 1XAV), b) *c-MYC* (PDB ID: 2JPZ), c) *BCL-2* (PDB ID: 2F8U), d) *VEGF* (PDB ID: 2M27).

(NHE III1) of the proto-oncogene *c-MYC*.[22] NHE III1 is situated upstream of the primary promoter 1 (P1) of the *c-MYC* gene and plays a pivotal role in regulating its transcription. Moreover, G4s have been studied in the promoter regions of various genes, including proto-oncogenes like *VEGF, BCL-2, PDGF-A, c-KIT, RET* and *HIF-1* as well as DNA repair gene *RAD17*, the human platelet-derived growth factor receptor beta PDGFR-β, etc (Table 1.1).

The presence of G4 structures in these promoter regions interferes with the interactions between DNA and its transcription factors, thereby exerting regulatory control over transcription machinery. Accumulating evidence from several studies demonstrates that G4 ligands have the capacity to reduce the expression of these genes, indicating that the presence of G4 structures may serve as a molecular switch influencing gene transcription.

Table 1.1: List of well explored DNA G-quadruplexes.

G-quadruplex	**Sequence (5'-3')**	**Topology**
c-MYC	$T_2ATG_4A\ G_3TG_4AG_3TG_4A_2G_2$	Parallel
c-KIT-1	$CAGAG_3AG_3CGCTG_3AG_2AG_4CTG$	Parallel
c-KIT-2	$C_4G_3CG_3CGCGAG_4AG_4AG_2C$	Parallel
k-RAS	$G_3AG_3AG_3AAG_2AG_3AG_3AG_3A$	Parallel
VEGF	$C_3G_4CG_3CCG_4CG_4TC_3G2CG_4CG_2A$	Parallel
HIF-1α	$GCGAG_3CG_5AGAG_4AG_4CGCG$	Parallel
BCL-2	$GTCG_4CGAG_3CG_5A_2G_2AG_3CGCG_3CG_4A$	Anti-parallel/Mixed
h-TELO (Na^+)	$T_2AG_3\ T_2AG_3\ T_2AG_3\ T_2AG_3$	Anti-parallel
h-TELO (K^+)	$T_2AG_3\ T_2AG_3\ T_2AG_3\ T_2AG_3$	Mixed

Furthermore, G-quadruplexes have been identified in other biologically significant genomic regions, including but not limited to immunoglobulin switch regions and ribosomal DNA. Certain genetic disorders, such as Fragile X syndrome, are associated with repeat expansion sequences $(CGG)_n$ that can form G-quadruplex structures, affecting gene expression and protein function. G-quadruplexes can also be found in other regulatory

elements of the genome, including enhancers and silencers, where they may influence gene expression.

ii. RNA G-quaruplexes

DNA predominantly exists in a double-helical structure, the formation of G-quadruplexes by G4s necessitates the temporary disruption of canonical Watson–Crick base-pairing to allow for Hoogsteen base-pairing. Such opportunities arise during key cellular processes like DNA replication, transcription, and recombination when sections of double-stranded DNA transiently become single-stranded. Regions naturally occurring as single-stranded, and thus energetically favorable for G-quadruplex formation, are located at the very ends of telomeres and RNA. It has been documented that as many as 3,000 messenger RNAs (mRNAs) contain putative G4s within their 5' untranslated regions (UTRs).

Recent studies have demonstrated that mammalian telomeres can be transcribed to form telomeric repeat-containing RNA (TERRA). Notably, TERRA has been observed to adopt G-quadruplex structures within cells.[23, 24] In addition, G-quadruplexes formed in the telomere RNA (TERRA) can exert an influence on chromosome elongation (Figure 1.7 a). These G-quadruplexes adopt an antiparallel arrangement with guanosine and adopt a *syn/anti* conformation.[25]

Molecular mechanisms governing gene expression can operate both prior to and subsequent to the process of transcription. Post-transcriptional regulation mechanisms can lie within the sequence or structural features of RNA transcripts. It has been hypothesized that nucleic acid sequence motifs capable of forming G-quadruplexes may influence both transcription and translation. Investigation into the prevalence of G-quadruplex motifs in the vicinity of mRNA untranslated regions (UTRs) suggest enrichment of G4s toward the 5'-ends of 5'-UTRs, supporting the notion of a potential link between 5'-UTR RNA G4s and translational regulation. Additionally, G4 motifs at 3'-UTRs have also been found, particularly immediately following the termination of genes, especially in cases where another gene is in close proximity.[26] This observation suggests a possible role for G-quadruplexes in gene transcription termination.

Gene expression regulation encompasses diverse mechanisms, operating at both the level of DNA and RNA to influence protein expression.[27] Post-transcriptional regulation

involves the control of mRNA processing, nucleocytoplasmic transport, cellular and subcellular localization, translation efficiency, and stability. Studies have demonstrated that the genetic information essential for post-transcriptional control primarily resides in the 5' end and 3' end of UTRs of mRNA (Figure 1.7 b). These regulatory processes may involve both the primary nucleotide sequence and the secondary structural features of non-coding elements further implicating an integral role of G4 formation in these regions.

Secondary structures formed within the 5'- and 3'-UTRs can also function as regulatory elements by acting as target sites for RNA-binding proteins or small molecule metabolites or by directly interacting with the translation machinery. Some of the noteworthy G-quadruplexes which have been studied are listed in Table 1.2.

Recent advancements in methodologies, including the development of predictive algorithms and structure-based sequencing, have enabled the comprehensive detection and mapping of RNA G-quadruplex structures within transcriptomes on a large scale. Numerous genomic studies have consistently revealed that a substantial portion of RNA G-quadruplexes predominantly forms within the UTRs of mRNA molecules, where they recruit translational proteins and modulate translation. In vitro experiments have confirmed the involvement of G-quadruplexes in gene translation regulation.

For instance the oncogene *NRAS*, harbors a characteristic G-quadruplex-forming sequence in its 5'-UTR.[28] The formation of this G4 structure has been demonstrated to inhibit gene expression. Furthermore, a G-quadruplexes in its 5'-UTR of proto-oncogene *VEGF-A* has been shown to play a pivotal role in regulating cap-independent translation initiation.[29] Notably, the translational protein eIF4A can recognize the CGG repeat sequence within the UTR and facilitate the progression of T cell acute lymphoblastic leukemia by unwinding the G-quadruplexes within this repeat.

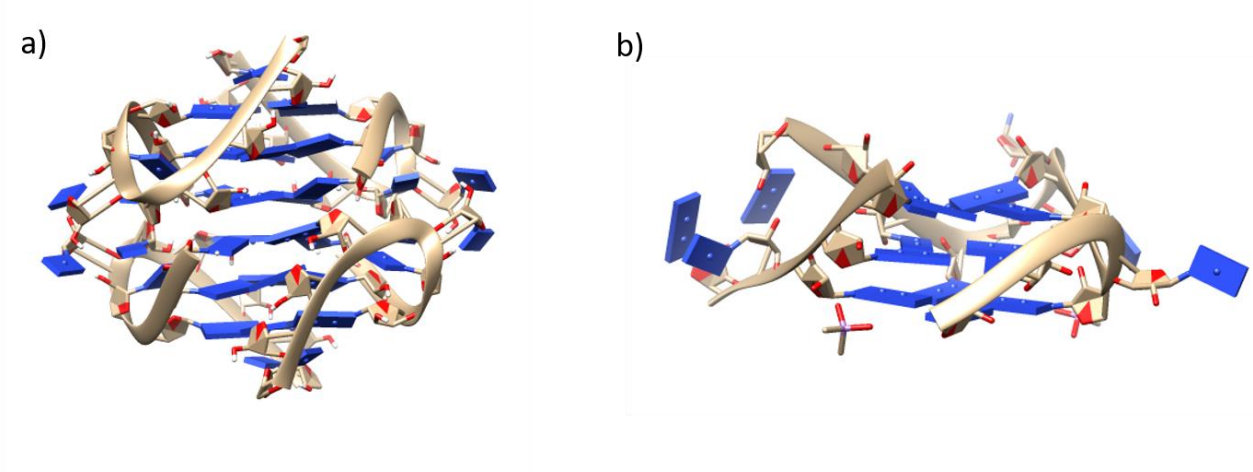


Figure 1.7. The PDB structures of RNA G-quadruplexes a) TERRA (PDB ID: 2M18), b) *n-RAS* (PDB ID: 7XSP).

In addition to the UTRs of mRNA, G-quadruplexes located in alternative splicing sites may function as cis-elements involved in post-transcriptional regulation as well. For example, G-quadruplexes forming within the sixth intron of the *hTERT* gene act as intron splicing silencing elements, leading to a reduction in splicing efficiency.[30]

Table 1.2: List of well explored RNA G-quadruplexes

G-quadruplex	**Sequence (5'-3')**	**Nature**
Telomeric RNA G4 (TERRA)	$G_3U_2AG_3U_2AG_3U_2AG_3U_2A$	Parallel
NRAS	$G_3AG_4CG_3UCUG_3$	Parallel
FMR1	$(CGG)_n$	Parallel
BCL-2	$G^5CCGUG_4UG_2GAGCUG_4$	Parallel
VEGF	$G_2AG_2AG_5AG_2AG_2A$	Parallel

Owing to their relevance in the cellular context, both RNA and DNA G-quadruplexes have emerged as structures of interest and are active areas of research, with investigations focused to understand their biological functions, potential applications and potential in therapeutics in fields like chemical biology and medicinal chemistry.

1.6.G-quadruplex –Functions and significance

The formation of G-quadruplexes in different regions of the DNA or RNA is implicated in critical functions of the cells (Figure 1.8). The formation of G-quadruplexes can regulate access of telomerase enzyme to telomeres, the ends of chromosomes. Telomerase, a complex consisting of non-coding RNA and reverse transcriptase, plays an important role in extending the 3' ends of chromosomes, particularly in cancer cells, stem cells, and germline cells, to prevent telomere shortening and maintain genome stability. The presence of telomeric DNA G4s hinders telomerase access to the 3' end of the telomere, thereby inhibiting telomere extension.[31] In addition, G4s have the capacity to interfere with various nucleic acid functions, including DNA replication, transcription, and translation. Moreover, there are proteins capable of resolving G4 structures within the cell to regulate the folding and unfolding of G-quadruplexes, which are formed transiently. Many of these G4-resolving proteins belong to well-known helicase families, such as the RecQ-like helicases and the DEAD box or DEAH box helicase families.

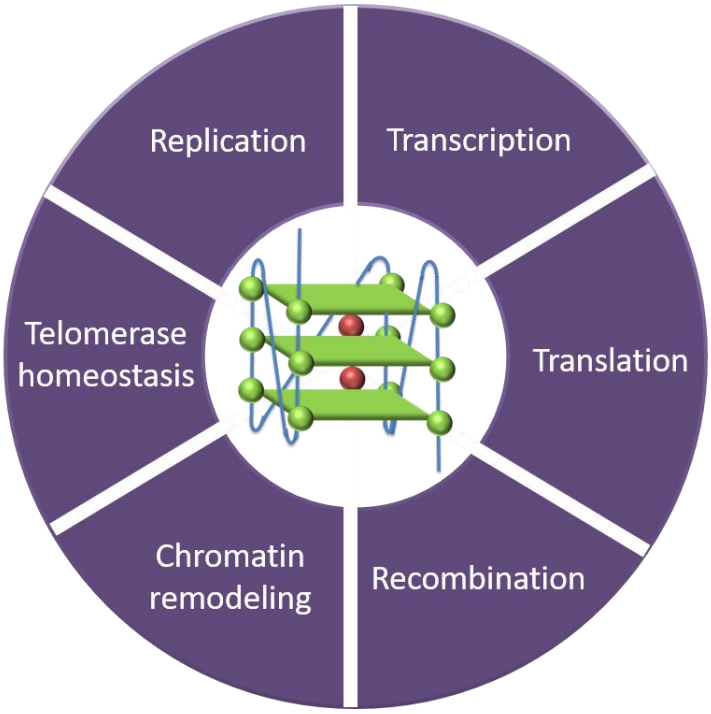


Figure 1.8. A schematic representation of G-quadruplexes involved in different biological processes.

1.7.Regulation of G4 dynamics in cells and disease physiology

Inside living cells, G4 structures are unwound by DNA/RNA helicases, and the resultant single-stranded DNA and RNA may be kept in this unfolded state by proteins that specifically bind to single-stranded nucleic acids. If the G4s are not resolved, it may lead to excessive G4 formation which can have detrimental effects on various biological processes including replication, transcription and translation. The class of helicases that exhibit a preference for unwinding G4 structures are referred to as G4 helicases. These helicases primarily target DNA G4s or RNA G4s, and approximately 20 G4 helicases have been identified to date, including DHX36, WRN, BLM, FANCJ, RTEL1, ATRX, eIF4A, among others.[32]

For instance, the Pif1 helicase in *Saccharomyces cerevisiae* is known to facilitate DNA replication by unwinding DNA-G4s. G4 helicases play a crucial role in replication, and any dysfunction in these helicases can lead to genome instability. Deficiencies in BLM and WRN, which are DNA G4 helicases, result in autosomal recessive syndromes known as Bloom syndrome and Werner syndrome, respectively.[33]

In contrast to G4-resolving proteins, G4-forming and stabilizing proteins have also been reported.[34] One of the well-established G4-forming and stabilizing proteins is nucleolin, which plays a role in ribosome biogenesis. Nucleolin promotes the formation of G4 structures in the *c-MYC* promoter region and stabilizes these G4 structures through its RNA-binding domains and the C-terminal arginine-glycine-glycine domain.[35] As a consequence, *c-MYC* transcription is repressed. Nucleolin also binds to G4 structures in the long terminal repeat promoter of the human immunodeficiency virus-1 (HIV-1) genome, leading to the stabilization of these G4 structures and the transcriptional repression of the promoter. These observations suggest that nucleolin-mediated G4 stabilization plays a regulatory role in oncogene expression.

Furthermore, excessive amplification of G4-forming sequences, such as the hexanucleotide repeat GGGGCC expansion in the C9orf72 gene, results in the accumulation of G4 structures in both DNA and RNA. This accumulation leads to abortive transcription and is believed to contribute to the pathogenesis of amyotrophic lateral

sclerosis and frontotemporal dementia. Mechanistically, G4 structures derived from the C9orf72 hexanucleotide repeat interact with nucleolin, causing mislocalization of nucleolin and nucleolar stress.[36] These findings indicate that the excessive stabilization and accumulation of G4 structures can lead to the development of diseases.

1.8.Role of G-quadruplex in cancer physiology

Cancer is one of the major diseases that pose a serious threat to human life and health. Despite significant breakthroughs anti-cancer therapy and development of drugs, the intricate etiology of cancer continues to present challenges. The quest for innovative anti-cancer medications characterized by elevated selectivity and minimal side effects remains the principal goal in for most drug candidate investigations. Consequently, there is ongoing studies targeted at focused at identifying novel targets and therapeutic strategies.

Numerous conventional chemotherapeutic agents exert their effects on both normal and cancerous cells, as they interact directly with duplex DNA. In this context, the development of novel pharmaceuticals that interact with unique G-quadruplex structures represents an alternative research approach. Such an approach holds the potential to enhance drug selectivity and precision, particularly within distinct genomic regions.

Furthermore, that alterations in the expression of oncogenes are widely recognized as pivotal hallmarks of cancer. Hanahan and coworkers delineated six fundamental cellular and microenvironmental processes as hallmarks of cancer.[37] These processes encompass self-sufficiency for growth signals, insensitivity to anti-growth signals, sustained angiogenesis, limitless replicative potential, evasion of apoptosis, and tissue invasion and metastasis. Notably, each of these categories is associated with a regulation and dysfunction of critical proteins such as such as *c-MYC*, *c-KIT*, and *kRAS* (self-sufficiency), *VEGF-A* (angiogenesis), *BCL-2* (evasion of apoptosis), *hTERT* (limitless replication), *pRb* (insensitivity) and *PDGF-A* (metastasis), etc (Figure 1.9).[38] These genes contain at least one G-quadruplex sequence in their core or proximal promoter regions and their functions have been reported to be influenced by these G4 secondary structures.

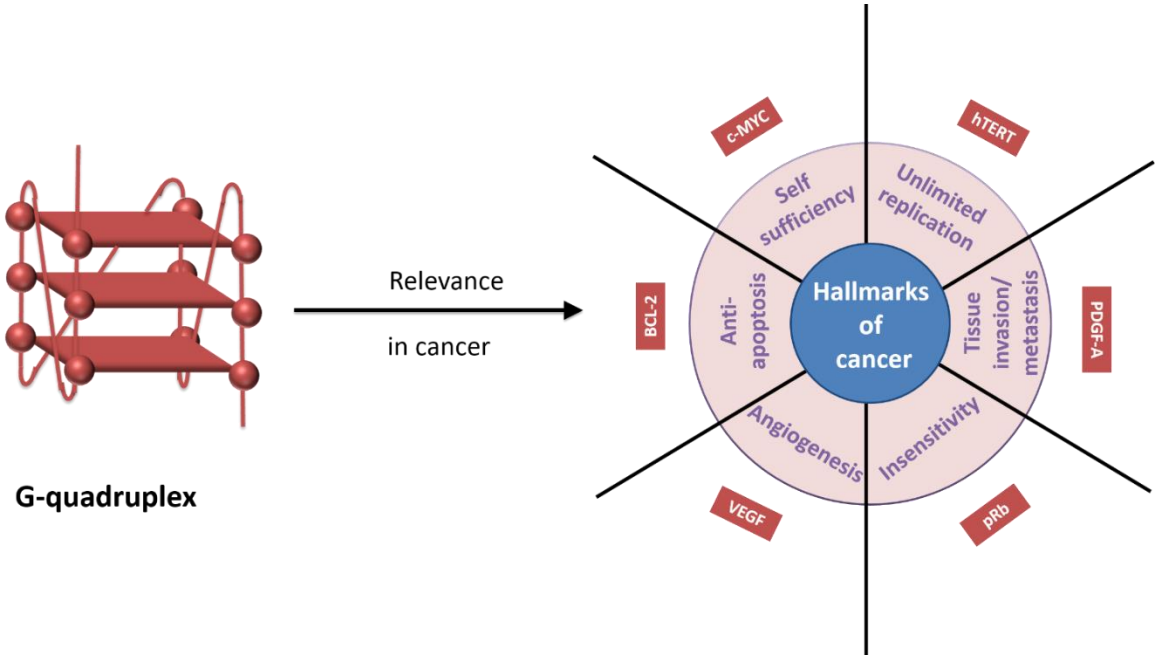


Figure 1.9. A schematic representation of the relevance of G-quadruplex formation and activity in cancer-critical genes which are recognized as hallmarks of cancer

1.9. G-quadruplex as therapeutic targets

I. DNA G-quadruplexes as therapeutic targets

From a historical perspective, DNA was the first defined target for the development of anticancer drugs. However, with the emergence of new molecular targets — such as cell surface receptors and kinases offering greater selectivity for cancer cells, DNA-targeted drug development declined, despite their continued pivotal role in the majority of treatment protocols.

A revival in therapeutics targeted at DNA emerged with the recognition that telomeres can form the unique structures of G-quadruplexes. Telomeres present at the ends of eukaryotic chromosomes, consist of extensive repetitive segments rich in guanine, which may self-assemble to form G-quadruplexes. These telomeric sequences serve as protective

shields against degradation and help maintain genomic stability and serve as substrate for telomerase extension.

G-quadruplexes present in promoter regions of genes (*c-MYC, kRAS* and *KIT, BCL-2, VEGF, RB1, HIF1A, PDGFA, TERT, PDGFRB, etc)* also present attractive therapeutic targets to modulate the function of these genes.[39, 40] Some of these telomeric and promoter G-quadruplexes as therapeutic targets have been discussed in more details.

i. Telomeric G-quadruplex

In normal cellular physiology, telomeres consist of single-stranded overhang regions containing repetitive hexanucleotide DNA sequences of $[TTAGGG]_n$ which serves as a critical mechanism for maintaining chromosomal stability and preventing end-to-end fusion. In somatic cells, telomere length naturally shortens with each cell division, ultimately contributing to cellular aging, senescence, and apoptosis. To counteract this shortening, the enzyme telomerase, a reverse transcriptase, actively maintains telomere length. Telomeres are further protected from degradation by shelterin, a complex composed of six subunits (TRF1, TRF2, POT1, RAP1, TIN2 and TPP1).[41, 42] However, in approximately 85–90% of malignant cells, telomerase activity is significantly upregulated.[43, 44] The increased activity of telomerase enzyme results in the lengthening of telomeres, resulting in cell immortality. Thus, by adding G-rich repeats to the ends of telomeres, telomerase, plays a central role in the process of tumorigenesis, as it induces cellular immortalization.

The highest concentration of potential G4s in eukaryotic chromosomes are the telomeres with multiple repeats of [TTAGGG]. Thus, at the 3' end of the telomeric G-rich sequence, monomeric or multimeric G-quadruplex structures can form. Small molecules have the capacity to stabilize these G4 structures, indirectly inhibiting telomerase activity (Figure 1.10). This inhibition arises from the disruption of the 3' single-stranded telomeric DNA, which is essential for hybridization with the RNA subunit of the telomerase enzyme. This hybridization is a prerequisite for the addition of [TTAGGG] repeats by the telomerase reverse transcriptase. Consequently, the hexanucleotide repeats within telomeres represent a promising target in anticancer therapy. As a result, the strategy of

stabilizing telomeric G-quadruplexes using small molecules has emerged as an attractive approach in the development of anticancer drugs.[45, 46]

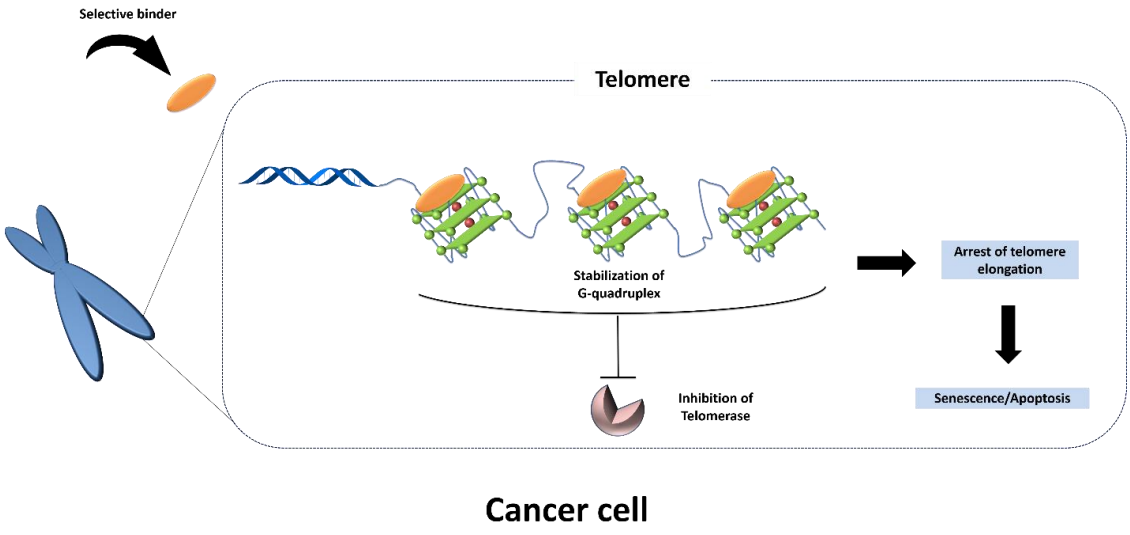


Figure 1.10. Schematic representation of the formation and stabilization of telomeric G-quadruplexes by therapeutic agents leading to arrest of telomeric elongation process and resulting in senescence.

ii. In promoter oncogenes

G-quadruplexes are highly enriched in the promoter regions of genes. In particular, the promoter regions of oncogenes have been reported to contain G-rich regions capable of adopting the folded G4 structure which impedes the recruitment and activity of DNA and RNA polymerase resulting in the stalling of these critical processes in the cell.[47] The inhibition of gene expression is higher for ligand mediated promoter G4 stabilization which further stabilizes the transiently formed G4 structures (Figure 1.11). G4 structures present in promoter regions and their activity has been discussed in details below:

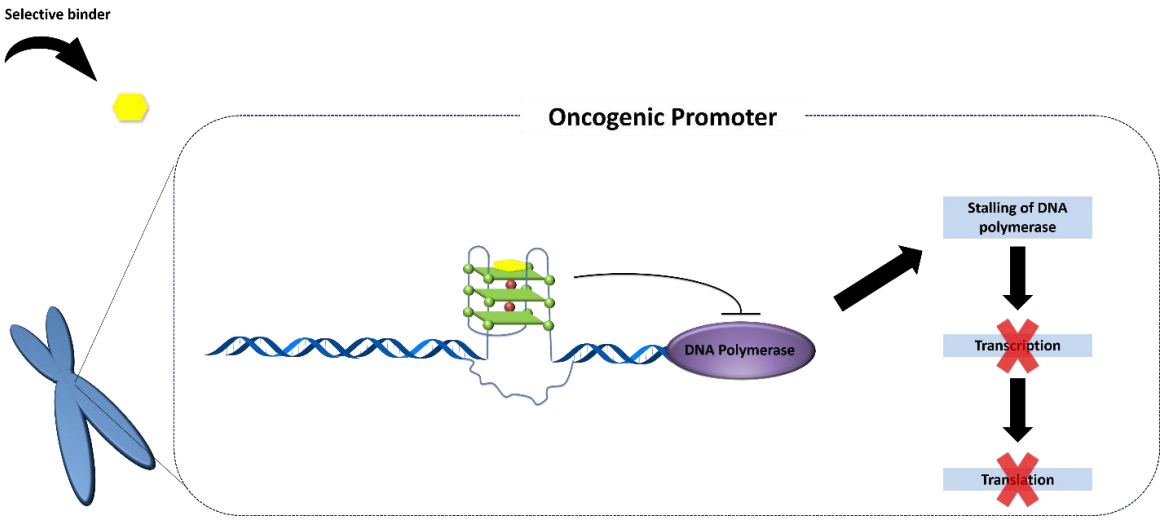


Figure 1.11. Schematic representation of the formation and stabilization of G-quadruplexes formed in promoter regions of oncogenes by therapeutic agents. This stabilization impedes polymerase activity and downregulates the processes of transcription and translation in cells.

a. c-MYC G-quadruplex

c-MYC encodes a versatile transcription factor that can function both as a transcriptional activator for certain genes involved in cell proliferation and as a repressor for others associated with growth arrest. This results in a diverse array of *c-MYC*-responsive genes that collectively regulate critical cellular processes, including metabolic alterations, cell proliferation, and the development of metastatis. In tumor cells, *c-MYC* protein function is typically increased due to the activation of upstream oncogenic pathways. Given that *c-MYC* is overexpressed in a variety of human malignancies, particularly in approximately 80% of solid tumors, suppressing the activity of this gene represents an effective strategy in cancer therapy.[48] However, targeting the *c-MYC* protein with small molecules has been traditionally considered challenging due to its brief half-life, large size, and lack of a well-defined structural conformation.

Within the *c-MYC* promoter region lies the nuclease hypersensitive element (NHE) III1, positioned approximately -142 to -115 base pairs upstream of the P1 promoter. This

region contains a G-quadruplex-forming sequence. The G-quadruplex present in the NHE IIIof *c-MYC* promoter was the first among many to be discovered in the prmoter regions of oncogenes. Ligands capable of interacting with G-quadruplex structures may contribute to the suppression of downstream *c-MYC* expression by stabilizing the G-quadruplex.[49, 50] Over the past decades, there has been a focus on studying G4-interacting ligands targeting *c-MYC* with the goal of developing potential antitumor therapies.

To investigate whether G4 ligands directly affect the *c-MYC* G-quadruplex in cells, the Brooks group devised an exon-specific RT-qPCR assay employing two pairs of Burkitt's lymphoma cell lines, CA46 and RAJI . This assay takes advantage of the unique reciprocal translocation of the *c-MYC* gene locus in these two cell lines. In CA46, exon 1 is separated from exons 2 and 3 due to the translocation, while in RAJI, all three exons and the PQS remain in tandem following the translocation. Consequently, when using primers specific for exons 1 and 2 in a qPCR, the *c-MYC* G-quadruplex-mediated control by G4 ligands can be demonstrated in a cellular context. Indeed, a G4-interacting ligand known as **GQC-05** was found to directly suppress gene expression in a manner mediated by the *c-MYC* G-quadruplex using this assay.[51] This methodology has since become a standard approach for investigating the intracellular activities of small molecules targeting the *c-MYC* G-quadruplex.

b. BCl-2 G-quadruplex

The *BCL-2* gene plays key role in regulation of apoptosis, with its encoded product located on the cytoplasmic side of the mitochondrial outer membrane, where it exerts control over the mobility of apoptosis-inducing proteins by regulating mitochondrial membrane permeability.[52] Elevated levels of *BCL-2* have been associated with abnormal tumor growth in various human diseases, particularly solid tumors such as lymphomas, non-small-cell lung cancer, myeloma, and melanoma, making it an excellent target for cancer therapy. Multiple strategies have been explored to downregulate *BCL-2* expression in cancer cells for therapeutic purposes, including the use of small molecules to disrupt protein-protein interactions, peptidomimetics and antisense oligonucleotides. Notably,

BCL-2 overexpression is considered a primary contributor to chemoresistance, particularly in lymphocytic cancers.[53]

Amplification and translocation of *BCL-2* genes are common mechanisms leading to its overexpression in human cancer cells. The human *BCL-2* gene consists of P1 and P2 promoters with multiple transcription start sites. While the primary transcriptional regulation is less dependent on a TATA-box in promoter 2, the P1 promoter, situated 1386–1423 nucleotides upstream of the translation start site, has been extensively implicated in the control of *BCL-2* transcription. A GC-rich element exists 1490–1451 nucleotides upstream of the P1 promoter. The presence of multiple G-quadruplex structures in this region has been thoroughly elucidated by researchers.[54] These findings have prompted investigations into the development of cancer therapies by stabilizing the *BCL-2* G-quadruplex structures formed in the regulatory elements and consequently reducing promoter activity, similar to the approach taken with small molecules targeting *c-MYC* G-quadruplexes.

c. *c-KIT* G-quadruplex

The *c-KIT* proto-oncogene encodes a receptor tyrosine kinase that is activated by binding to dimerized stem cell factors (SCF), subsequently initiating processes like proliferation, differentiation, and survival in hematopoietic precursor cells. Dysregulation of the Kit receptor due to overexpression or mutations has been implicated in various diseases, including gastrointestinal stromal tumors (GIST), mastocytosis, and acute myelogenous leukemia (AML).[55] While the kinase inhibitor Imatinib (Glivec) has been developed as an FDA-approved drug for GIST, long-term exposure often leads to secondary mutations at exons 13, 14, or 17, encoding tyrosine kinase domains. Notably, resistance to the drug resulting from mutations at exon 17 significantly diminishes its therapeutic effectiveness. Therefore, a strategy to suppress *c-KIT* expression is highly desirable.

The human *c-KIT* promoter lacks both TATA and CCAT boxes. Instead, the region located within 200 base pairs upstream from the transcription start site (TSS) is rich in GC

content and involves several transcription factors, such as MAZ in human normal fibroblasts and SP1 in hematopoietic cells and carcinomas. Within this region, two well-defined G-quadruplex structures have been identified, to play a role in the regulation of *c-KIT* gene transcription.[56-58] This insight has accelerated the development of ligands with a preference for *c-KIT* G4 structures. Modulating such structural dynamics with small molecules has proven effective in suppressing gene expression and inducing apoptosis. The feasibility of targeting the *c-KIT* promoter G4 with small molecules was further validated in several studies involving in vitro, in cellulo and in vivo techniques.

d. KRAS G-quadruplex

The *RAS* gene family, which includes *HRAS, NRAS*, and *KRAS*, was initially identified as driver oncogenes in human tumors and has long been recognized as a crucial therapeutic target. *KRAS* mutation is one of the most potent driver mutations in cancers such as pancreatic, colorectal, and lung cancers. Thus targeting of active *KRAS* with small molecules was initially seen as a promising strategy for combatting *KRAS* mutant tumors.[59] However, this approach has faced significant challenges and remains unsuccessful. G4-mediated promoter targeting has also been explored for suppressing overexpression of *KRAS*. The *KRAS* proximal promoter region contains numerous G-rich sequences, and multiple transcription factors interact with a G4 structure formed within this region.[60] A polypurine G-rich element, situated approximately 300 to 100 nucleotides upstream of the exon 0/intron 1 boundary in both murine and human genomes, is evidenced to contribute to promoter activity and includes several potential G4-forming sequences.

Other G4 structures formed within putative regulatory elements in the promoters of cancer-related genes have also been reported as potential targets for ligands interacting with G4 structures. These genes include *PDGFR-β, PDGF-A, STAT3, FGFR2*, among others.[61] For example, **GSA11129**, an agent capable of interacting with a G4 structure in the PDGFR-β gene promoter and shifting the equilibrium toward G4 formation, was found to reduce transcription levels, inhibit *PDGF-β*-induced cell proliferation, and impede migration. Similarly, the G-rich element in the proximal promoter of the *PDGF-A* gene forms a stable

G4 structure, even within a duplex DNA. Treatment with **TMPyP4**, a compound targeting the *PDGF-A* G4, reduced basal promoter activity, suggesting that ligands specific to this G4 may hold promise as a potential cancer therapy for gliomas, astrocytomas and sarcomas.

II. RNA G-quadruplexes as therapeutic targets

In addition to DNA G-quadruplexes playing important roles for the cellular functions, RNA G-quadruplexes also play pivotal roles in various critical cellular processes, encompassing mRNA translation and localization, alternative splicing, protein interactions and telomere maintenance. The exploration of RNA G-quadruplexes began with the investigation of the 5'-end of the HIV gag gene. A 127-nucleotide RNA fragment derived from *HIV-1* was observed to spontaneously dimerize, forming an interstrand G-quadruplex structure, which significantly impacted the viral life cycle.[62] Consequently, RNA G-quadruplex structures emerged as promising targets for antiviral interventions. Subsequently, the role of human telomeric DNA, transcribed into TERRA RNAs featuring repetitive UUAGGG motifs, which have the capacity to adopt stable G-quadruplex structures was elucidated.[23] These telomeric RNA G-quadruplexes play a crucial role in regulating telomere length and suppressing telomerase activity.

Moreover, RNA G-quadruplexes situated in the 5' UTR of *LRP5* and *FXR1* mRNAs were found to enhance alternative polyadenylation efficiency. Similarly, RNA G-quadruplexes were identified as mediators of alternative splicing in genes like *p53* and *PAX9*. In light of the vital biological functions attributed to RNA G-quadruplexes, they have also become attractive targets for small molecules (Figure 1.12).[63] Consequently, threrapeutic strategies to develop effective techniques for targeting RNA G-quadruplexes to elucidate their biological significance to further expand their applications has recently gained momentum. In this line of investigation, several small molecules have been designed to better understand and modulate the distribution and functional roles of RNA G-quadruplexes.

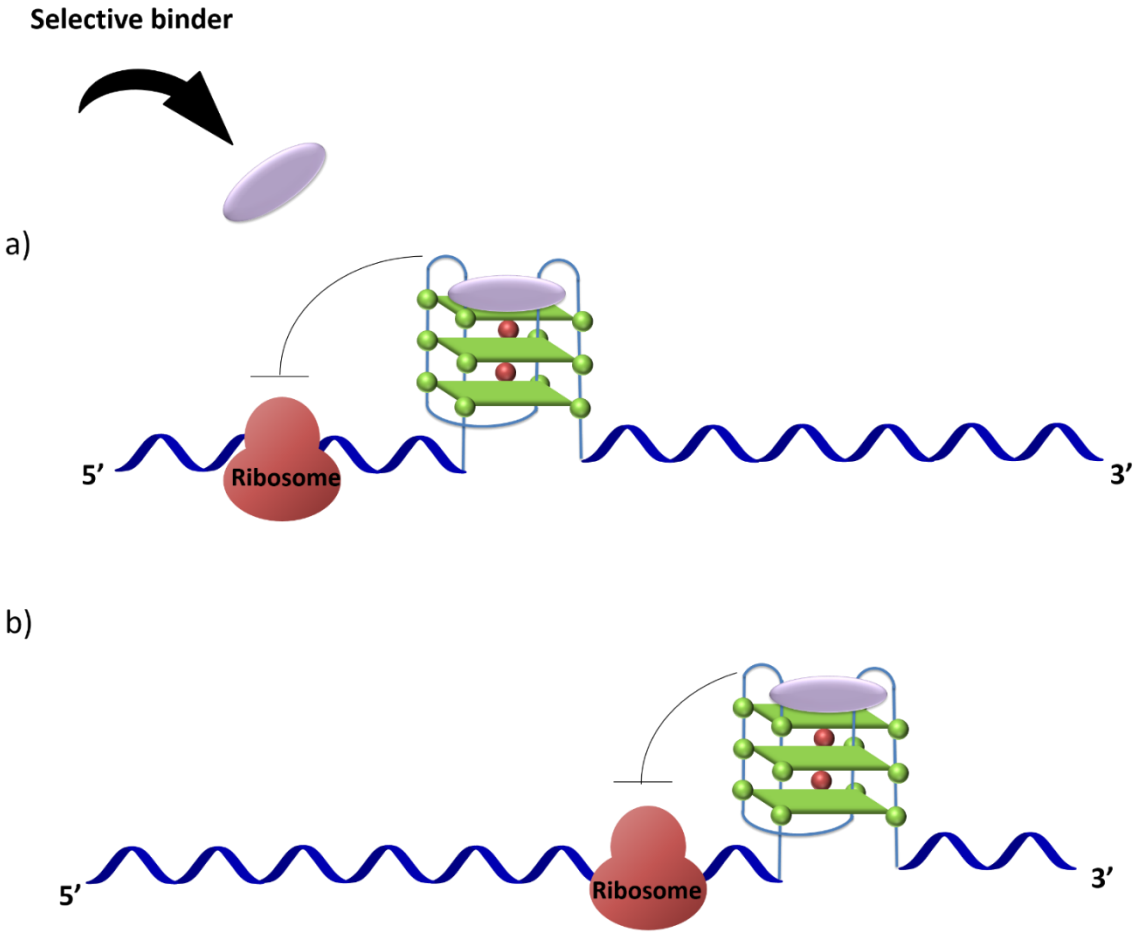


Figure 1.12. The formation and stabilization of G-quadruplexes formed in untranslated regions (3' UTR and 3' UTR) of mRNA inhibits the activity of ribosome and downregulates the process of translation in cells.

1.10. G-quadruplex interacting small molecules

Over the past few decades, there has been a significant surge of interest in biologically relevant non-canonical nucleic acid structures, such as G-quadruplexes, owing to their distinctive structural characteristics and crucial roles in biological processes. Consequently, these secondary structures have emerged as highly attractive targets for drug development. G-quadruplexes have been unequivocally implicated in regulating fundamental cellular activities, including DNA replication, DNA damage repair, as well as the transcription and translation of oncogenes and other genes associated with cancer.

Consequently, G-quadruplexes as therapeutic targets has evolved into a promising and innovative strategy for combating cancer. Diverse classes of small molecules have been strategically designed, synthesized, and evaluated as potential anti-cancer agents, with some directly binding to G-quadruplexes and others disrupting the interactions between G-quadruplex structures and their associated binding proteins.[64]

G-quadruplexes exhibit a diverse array of biological functions, spanning telomere maintenance, transcriptional regulation, translation modulation, replication control, DNA damage response, genome rearrangement, and epigenetic regulation. Consequently, the design of small molecules capable of interacting with G-quadruplexes has emerged as a promising avenue for the discovery of novel compounds with potential anti-tumor activities. Over the past decades, numerous small molecules that interact with either DNA G-quadruplexes or RNA G-quadruplexes have been reported as promising anti-cancer drug candidates in preclinical studies. These molecules exert have demonstrated diverse effects on cells due to the modulating the biological functions of G-quadruplexes:

1. Inhibition of telomerase activity and the disruption of unlimited tumor cell proliferation result from the stabilization of G-quadruplexes at telomeres.
2. Suppression of oncogene expression by stabilizing DNA G-quadruplexes within promoters.
3. Decrease in translation efficiency by stabilizing RNA G-quadruplexes
4. Blocking of replication forks and induction of single-stranded DNA gaps or breaks in tumor cells induced by small molecule mediated interactions with G-quadruplexes.

To date, extensive research efforts have explored a wide array of ligands designed to target G-quadruplex structures, with many of these compounds documented in the G4 Ligands Database 2.2 (G4LDB, http://www.g4ldb.com/).[65] Figure 1.13 highlights some noteworthy G-quadruplex ligands. These well-established G-quadruplex ligands typically share fundamental structural attributes which render them as highly selective molecules designed to interact with G-quadruplex structures over double stranded or single stranded nucleic acids. These key structural features include:

i. Containing an aromatic core that facilitates π–π stacking interactions with the planar G-tetrads which is a unique feature of G-quadruplexes.

ii. Incorporating one or more positively charged moieties capable of interacting with the negatively charged phosphate groups along the DNA or RNA backbone within the grooves and loops of G-quadruplexes *via* electrostatic interactions.

Based on these characteristics, various classes of small molecules have been explored (Figure 1.13). Earliest reported G-quadruplex binding ligands include bisamido-anthraquinone, **TMPyP4** and **PIPER** which laid the initial foundation for subsequent designs and studies.[44] Progressively modifications of the core and the sidechains of the initial ligands: from anthraquinone to fluorenone, then acridone and acridine led to the design of **BRACO-19**. **BRACO-19** showed interaction with G4s, antiproliferation of cancer cells, and strong potency for telomerase inhibition.[66, 67] The tetracationic porphyrin, **TMPyP4** has been extensively studied by leading G-quadruplex research groups and is the pivotal example of N-Methylated aromatic G-quadruplex ligands.[22, 68, 69] Several structurally-related ligands have been evaluated over the past years: the porphyrin **TQMP68** and the porphyrazine **3,4-TMPyPz,** etc which have shown G4 binding properties.[70] The perylene diimide **PIPER** has shown moderate telomerase inhibition as well.[71] An important breakthrough in the porphyrin series came with the design of a diselenosapphyrin **Se2SAP**, with an expanded porphyrin core. **Se2SAP** showed strong selectivity for quadruplex-DNA (~50-fold) over duplex-DNA and was found to discriminate among the various types of DNA G-quadruplexes.[72]

Beyond the macrocyclic porphyrins, several small molecules have been reported with exceptional properties. Among the first was N-methylated pentacyclic acridinium, **RHPS4** reported in 2000 has been found to bind to the telomeric G4 with high affinity.[73] **Pyridostatin** (**PDS**) was designed with emphasis on electron rich aromatic surface and was found to increase telomere fragility in BRCA2-deficient cells *via* stabilization of the G-quadruplex in the telomeric region, and thus reduce proliferation of homologous recombination-defective cells by inducing double-strand breaks.

Potential anti-cancer activities has been demonstrated by pyridodicarboxamide (**PDC**) core containing series such as **307A** and **360A**. **360A** exhibit a high degree of

quadruplex-stabilization, quadruplex-over duplex selectivity, telomerase inhibition, growth arrest and apoptosis in cancer cells.[74] Another extension of this family of ligands was achieved by the synthesis of phenanthroline analogues **Phen-DC** that show a perfect geometrical match with a G- tetrad and a higher selectivity than that of **telomestatin**, further confirming the potential of the potential of bisquinolinium compounds as selective G4 interacting ligands.[75] In the category of neutral ligands **telomestatin** has emerged as a class of its own. This natural molecule has been isolated from *Streptomyce*s *annulatus* in 2001 and has been subsequently extensively studied since it appears to be one of the most selective binders of G-quadruplex.[76]

Quarfloxin is a fluoroquinolone derivative, and the first G-quadruplex interactive agent to enter human phase II clinical trials.[77] It has been shown to bind to G-quadruplex DNA and selectively disrupt interaction of rDNA G-quadruplexes with the nucleolin protein, resulting in Pol I transcription and inducing apoptosis in cancer cells. To date, about 3000 different G4-targeted ligands have been identified. Ongoing research efforts are based on rational design and synthesis of small molecules with optimum functionalizations based on these key ligands to achieve selective recognition and anti-cancer activities.

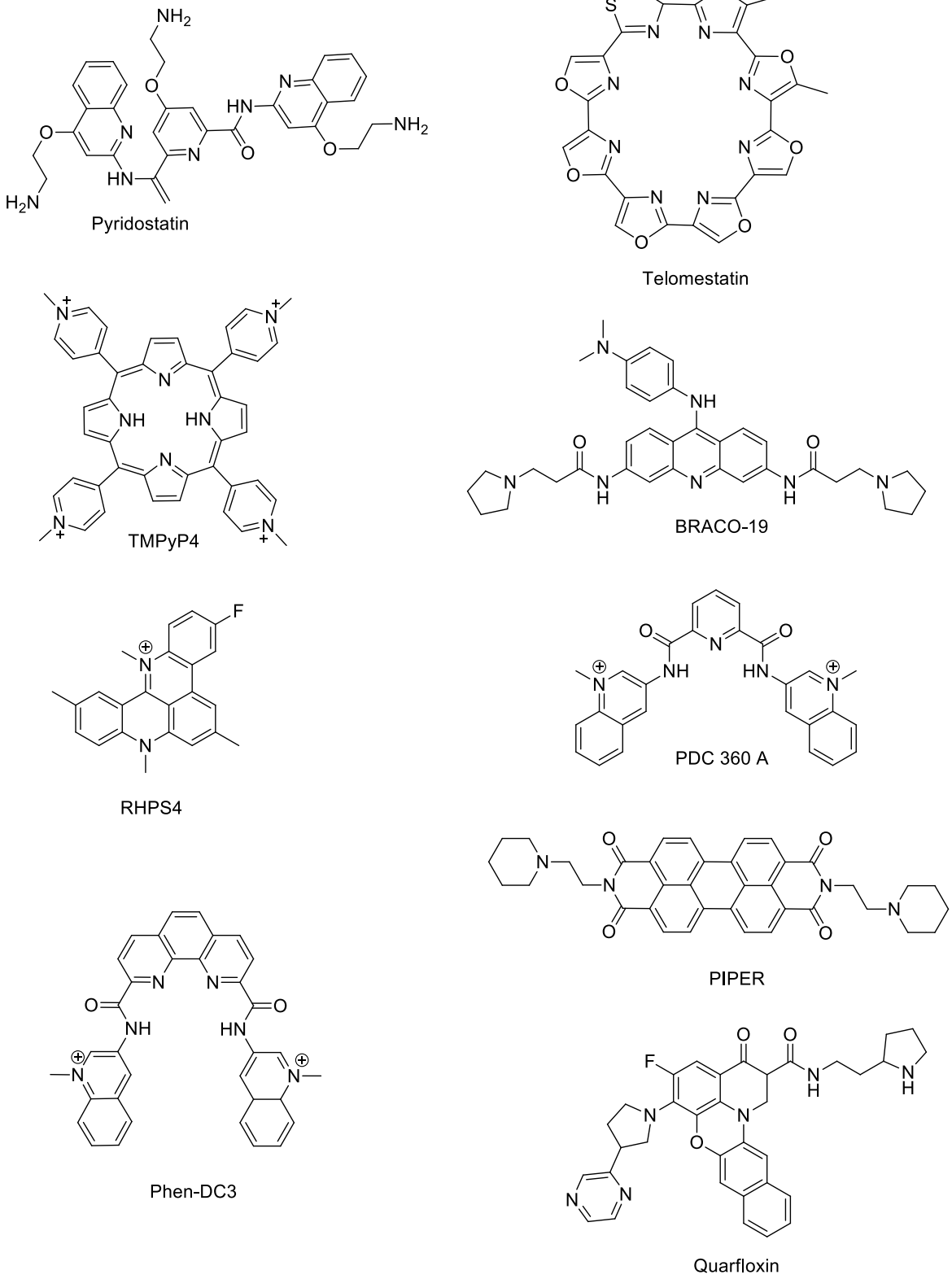


Figure 1.13. The chemical structure of G-quadruplex interacting small molecules.

1.11. Exploration/predicting of novel G-quadruplexes in important genes

Computational methods have been employed to predict and detect G-quadruplexes (G4s) in DNA sequences. Early algorithms focused on identifying sequences containing the pattern of $G_{3–5}N_{1–7}G_{3–5}N_{1–7}G_{3–5}N_{1–7}G_{3–5}$, with four runs of guanines separated by variable loops which indicated the potential to adopt G4 structures (Figure 1.14).[13, 19] These initial computational analyses suggested that the human genome might contain more than 300,000 sequences capable of forming G4 structures.

These computational tools have proven valuable in pinpointing potential G4-forming regions within genes and have uncovered an enrichment of G4-forming sequences in genomic areas associated with gene regulation, specific cellular functions, and various diseases[21]. However, the early search algorithms had limitations; which could not account for structural variations, such as G4s with longer loops, bulges, or mismatches, nor could they identify G4s composed of two tetrads. Additionally, they did not consider the significance of flanking sequences. More recent computational tools have addressed some of these limitations. They incorporate factors like structural variants, consider the effects of flanking sequences based on metrics like g-fraction and g-skew, and explore the possibilities of higher-order G4 assemblies in the genome.[78] Furthermore, machine learning approaches have recently been employed to identify sequences G4s likely to form within genomic contexts and those with a propensity to fold in RNA sequences.[79]

Additionally, analysis of the complete human genome computationally has revealed that G-quadruplexes are not randomly distributed. G-rich sequences are highly enriched at the 5'-end of genes (5'-untranslated region, first exon and first intron compared to the coding regions, thereby suggesting proximity to the transcription start site (TSS) within genes.[80] Surprisingly, nearly half of all known genes possess a potential G-quadruplex structure in their promoter regions, strategically positioned for potential involvement in gene regulation. These G-quadruplexes exhibit greater thermodynamic stability compared to typical structures. In contrast, RNA has demonstrated an even greater propensity to form G-quadruplex structures than DNA, and this propensity is enhanced by the fact that RNA is typically single-stranded in biological systems. This single-stranded

nature of RNA results in higher probability of G-quadruplex formation as it does not compete with duplex formation, making it more likely to occur. The abundance of sequences in oncogenes predicted to form G4 structures influences gene expression and contributes to genomic instability. Novel computational prediction methods are routinely employed to discover new sequences capable of adopting the higher-order structure of G-quadruplex and corroborated by biophysical, structural and cellular investigations.[81]

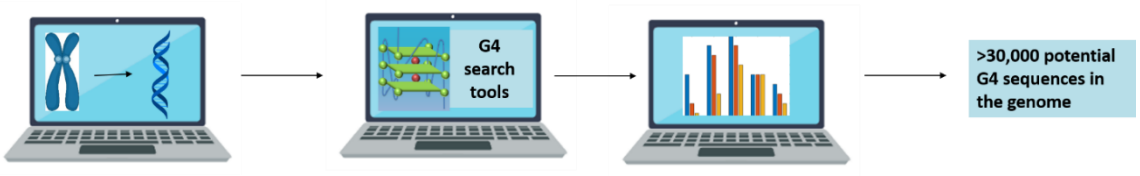


Figure 1.14. Computational algorithms are used to search for putative G-quadruplex forming sequences.

Chapter 2

Regulating the telomeric elongation by stabilizing the telomeric G-quadruplex using small molecules

Introduction

In eukaryotic somatic cells, the telomere region undergoes progressive shortening during replication, consequently initiating cellular senescence and apoptotic pathways.[31] However, in the context of cancer cells, a substantial upregulation of telomerase enzyme activity is observed, facilitating the extension of telomeres. Telomerase exhibits elevated expression in approximately 80-85% of cancer cells and primary tumors, playing a role in maintaining telomere length homeostasis, which contributes to its function as a promoter of tumorigenesis.[82, 83] This mechanism, in turn, results in cancer cell proliferation, evasion of apoptosis and cellular immortality. Telomere contains tandem repeats of G-rich $(TTAGGG)_n$ for several kilobases in the 3' overhang which can easily adopt the G4 structure.[44] The formation of G-quadruplex structures within this single-stranded DNA overhang of telomere can inhibit telomerase activity which extends the telomeric region by using the 3'overhange as substrate. This inhibitory effect of these G-quadruplexes can be further enhanced by specific small molecules that stabilize the G-quadruplex structure.[67, 84, 85] Consequently, G-quadruplexes have garnered significant attention as promising targets for therapeutic interventions.[86]

Investigations into the stabilization of the telomeric G-quadruplexes using small molecules such as **telomestatin**,[45] **pyridostatin**, **RHPS4**,[87] **tetrakis-(N-methyl-4-pyridyl) porphine** (**TMPyP4**), **BRACO-19**,[66] fluoro-indoloquinolines, and berberine has emerged as a promising avenue.[88] The aim is to inhibit telomerase activity and hinder telomere elongation specifically in cancer cells. Despite the identification of numerous molecules with potential, the selective recognition of particular G-quadruplex motifs poses a critical challenge and cancer cell cytotoxicity that constrain the translational progress of these stabilizing ligands into clinical investigations.

Although Quarfloxin has advanced to phase 2 trials, it has been found to have unintended off-target binding, resulting in lack of specificity. Consequently, the quest to devise innovative designs that strike a balance between effective target recognition, permeability, and therapeutic impact is ongoing. Ensuring that induced responses do not escalate into excessive toxicity remains paramount, as overly aggressive interventions can lead to off-target effects, thereby hindering the effectivity of these approaches during

preclinical evaluation. Optimal treatment strategies may involve devising a method to induce graded response that influences related molecular mechanisms, presenting a more tunable and potentially efficacious therapeutic approach.

The indole alkaloid ellipticine was originally isolated in 1959 from *Ochrosia elliptica.*[89] Initially, it exhibited promising anticancer properties and laid the foundation for subsequent investigations. Over the next two decades, extensive research efforts in both the synthesis and biological evaluation of ellipticine and its analogs led to the development of commercial cancer chemotherapy agents.[90] However, the clinical success of these drugs in the pharmaceutical market was hindered by issues related to their toxicity. Nevertheless, research persisted in exploring the indoloquinoxaline system and its derivatives.

Further investigations revealed that **B-220a** exhibited notable antiviral activity against herpes simplex virus type 1 (HSV-1), cytomegalovirus (CMV), and varicella-zoster virus (VZV).[91]. Subsequent developments in a new series of indoloquinoxaline derivatives have demonstrated significant efficacy against autoimmune inflammatory diseases such as multiple sclerosis (MS), rheumatoid arthritis (RA), and other similar conditions.

In this chapter, rationally designed and synthesized library of indoloquinoxaline containing small molecules (**IQ1-9**) with different side chain modifications was studied to ascertain their binding and stabilization potential with different G-quadruplexes (Figure 2.1). This library of ligands was designed in an effort to enhance the DNA-binding properties of the indoloquinoxaline analog family and optimize their anticancer effects. The G4 stabilizing property of these ligands and their potential as anti-cancer agent was investigated. Biophysical investigations were used to screen for potential drug candidate molecules leading to **IQ1** and **IQ2** being found as the most promising binders of the telomeric G-quadruplex. Further investigations into cellular response of the ligands in the cancer cells were carried out. The ligands induce apoptosis *via* intrinsic apoptotic pathway and activating the 3/7 caspase, it also generates ROS as well as induces DNA damage and cell cycle arrest. In addition, it leads to the downregulation of genes related to the telomere integrity and anti-apoptosis in the cancer cells. Effective binding of ligands to target leading to induction of response in mechanisms of the telomere extension allows a greater degree of control to fine tune the activity of telomere.

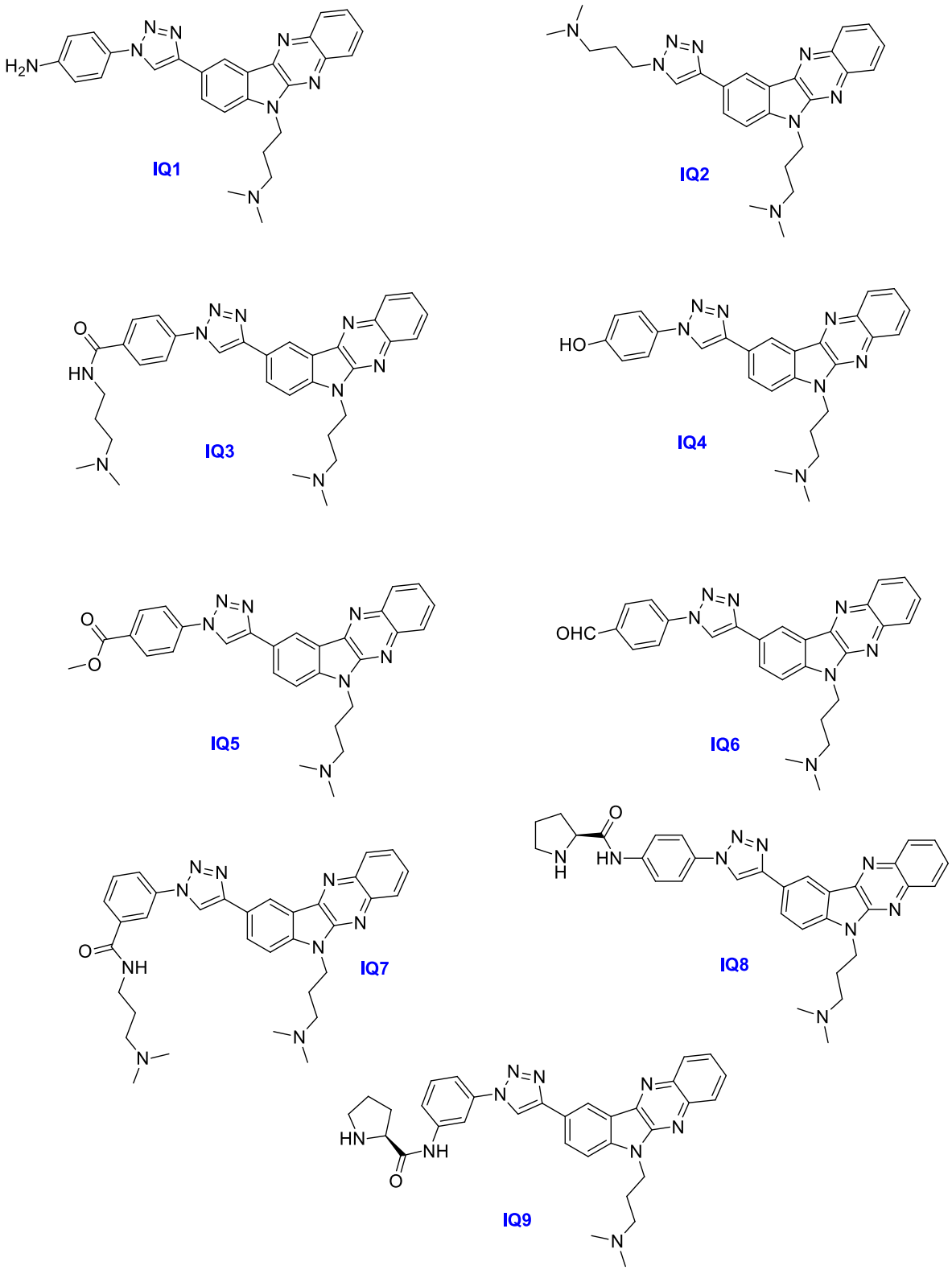


Figure 2.1. The chemical structure of the investigated indoloquinoxaline derivatives (**IQ1-9**) containing different side chain modification.

Results and Discussion

2.1. Indoloquinoxalines exhibit selective binding to telomeric G-quadruplex

FRET Melting assay

Fluorescence resonance energy transfer (FRET) is adistance-dependent process between a donor fluorophore, which when excited, can excite the acceptor molecule with its emission. When these two components are in close proximity, energy is transferred from the donor to the acceptor (Figure 2.2a). FRET, a reliable method allows for real-time monitoring of the structural stability and dynamics of biomacromolecules in solution. It is especially valuable for investigating G-quadruplex (G4) nucleic acids and their interactions with ligands.

Thus, FRET melting assay was performed to determine the stabilization potential of Indoloquinoxaline derivatives**,** containing dimethyl amino propyl side chain (Figure 2.2b). The assay aimed to measure the change in T_m of different dual fluorophore (FAM and TAMRA) tagged G-quadruplexes (*VEGF*, *BCL-2*, *c-KIT 1*, *h-TELO*) and the control double stranded DNA (*ds*DNA) in presence of each ligand at concentration of 1 µM. The concentration with half-maximal denaturation of folded DNA was used to find the ΔT_m values of the G-quadruplexes in presence of the ligands. Maximum ΔT_m (ΔT_m = 23 °C) was observed in case of **IQ3** (ΔT_m = 23.4°C)**, IQ2** (ΔT_m = 23°C)**, IQ7** (ΔT_m = 23.2°C) and **IQ9** (ΔT_m = 23.7 °C) for the telomeric DNA (*h-TELO*).A lower ΔT_m was observed for **IQ1** (ΔT_m = 13 °C) with *h-TELO*. The rest of the ligands showed ΔT_m of ≥ 10 °C. Thus, most of the ligands of the library showed higher stabilization of *h-TELO* G4 with varying degrees of stabilization. In addition, all the ligands of the library showed insignificant stabilization with the control dsDNA. The high degree of change in ΔT_m value of *h-TELO* suggests that **IQ1, IQ2, IQ3 and IQ7** can selectively recognize and stabilize the *h-TELO* G4 structure.

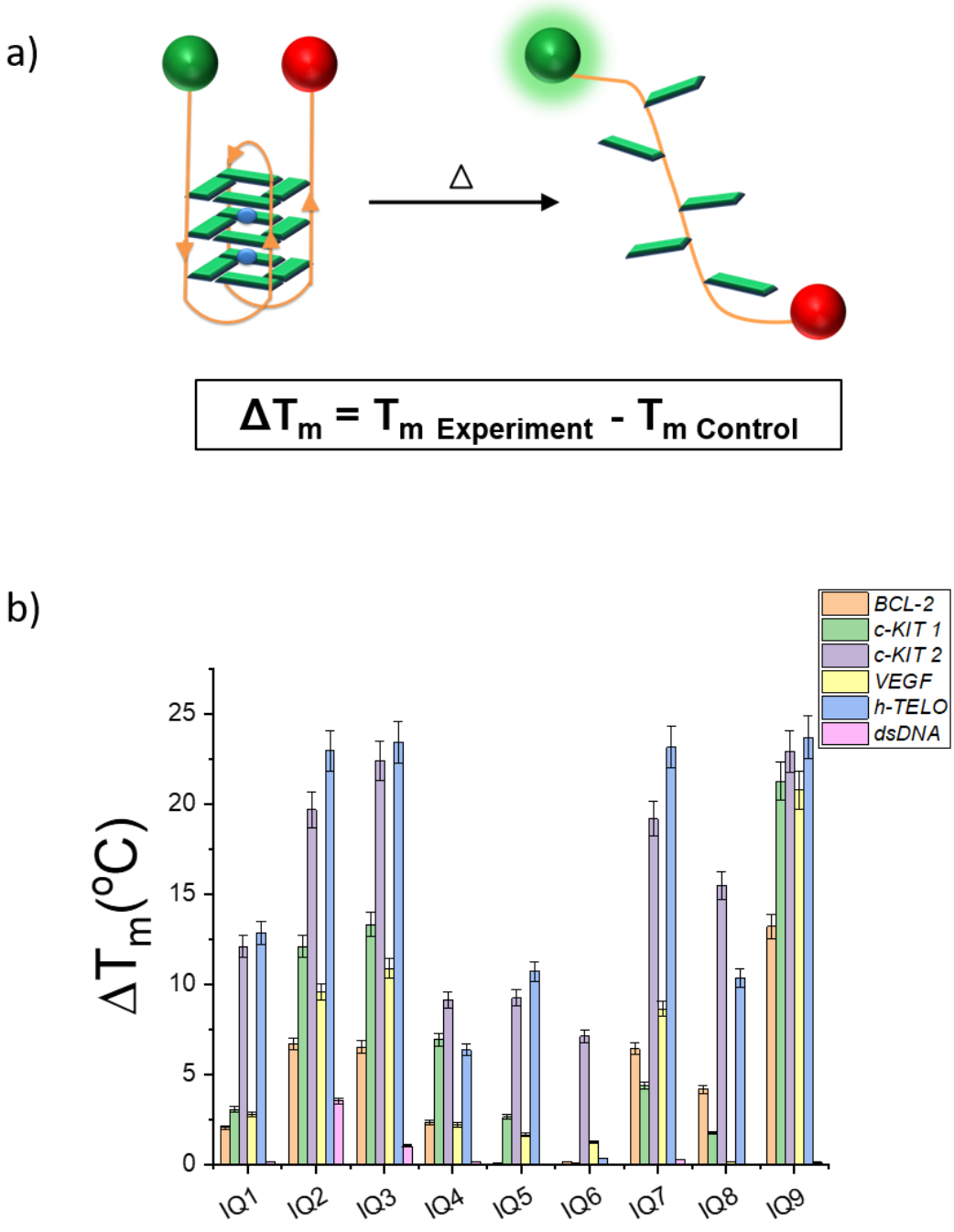


Figure 2.2. a) Principle of FRET melting assay using FAM (donor-green) and TAMRA (acceptor-red) to monitor the melting temperature of G-quadruplex on application of heat, b) FRET melting analyses of indoloquinoxaline ligands (1 μM) with different dual-fluorophore labeled G-quadruplexes (200nM) in 60mM potassium cacodylate buffer (pH 7.4).

Fluorescence Intercalator Displacement assay

The binding affinity of the lead ligands (**IQ1, IQ2, IQ3 and IQ7**) which exhibited maximum stabilization of *h-TELO* G4 was investigated by fluorescence intercalor displacement (FID) using thiazole orange (TO) as the fluorescent intercalator (Figure 2.3). The displacement of DNA-bound TO from DNA can be used as a measure of affinity of a ligand which can displace the bound TO from the G4 and results in the decrease of fluorescence of the TO intercalator. The DNA bound TO (λ_{max} = 501 nm) exhibits high fluorescence (fluorescence maxima = 534 nm) which gradually decreases with incremental additions of the ligand. The concentration of half-maximal displacement (DC_{50}) of TO by the ligands were obtained for all four ligands with the lead target G4 of *h-TELO* and *ds*DNA as a control DNA (Figure 2.4). The DC_{50} values were found for only **IQ1** (3.5 µM) and **IQ2** (1µM) with *h-TELO* G4. **IQ3** and **IQ7** were not able to displace **TO** upto the half-maximal value and the DC_{50} values could not be obtained for them (Figure 1b). On the other hand, none of the ligands have a high affinity for the control *ds*DNA as the intercalator could not be displaced enough to obtain a DC_{50} value. The significantly low DC_{50} value of **IQ2** (1 µM) and the moderately low DC_{50} value of **IQ1** (3.5 µM) suggests high affinity for *h-TELO* G4. **IQ2** showed a high binding interaction with *h-TELO* and **IQ1** also exhibited affinity towards the telomeric G4 sequence. As only **IQ1** and **IQ2** emerged as a potentially selective ligand with high affinity for the telomeric G4, they were taken for further investigations to decipher their effect on the telomere integrity.

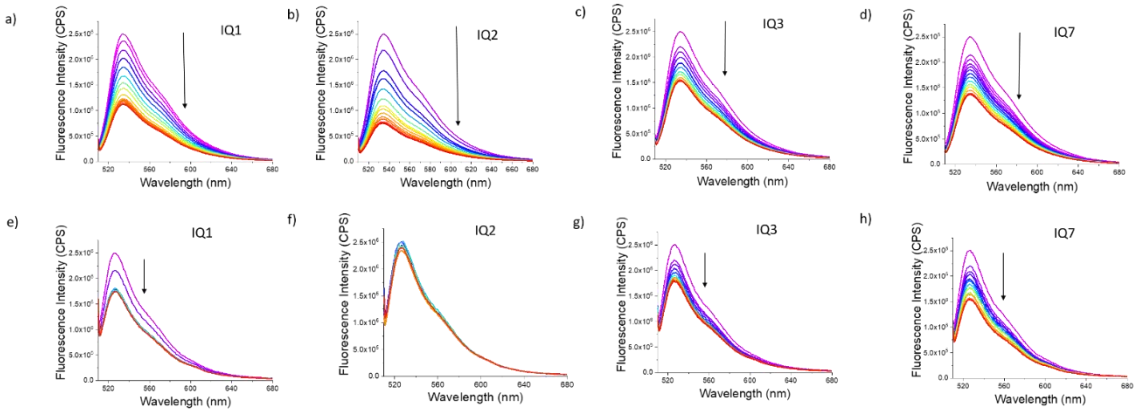


Figure 2.3. Fluorescence displacement assay results showing the DC_{50} values of indoloquinoxaline ligands with a-d) *h-TELO* G-quadruplex and e-h) *ds*DNA. All experiments were performed in 100mM Tris 100mM KCl buffer (pH 7.4).

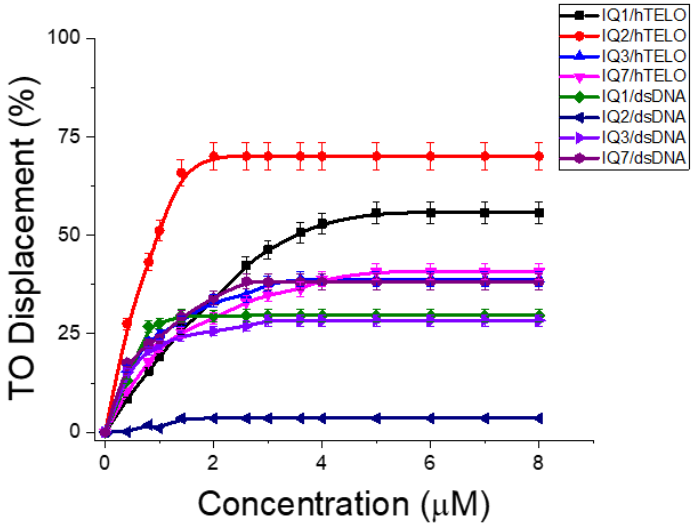


Figure 2.4. Fluorescence Intercalating Displacement assay analyses of Thiazole Orange bound DNA with indoloquinoxaline ligands. All experiments were performed in 100mM Tris 100mM KCl buffer (pH 7.4).

Isothermal Titration Calorimeter experiments provide valuable information about the thermodynamics and binding parameters of drug-target interactions. Thus, ITC was employed to study the real time molecular interactions of the ligands **IQ1** and **IQ2** with different G-quadruplexes and the dsDNA (Figure 2.5 and 2.6). The G-quadruplexes were titrated with a total of 19 injections at regular intervals to monitor the dynamic heat compensation process. The integrated heat plots or isotherms showed that the ligands have binding affinities to specific G-quadruplexes. **IQ2** shows the minimum binding constant (K_d) of 2.7 μM with *h-TELO* where the molecular reaction is exothermic in nature suggesting the highest binding affinity of **IQ2** with *h-TELO* G4. The reaction also demonstrated very low G value of -31.8 kcalmol^{-1} and the reaction stoichiometry of n = 1. **IQ2** showed relatively higher K_d values of 6 μM with *BCL-2* and 4.7 μM with *c-KIT 2* G-quadruplexes whereas satisfactory K_d values were not found with *K-RAS*, *c-MYC*, *VEGF* and the control *ds*DNA. On the other hand, molecular interactions of **IQ1** with the G4s yielded much higher K_d values. The K_d for **IQ1** with *h-TELO* was found to be 4.95 μM whereas with *k-RAS, BCL-2* and *c-KIT 1*, the values were observed to be 5.7, 12.9, and

29.4 µM respectively. The k_d of **IQ1** with the rest of the studied G4s (*c-MYC*, *c-KIT 2* and *VEGF*) and *ds*DNA could not be determined. The collective result of ITC analyses suggest that both **IQ2** and **IQ2** has highest binding affinity with *h-TELO* G-quadruplex where the **IQ2** ligand was found to be a more selective and better binder of the telomeric G4.

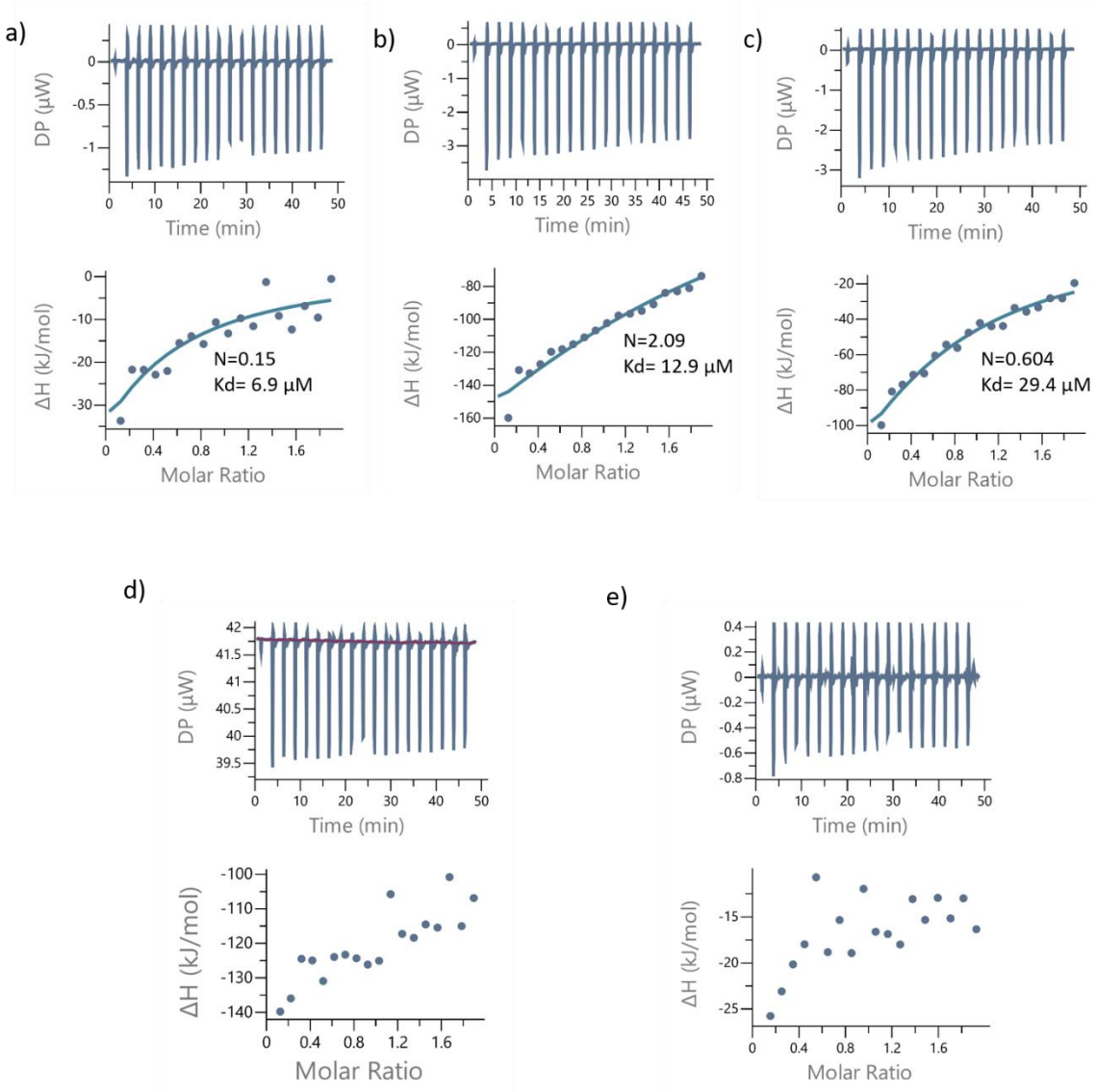


Figure 2.5. ITC heat curves and binding isotherms of a) **IQ1**/*h-TELO,* b) **IQ1**/*BCL-2*, c) **IQ1**/*c-KIT 1*, d) **IQ1**/*VEGF*, e) **IQ1**/*dsDNA*. All experiments were performed in 100mM Tris 100mM KCl buffer (pH 7.4).

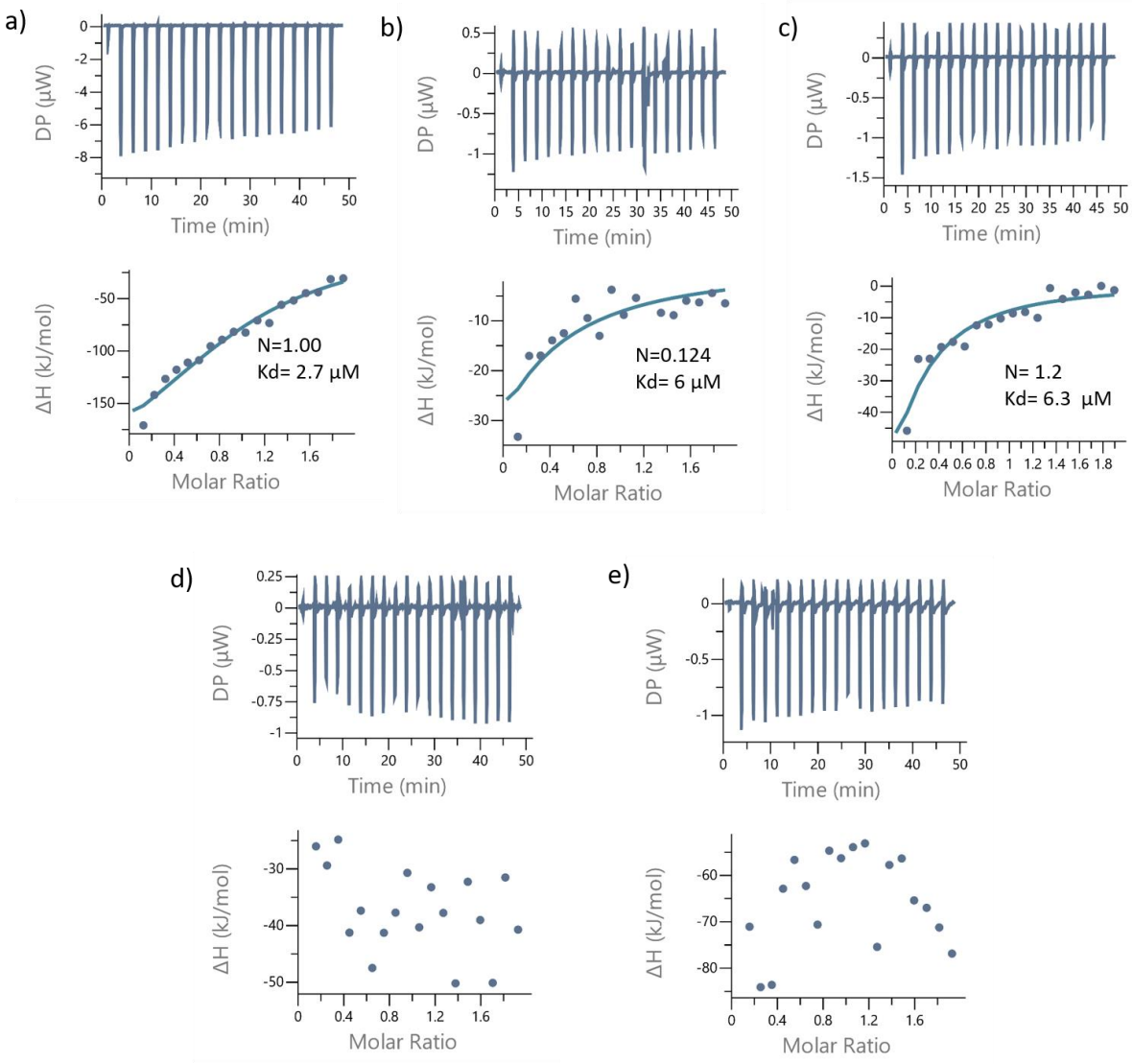


Figure 2.6. ITC heat curves and binding isotherms of a) **IQ2**/*h-TELO*, b) **IQ2**/*BCL-2*, c) **IQ2**/*c-KIT 1*, d) **IQ2**/*VEGF*, e) **IQ2**/*dsDNA*. All experiments were performed in 100mM Tris 100mM KCl buffer (pH 7.4).

Circular Dichroism Spectroscopy

Circular Dichroism spectroscopy was used to confirm the secondary G-quadruplex structure of the telomeric sequence. The hybrid G-4 structure in presence of K^+ displayed the characteristic positive peaks at 262 nm and 290 nm as well as the negative minima at -240 nm. The G4 folded structure was further titrated with increasing concentrations of **IQ1** and **IQ2** to monitor their effects on the folded conformation (Figure 2.7). It was observed that the peak intensity of *h-TELO* did not undergo alterations in presence of the ligands even at high concentrations as 10 equivalents compared to the DNA concentrations. Thus, it was found that both the molecules **IQ1** and **IQ2**did not disrupt or perturb the folded G4 structure.

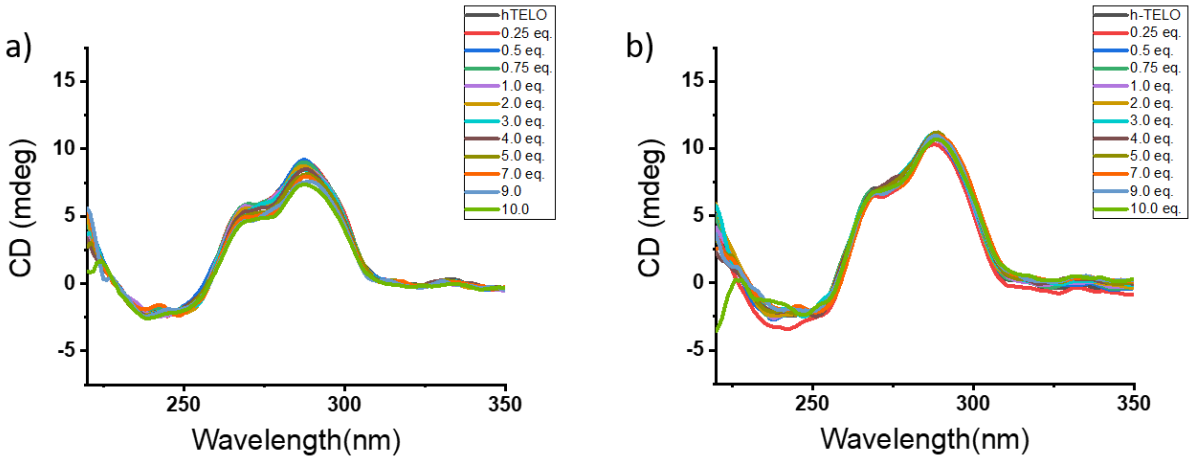


Figure 2.7. Circular dichroism of *h-TELO* G4 (15µM) titrated with increasing concentrations of a) **IQ1**and b) **IQ2**. All experiments were performed in 100mM Tris 100mM KCl buffer (pH 7.4).

2.2.Mode of interaction of indoloquinoxaline derivatives

Nuclear Magnetic Resonance Titration

The interaction of **IQ2** with the h-Telo G-quadruplex structure was investigated by NMR, recording 1D ^{1}H NMR spectra at different [ligand]:[DNA] ratio. The titration was performed in 5% d_6-DMSO/95% H_2O for better solubility of the ligand. The stabilized *h-TELO* sequence Tel24 adopting Hybrid-1 conformation was used for NMR investigations

to reduce dynamical exchange between G4 conformation and gain a better resolution of NMR signals.]

Addition of 0.25 equivalents **IQ2** result in line broadening of the imino signals of the residues G3, G17 and G21 and affects the aromatic signals of G16 and A20 (Figure 2.8). Further addition increases the observed spectral changes with additional moderate line broadening of the imino signals of the residues G5 and G11. Further, impact on the aromatic signals of G4, T18, T19 and A24 as well as spectral changes in the overlapping region of T6 and T12 and T1, G3, A14H2 and G23 can be observed. The results of the NMR titration hint to specific binding of **IQ2** with a 1:1 ratio [Ligand:DNA]. However, additional interaction of **IQ2** with the Hybrid-2 minor conformation can be detected as well.

Imino and aromatic signals affected by addition of ***IQ2*** have been marked with spheres on the Hybrid-1 NMR-structure of *h-TELO* (PDB ID: 2GKU). The strongest impact upon binding experience residues forming the tetrad at the 5´-end and the 5´-capping structure indicating interaction of **IQ2** with the lower tetrad of *h-TELO*.

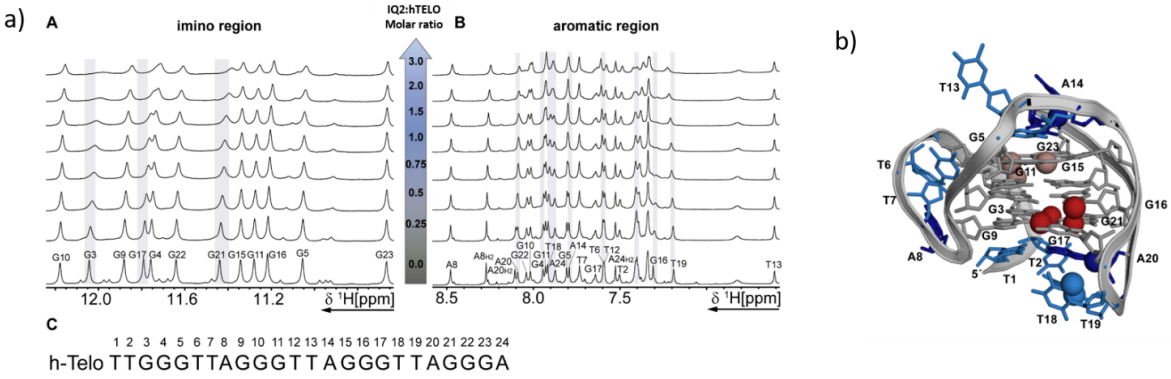


Figure 2.8. a) 1D ^{1}H NMR spectra of the imino (A) and aromatic (B) region of *h-TELO* in 25 mM Tris·HCl buffer containing 100 mM KCl at different [**IQ2**:DNA] molar ratios. The spectra of the imino region have been recorded with jump-return-echo water suppression[2] and the spectra of the aromatic region with gradient-assisted excitation sculpting. The partial assignment of h-Telo has been transferred from Luu et al. 2006 and Bessi et al. 2015 (C) Sequence and numbering of *h-TELO*. b) Structure of h-Telo (PDB: 2GKU) with displayed location of changes in imino and aromatic region of the NMR-spectra (spheres)

upon ligand addition. The residues are coloured in blue, dark blue and grey for dT, dA and dG, respectively. Strongly affected imino signals are highlighted with red spheres and moderately affected imino signals are highlighted with dusky pink spheres.

Interestingly, stepwise addition of **IQ1** especially affected imino signals stemming from the minor Hybrid-2 conformation of *h-TELO* with only minor effects on the major Hybrid-1 conformation suggesting specific binding of **IQ1** to a certain G4-topology (Figure 2.9).

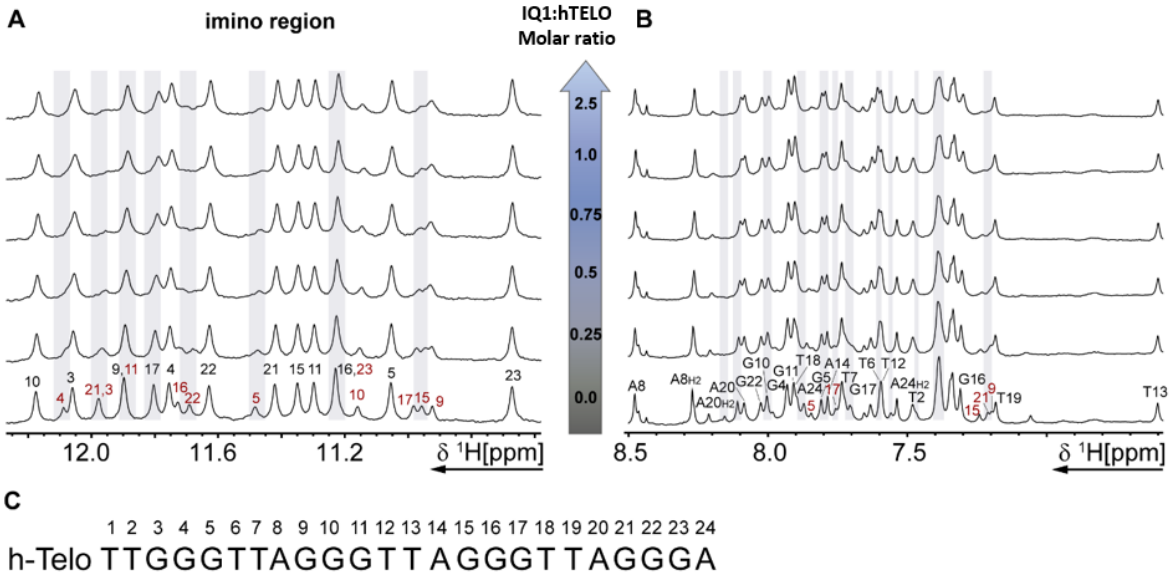


Figure 2.9. 1D ^{1}H NMR spectra of the imino (A) and aromatic (B) region of *h-TELO* in 25 mM Tris·HCl buffer containing 100 mM KCl at different [**IQ1**:DNA] molar ratios. The spectra of the imino region have been recorded with jump-return-echo water suppression and the spectra of the aromatic region with gradient-assisted excitation sculpting. The partial assignment of *h-TELO* has been transferred from Luu et al. 2006 and Bessi et al. 2015. Signals of the major and minor conformation are assigned in black and red, respectively. (C) Sequence and numbering of *h-TELO*.

2.3.Induction of cytotocity by IQ molecules in cancer cells by cell membrane penetration

Cytotoxicity assay

Assessment of the cytotoxicity was carried out for the synthesized molecules which showed high ΔT_m with the G-quadruplexes and in particular the telomeric G-quadruplex. The **IQ** compounds with highest ΔT_m, namely **IQ1, IQ2, IQ3 and IQ7**were examined to investigate if they induce any cytotoxicity in cancer and normal cells. The effect of the ligands on the proliferation of cervical cancer cell line, HeLa and normal cell line HEK 293 were assessed to observe their cytotoxicity (Figure 2.10). HeLa was taken as a representative telomerase positive cell line without telomere shortening and thereby abundance of the telomeric G-rich sequence.On the other hand, HEK 293 which has relatively very low expression of telomerase was taken as a control cell line where normal telomere shortening takes place and the possibility of telomeric G4s in much lesser. Interestingly, among the four lead ligands only **IQ1** and **IQ2** were found to be have IC_{50} values of 6 µM and 9.3 µM respectively in HeLa cells. In addition both **IQ1** and **IQ2** were found to be non-toxic to non-cancerous cells of HEK293 at high dose treatment of upto 100 µM. Thus, our library of IQ molecules now contains two lead small molecules which have potential drug-like properties. Both ligands **IQ1** and **IQ2** cause deleterious effects on cancer cells without causing any cytotoxicity in normal cells.

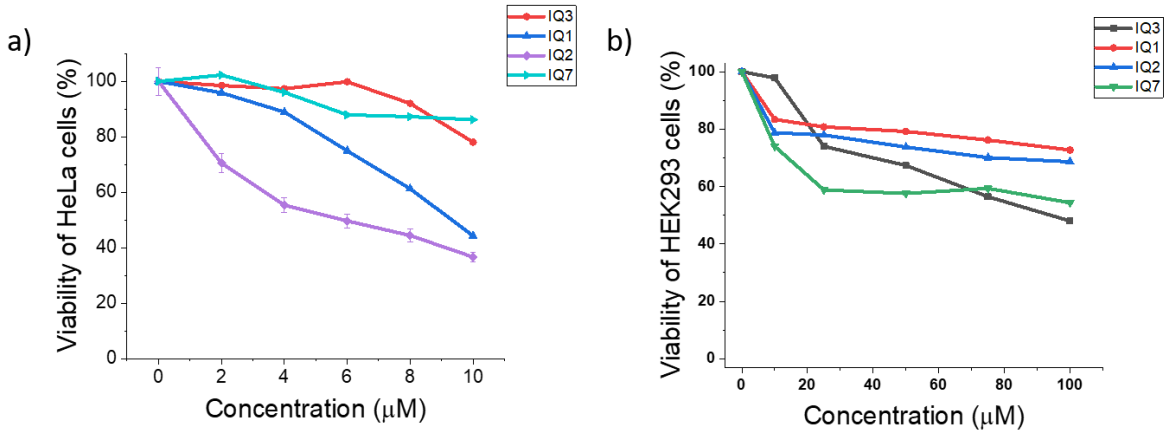


Figure 2.10. Cell cytotoxicity analysis of **IQ1, IQ2, IQ3 and IQ7** in a) HeLa, b) HEK 293 cell lines after 24 h of treatment.

Cell Imaging

In accordance with the promising cytotoxic effects of **IQ1** and **IQ2** observed in cancer cells, both these molecules were selected for subsequent studies in cells to observe which ligand has the most potent effect as an anti-cancer agent. The capability of **IQ1** and **IQ2** to stabilize the *h-TELO* G4 observed in biophysical assays as well as possessing cytotoxic effect in cancer cells were investigated for their cell membrane permeability (Figure 2.11). HeLa cells were treated with **IQ1** and **IQ2** at the concentration of 2 µM. The fixed cells were visualized under confocal microscope and exhibited inherent red fluorescence within the treated cells. This suggests that the **IQ1** and **IQ2** could easily penetrate the cell membrane as well as the nuclear membrane and localize in the cancer cells.

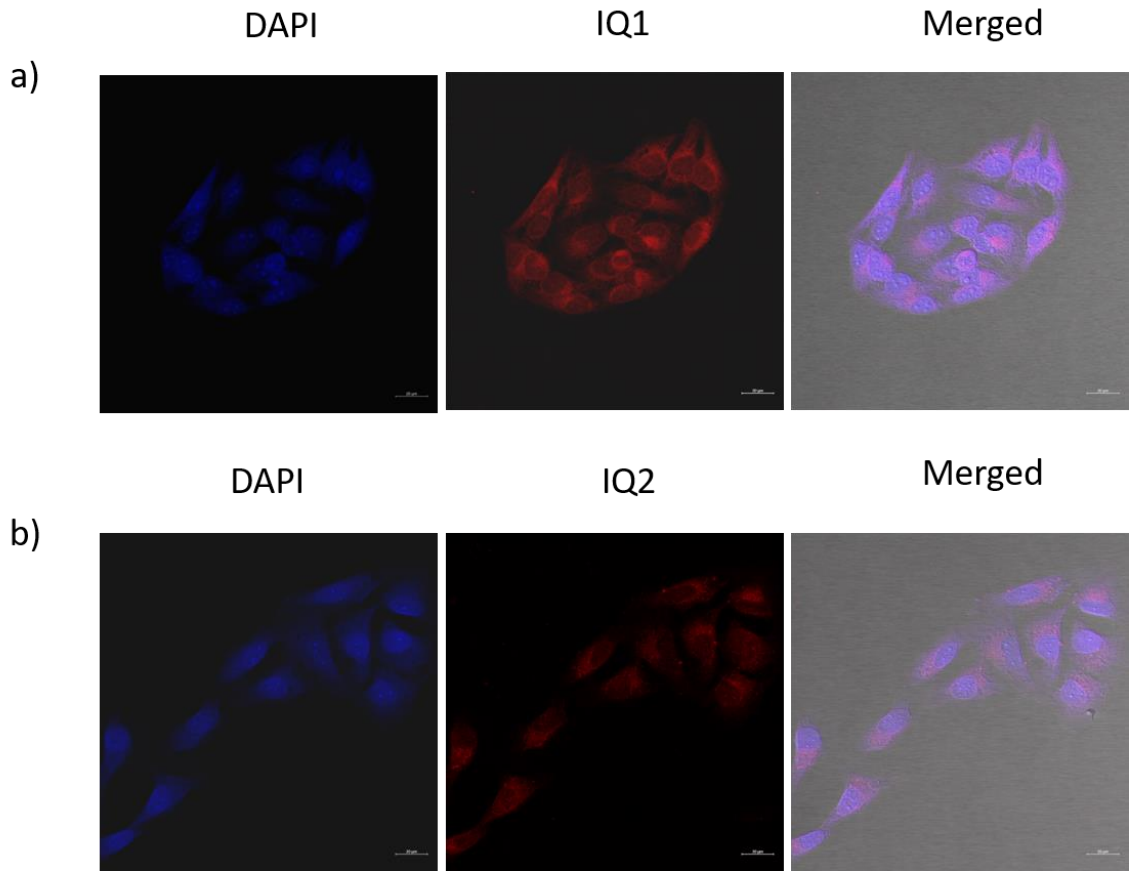


Figure 2.11. Confocal microscopy images HeLa cells treated with 2 µM of indoloquinoxaline ligands a) **IQ1** and b) **IQ2**.

2.4.IQ molecules elicit selective cellular response by modulating gene regulatory process

qRT-PCR

Changes in gene expression of relevant genes related to telomere function and induction of apoptosis upon ligand treatment were studied using qRT-PCR. The relative mRNA expression of *c-MYC*, *BCL-2* and *h-TERT* was assessed to evaluate the effect of the ligands **IQ1** and **IQ2** on the transcription of these genes. HeLa cells were treated with different doses of **IQ1** (4.5 and 9 µM) and **IQ2** (3 and 6 µM) and the mRNA was quantified and normalized against expression of *18s* (Figure 2.12). The expression of *h-TERT* showed a significant decrease of 31.2 and 59.9% after treatment with 3 and 6 µM concentrations of **IQ2** respectively.

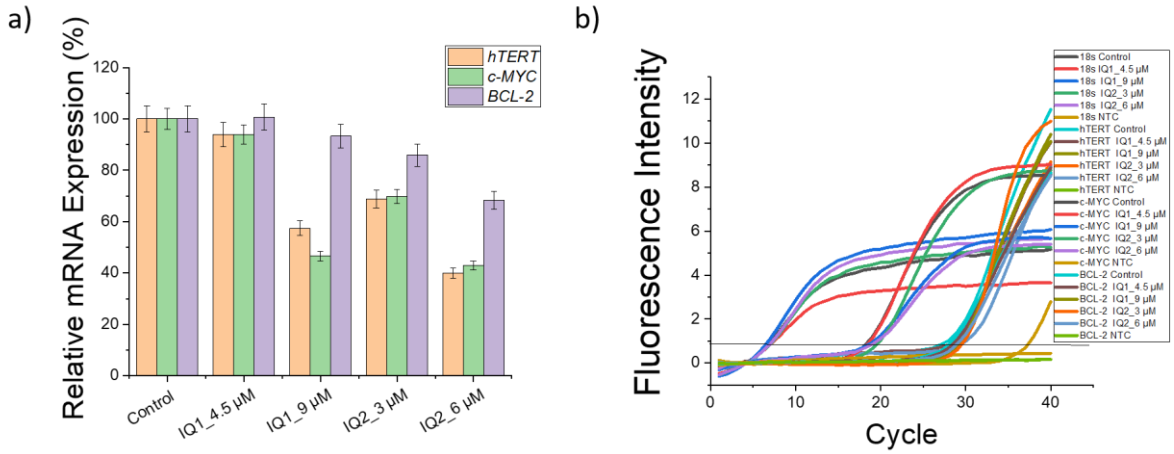


Figure 2.12. a) qRT-PCR results of *hTERT*, *c-MYC* and *BCL-2* genes in **IQ1** and **IQ2** treated HeLa cells, b) SYBR green sigmoidal graphs of qRT-PCR.

On the other hand, *h-TERT* decrease was lesser in case of **IQ1** with only 7% at 4.5 µM dose 42.3% at the higher dose of 9 µM. In addition *c-MYC*, which is reported to enhance expression of *hTERT* was also downregulated in a dose dependent manner by **IQ2** (30.3 and 57.1%) and **IQ1** (7 and 53.3%). In addition, anti-apoptotic gene *BCL-2* was downregulated by **IQ2** (14.2 and 31.7%) which may suggest activation of the apoptotic pathway. In contrast, treatment with **IQ1** induced a downregulation of only 6.7% at the

higher dose of 9 µM without much change at the lower dose of 4.5 µM. These suggests that the expression of *h-TERT* is significantly downregulated by **IQ2** and to some extent by **IQ1** which may result in the depletion of the telomerase enzyme in the telomerase positive HeLa cells.

Table 2.1. Observed values of cycle threshold (***C*T**):

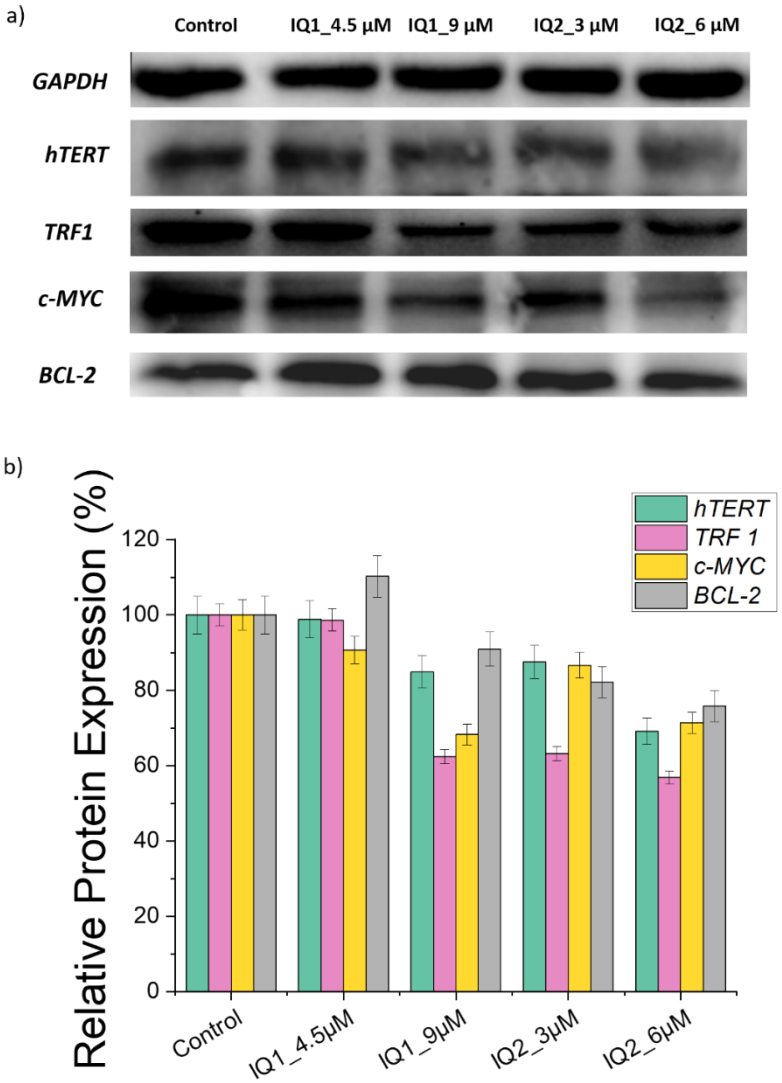
Sample	***18s* (°C)**	***c-MYC* (°C)**	***BCL-2* (°C)**	***hTERT* (°C)**
Control	6.6	17.3	27.9	27.4
IQ1 (4.5 µM)	6.91	17.7	28.2	27.8
IQ1 (9 µM)	6.3	18.1	27.7	27.9
IQ2 (3 µM)	6.78	18	28.3	28.12
IQ2 (6 µM)	6.78	18.7	28.63	28.9

Western Blotting Assay

To confirm the observations of mRNA expression, western blotting assay was performed. As telomeric region is a non-coding part of the genome, the expression of several related proteins that work in tandem with the telomere were assessed. The protein expression of important genes like *h-TERT* (a component of telomerase enzyme), *TRF 1* (a component of the shelterin complex, *c-MYC* (enhancer of *hTERT* expression by binding to the *hTERT* promoter) and *BCL-2* (anti-apoptotic gene) were assessed. *hTERT* is an important component of the telomerase enzyme. In addition, *BCL-2* gene which is relevant for the apoptotic pathway with anti-apoptotic activities was assessed to shed light on the apoptotic activity of the ligands. Western blotting assay was undertaken to estimate the level of protein expression of these genes after the cancer cells were treated with increasing doses of **IQ1** (4.5 and 9 µM) and **IQ2** (3 and 6 µM) (Figure 2.13). The *GAPDH* gene was taken as the housekeeping gene and used for the normalization of results for all genes. The results show that the expression of both *h-TERT* and *TRF 1* and *c-MYC* decreased in a dose dependent manner with **IQ1** and **IQ2**. Treatment of telomerase positive HeLa cells with **IQ2** at doses of 3 and 6 µM resulted in significant downregulation of *hTERT* to 87.4 and 69.1%, *TRF 1* to 63.1 and 56.8%, *c-MYC* to 86.6 and *BCL-2* to 71.35% and 75.0%.

Whereas, **IQ1** (4.5 and 9 µM) also exhibited decreased expression of all genes (*hTERT* = 98.8 and 84.4%, *TRF1* = 98.6 and 62.4%, *c-MYC*= 90.6 and 68.4 and *BCL-2* = 110.2 and 90.8%) the effect of treatment with **90** elicited a more potent response. These results suggest that both **IQ1** and **IQ2** cause imbalance of telomere integrity which result in apoptosis. In addition the telomerase component *hTERT* is also decreased leading to the potential decrease of telomerase activity and loss of cell immortalization.

Figure 2.13. a) Western Blot of **IQ1** and **IQ2** treated HeLa cells. b) Bar graph of



observed protein expression in and normalized with housekeeping gene *GAPDH*.

2.5.IQ molecules trigger DNA damage and apoptosis in cancer cells

Cell cycle assay

The frequency of G-quadruplexes at different phases in the cell cycle is well reported with an arrest of cell cycle in the presence of a G-quadruplex-interactive compounds. The effect of the ligands on cell cycle Hela cells were treated with **IQ1** and **IQ2** for 24 h respectively, and then analyzed by flow cytometry to determine the influence of the **IQ** compounds on cell cycle (Figure 2.14). The cells treated with **IQ2** showed a dose dependent cell cycle arrest in G2/M phase with an arrest of 18.2% and 28.6% cells when treated with 3 μM and 6 μM concentrations of **IQ2**. On the other hand, it was observed that treatment of **IQ1** did not result in cell cycle arrest in any of the phases. This suggests that **IQ2** is a novel ligand which arrests cells in the G2/M phase of the cell cycle, which is the late period of DNA synthesis to inhibit cell growth and cell division. Interestingly, **IQ2** was found to arrests cells in the M phase and is novel in this regard as most G4 ligands tend to arrests the G1 phase or the S phase of the cell cycle.

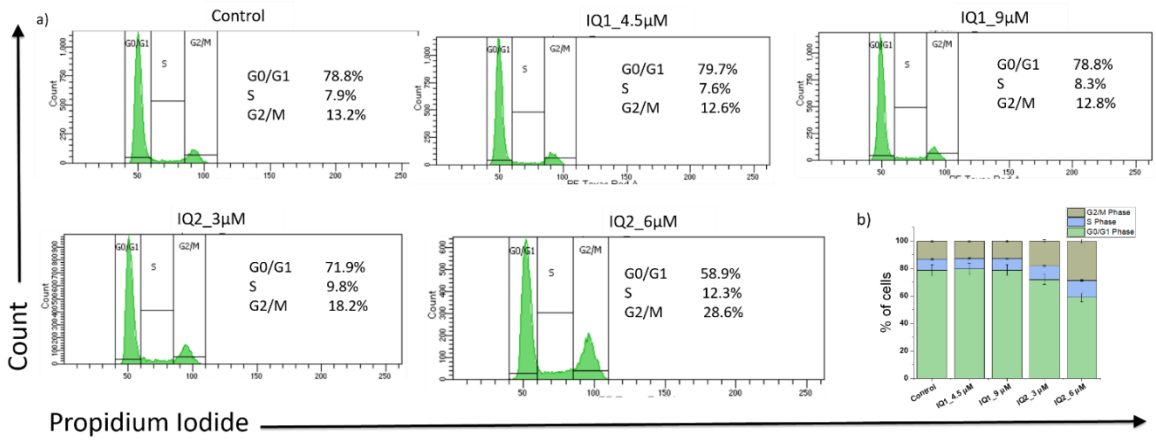


Figure 2.14. a) Flow cytometry graphs of cell cycle studies of HeLa cells treated with increasing doses of **IQ1** and **IQ2**. b) Bar graph of observed percentage of cells in each phase of the cell cycle

Apoptosis assay

Having demonstrated that **IQ** derivatives can arrest cell cycle, investigation of the **IQ** derivatives to induce cellular apoptosis was carried out (Figure 2.15). To evaluate the apoptosis induced by **IQ1** and **IQ2**, Annexin V-FITC/PI double staining assay was carried out and analyzed in flow cytometer. After incubation with 3 µM and 6 µM of **IQ2** for 24%, 13.1% and 25.2% of the cell populations were found to be apoptotic. Whereas treatment of 24h treatment of cells with **IQ1** at concentrations of 3 and 6 induced apoptosis in 9.9% and 10.5% cells. This result suggests that **IQ2** has a higher ability to induce apoptosis in Hela cells.

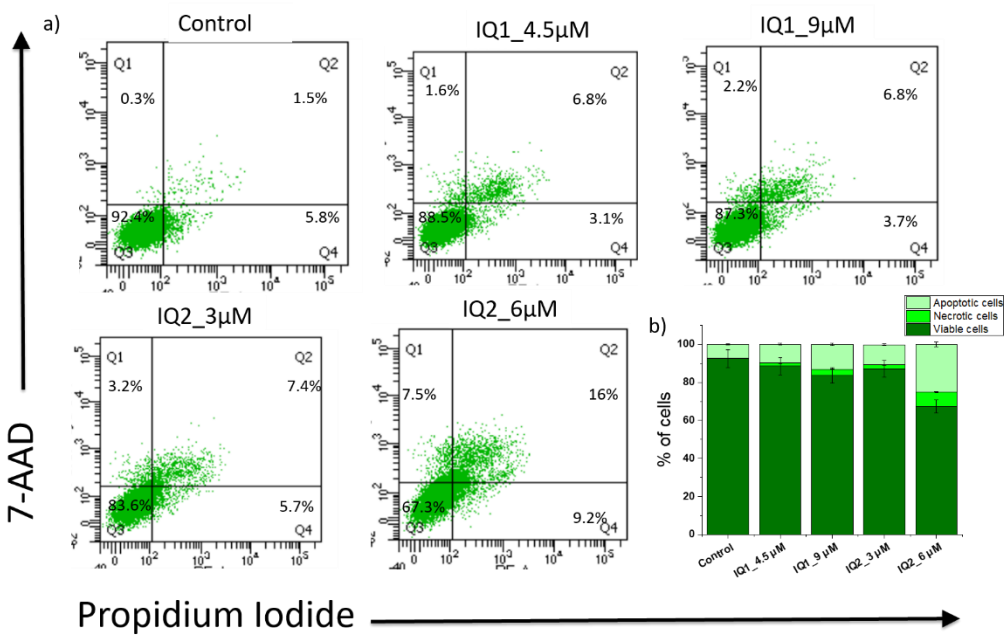


Figure 2.15. a) Flow cytometry graphs of apoptosis of HeLa cells treated with increasing doses of **IQ1** and **IQ2**. b) Bar graph of observed percentage of cells in necrosis and apoptosis.

Caspase assays

Having observed that the **IQ** treatment in HeLa cells induce apoptosis, investigation was carried out to find the mode of cell death. Caspase 8 is activated in the extrinsic apoptotic pathway whereas Caspase 9 is activated in the intrinsic apoptotic pathway. Both

these enzymes in turn activate caspase3/7 where Caspase 3 and 7 are key apoptotic enzymes that function together to initiate cellular apoptosis. The detailed evaluation of the apoptotic pathway was carried out for the **IQ** treated HeLa cells by using Caspase 8 and Caspase 3/7 assays and analyzed in flow cytometer to measure the extent of cells where Caspase 3/7 and Caspase 8 were initiated (Figure 2.16). Upon analysis it was observed that the Caspase 8 was not significantly activated either in cells treated with **IQ1** (1% and 2%) and **IQ2** (2.7% and 3.2%) at concentrations of 3 μM and 6 μM respectively. However, it was observed that incubation of HeLa cells with **IQ2** resulted in significant dose dependent caspase activity with 23.7% and 35.1 % caspase 3/7 positive cells at 3 μM and 6 μM. Whereas, in cells treated with **IQ2** caspase 3/7 was activated to a lesser extent at 4.5 and 9 μM with only 10.4% and 14.7% of cell population. Thus, it was found that **IQ2** induces apoptosis through activating the caspase 3/7 without the activation of Caspase 8 suggesting that the induction of apoptosis is through the intrinsic apoptotic pathway.

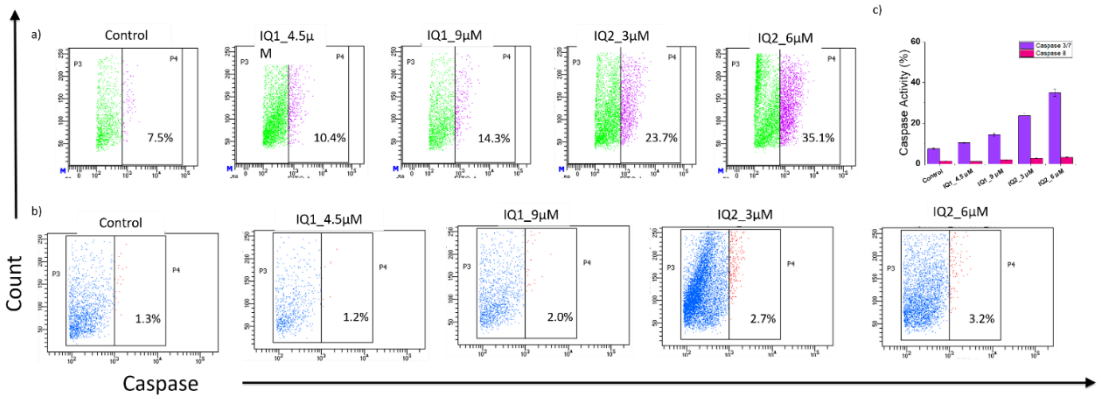


Figure 2.16. a) Flow cytometry graphs of HeLa cells treated with increasing doses of **IQ1** and **IQ2** showing activation of Caspase 3/7. b) Bar graph of observed percentage of cells in Caspase 3/7 and 8, c) Flow cytometry graphs of HeLa cells treated with increasing doses of **IQ1** and **IQ2** showing activation of Caspase 3/7 and Caspase 8.

DNA Damage assay

Double strand breaks (DSBs) can be induced by mechanisms such as ionizing radiation or cytotoxic agents. Measuring damage to DNA due to potential drug molecules is central in many studies of cancer biology. DSBs can be readily detected by using γ-H2AX antibody in flow cytometry to allow the assessment of DNA damage. HeLa cells treated with **IQ1**and **IQ2** were evaluated for DNA damage (Figure 2.17). It was observed that both **IQ1** and **IQ2** induce a dose dependent DNA damage. **IQ2** was found to demonstrate a much higher level of DNA damage of 44% and 62% at both 3 μM and 6 μM concentrations respectively. **IQ1**also showed similar DNA damage of 45% at 4.5 μM but the second dose of 9 μM showed a lesser degree of DNA damage of 54% compared to **IQ2**. Thus, both the **IQ** molecules elicit DNA damage in cancer cells with **IQ2** showing a higher level of damage at the IC_{50} concentration of 6 μM.

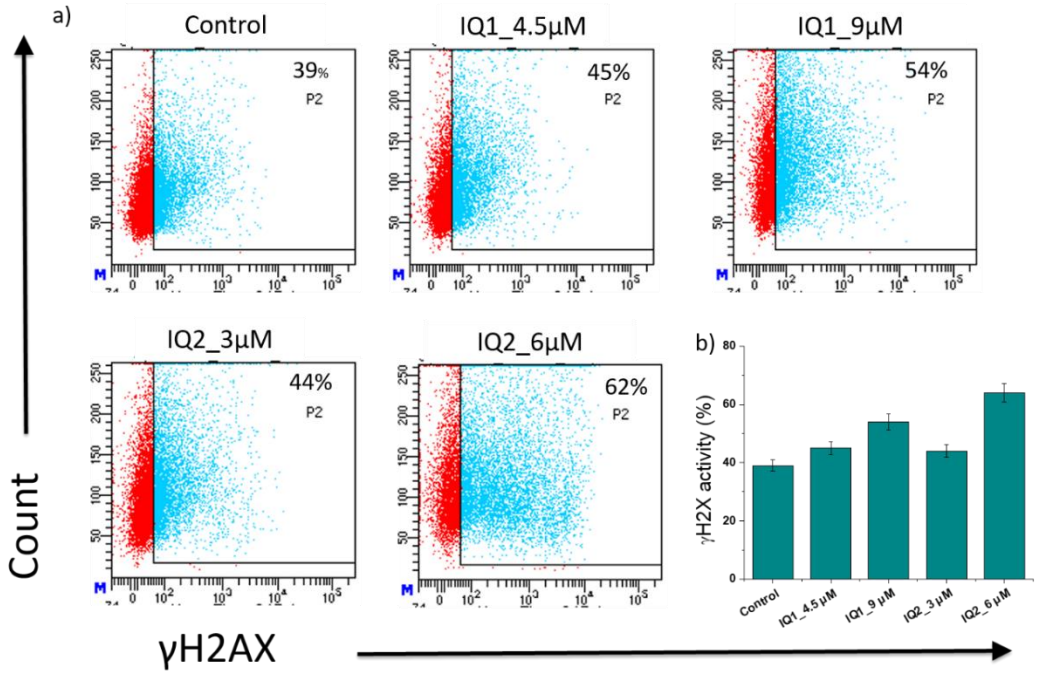


Figure 2.17. a) Flow cytometry graphs of HeLa cells treated with increasing doses of **IQ1** and **IQ2** showing double stranded DNA damage. b) Bar graph of observed percentage of cells showing DNA damage

ROS generation detection

Reactive oxygen species (ROS) generation was performed to assess the generation of oxidative species leading to apoptosis. It was observed that the treatment of Hela cells with the IQ ligands lead to increase in the ROS generation signified by the shift of the shift of the cell population stained with DCFDA (Figure 2.18). **IQ1** showed a similar ROS generation of 15.6 and 17.8% at concentrations of 4.5 µM and 9 µM. Whereas in Hela cells treated with **IQ2**, a concentration dependent response was observed with increase of ROS generation from 20.33 to 30.8% at concentrations of 3 µM and 6 µM. H_2O_2 was taken as a positive control and exhibited 48% ROS damage. This suggests that **IQ2** induces ROS generation which may consequently play a role in driving the cancer cells towards apoptosis.

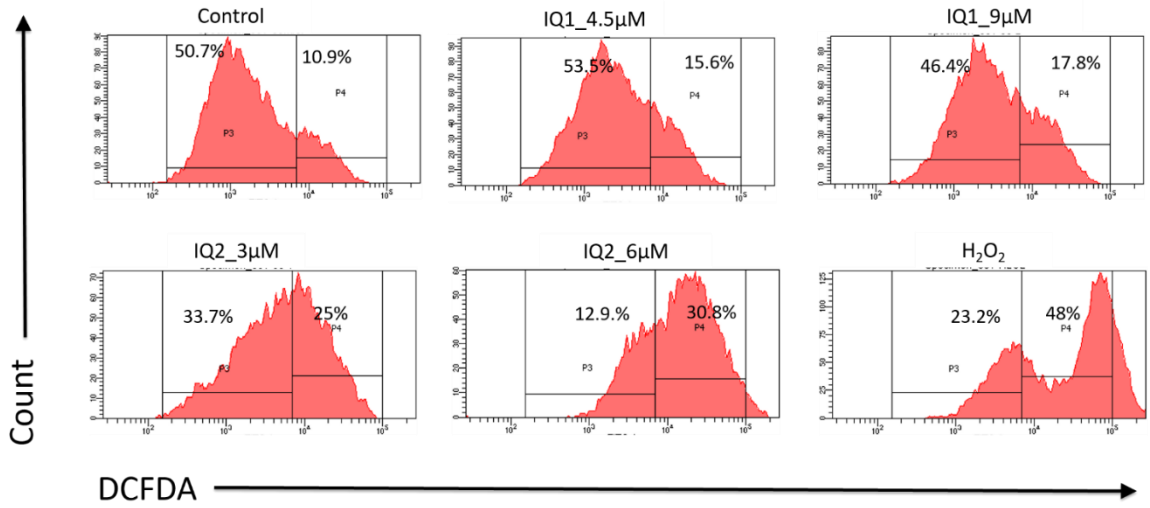


Figure 2.18. a) Flow cytometry graphs of HeLa cells treated with increasing doses of **IQ1** and **IQ2** showing generation of reactive oxygen species.

2.6.IQ molecules hamper the activity of telomerase to arrest telomere elongation

Telomerase Repeated Amplification Protocol (TRAP)

The telomere elongating telomerase enzyme is overexpressed in a various human cancers. TRAP–LIG assay is a reliable and reproducible method for measuring of telomerase activity in telomerase positive cells to assess the level of telomerase inhibition. TRAP-lig assay was performed to assess the activity of telomerase upon treatment with **IQ1** and **IQ2** at different doses (Figure 2.19). The telomerase enzyme was isolated from HeLa cells and subjected to PCR extension in presence of the ligands at respective doses. It was observed that the relative telomerase activity decreased in a dose dependent manner on treatment with **IQ2**. With treatment at 3 μM concentration of **IQ2**, the telomerase activity decreased from 100% in untreated to 56.58% in treated sample. Further reductions were observed at increasing dose 6 μM which were found to be 42.8%. A dose dependent reduction was also observed after treatment of **IQ1** at different concentrations. The telomerase activity was found to be, 91.87% and 80.4% at the 2 different doses of **IQ1** (4.5 and 9 μM). According to the result, the telomerase inhibition by **IQ2** was much greater than **IQ1**. This suggests that **IQ2** has the selective affinity as a potent G4-binder for the telomeric G4 exclusively and consequently leads to blocking of the telomeric region which usually serves as the substrate for telomerase-mediated extension of the telomere in cancerous cells.

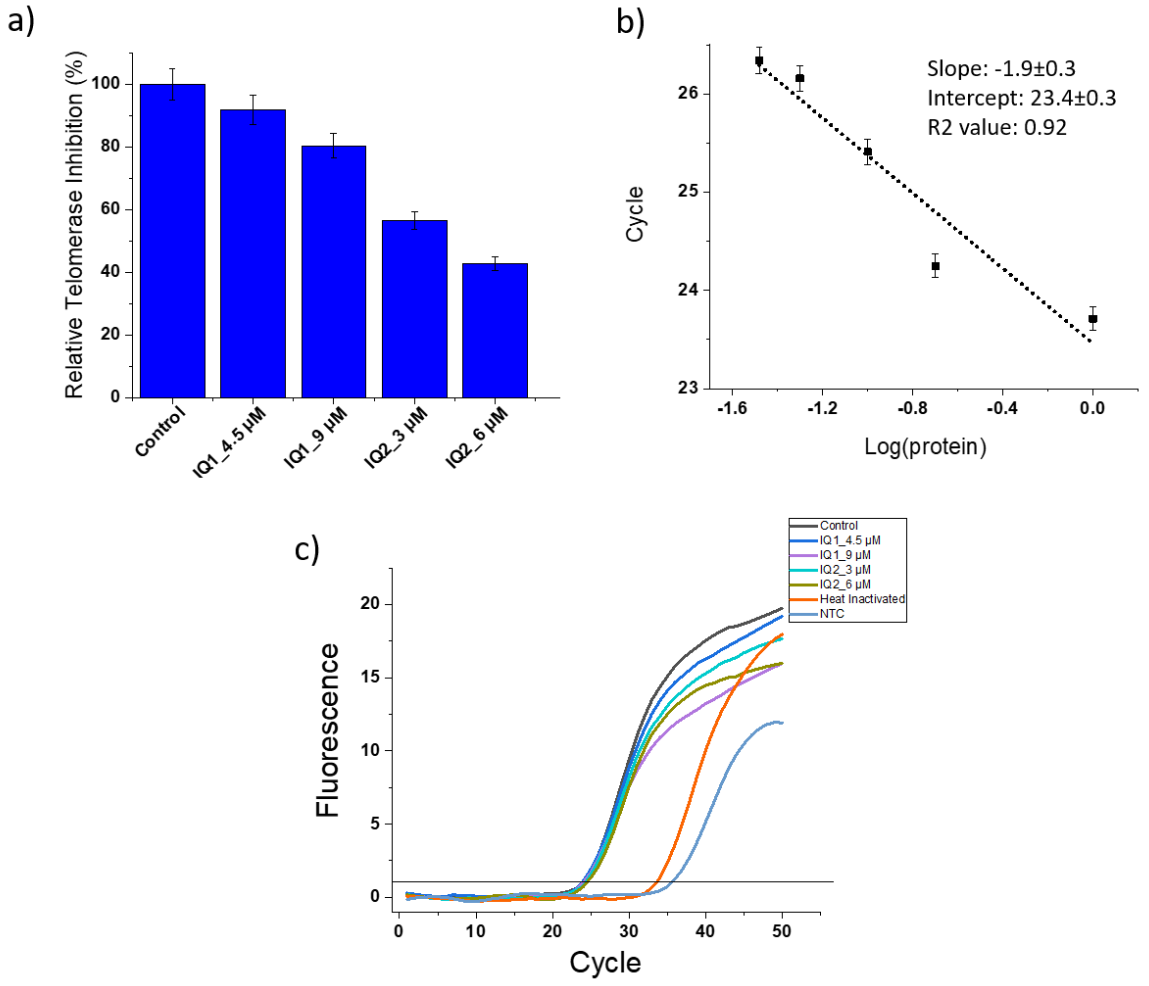


Figure 2.19. a) Analysis of telomerase inhibition in HeLa cells treated with increasing doses of **IQ1** and **IQ2**, b) q-TRAP standard curve, c) q-TRAP SYBR greem amplification graphs of **IQ1** and **IQ2** treated HeLa cell lysates subjected to telomerase extension.

Conclusion

In summary, both **IQ1** and **IQ2** exhibit anti-cancer properties which can be used to target cellular telomeric G-quadruplex. In addition, both ligands show that they can efficiently reduce telomerase activity which is an overexpressed hallmark of most cancer types. In particular, **IQ2** elicits a higher range of cellular responses including arrest of cancer cells in G2/M phase of cell cycle, ROS generation, apoptosis, activation of 3/7 caspase, greater DNA damage and significant reduction in telomerase activity (Figure 2.20). Thus, **IQ2** is

a molecule that affects a range of cellular activities, all of which summate to deleterious effects on the cell and result in cell death. On the other hand, **IQ1** is also a promising ligand but the aromatic side chains of **IQ1** may potentially hinder critical processes such as cellular uptake, nuclear penetration, stable binding interaction and eliciting of cellular response. The aliphatic nature of the side chain in **IQ2** may promote these critical processed resulting in lower IC_{50} value and greater cellular response.

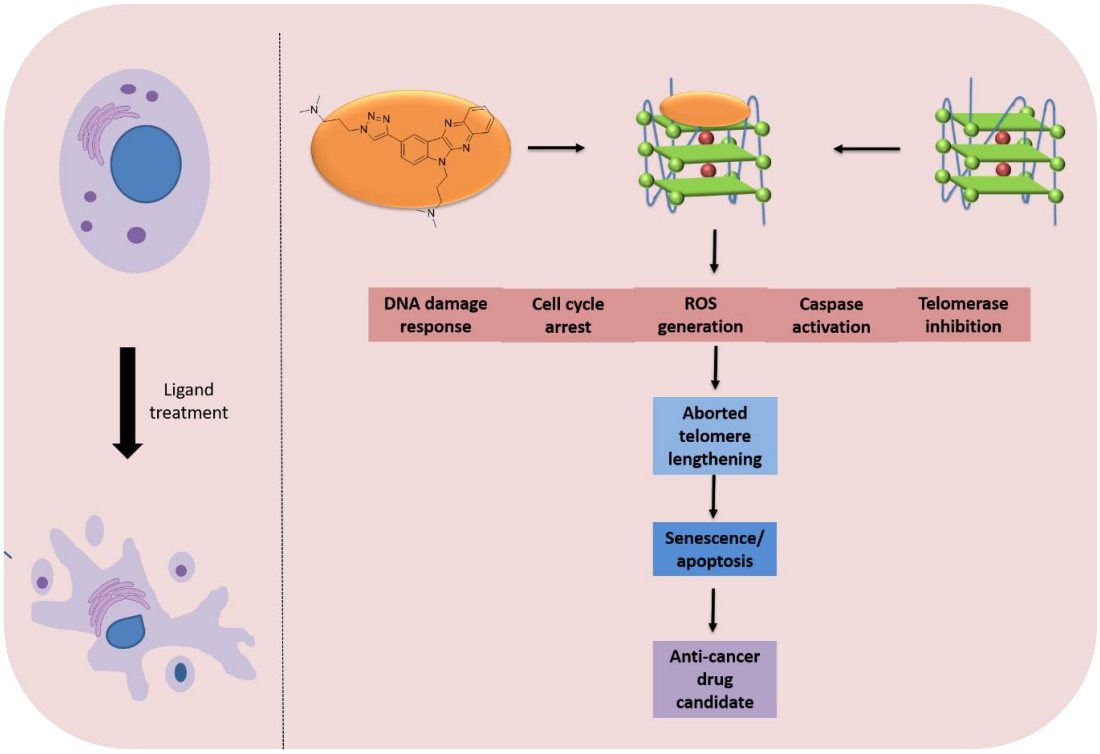


Figure 2.20. Summary of cellular responses induced in cancer cell by indoloquinoxaline derivative **IQ2** by stabilization of the telomeric G-quadruplex.

The induction of apoptosis in cancer cells by small molecules containing the indoloquinoxaline scaffold holds significant therapeutic potential. By specifically targeting cancer cells and promoting their programmed cell death, small molecules offer a promising

approach to combat cancer. However, further research is needed to pinpoint and elucidate the additional cellular mechanisms at play, optimization of side chains, decreasing dosage concentrations, and evaluate their efficacy and safety in preclinical and clinical studies.

Chapter 3

Exploring alternative therapeutics for *c-KIT* overexpressing cancer cells by targeting *c-KIT* G-quadruplex using specific ligand

Introduction

c-KIT proto-oncogene encodes a transmembrane receptor tyrosine kinase (RTK), primarily expressed in hematopoietic cells and myeloid cells, which responds to stem cell factor (SCF).[92, 93] *c-KIT* regulates cell proliferation, survival, migration, pigmentation, and hematopoiesis.[94, 95] As *c-KIT* plays a pivotal role in establishing normal cell function, *c-KIT* mutations or overexpression result in oncogenic cellular transformations including gastrointestinal stromal tumors (GISTs), melanomas, mastocytosis, and acute myeloid leukemia (AML).[96, 97] Therefore, *c-KIT* has emerged as a promising molecular target for anti-cancer therapy.[98] Attempts to inhibit *c-KIT* overexpression have led to the development of small molecule inhibitors like the Food and Drug Administration (FDA) approved drug imatinib mesylate (Gleevec). Subsequently, commercially available drugs like dasatinib, sunitinib, nilotinib, and sorafenib have been developed for *c-KIT* inhibitors.[99-101]

Using natural or synthetic molecules to modulate aberrant gene expression by targeting non-canonical DNA conformations has emerged as an alternative strategy for potential cancer therapeutics.[98, 102-106] The human *c-KIT* oncogene contains two guanine-rich (G-rich) sequences *KIT 1* and *KIT 2*, upstream of the transcription starting site.[107-109] These G-rich sequences can adopt G-quadruplex topologies and modulate cellular functions.[110] G-quadruplexes consist of stacked G-tetrad subunits, where each tetrad contains four coplanar guanines associated by Hoogsteen hydrogen bonds together with potassium or sodium counter-ions coordinated between the G-tetrads.[80, 111, 112] Recent studies elucidate the regulatory role of G-quadruplexes in telomeres (*h-TELO*) and promoter regions of genes like *c-MYC*, *BCL-2*, *c-KIT*, etc.[39, 113-126] Several studies have shown that the *c-KIT* quadruplexes can be stabilized by appropriate small molecules, inhibiting *c-KIT* transcription.[127-132]

Small molecules capable of stabilizing the G-quadruplex structure hold the potential for the discovery of novel anticancer agents.[56, 57, 127, 133-139] In most cases, the quadruplex stabilization by a ligand (usually a planar aromatic molecule) occurs *via* π–π stacking and electrostatic interactions with the terminal G-tetrad of the quadruplex. This external stacking mode of binding represents a unique feature of quadruplex recognition compared to other DNA structures. Electrostatic interactions between positively charged ligands and

the phosphate backbone also strongly participate in the stabilization of G4. Selection of aromatic ring system for optimal stabilization of specific G-quadruplex is a key feature in promoting π–π interaction with the planar G-tetrad of G4 to drive the selection of a specific G-quadruplex over other G4s. Relying on these unique features of G4 binders, aa molecular library of small molecules were synthesized consisting of 14 small molecules containing a fused heteroaromatic core of carbazole or dibenzofuran chemical scaffold with triazolyl side chains.

Carbazole ring system is a key structural scaffold of many biologically active compounds, including synthetic and natural products.[140] Many studies have shown that carbazole containing ligands possess anticancer, antibacterial, anti-inflammatory, antifungal, anti-HIV, hepatoprotective, antiprotozoan, or topoisomerase II inhibition ability.[141-143] A carbazole derivative is known to bind *c-KIT* G-quadruplex but it has not been explored in detail.[144] Heteroaromatic dibenzofurans, containing a central furan ring between two benzene rings, have also been devised to exhibit potent anticancer, anti-bacterial, anti-fungal, anti-HIV and anti-inflammatory activities in medicinal chemistry.[145]

Rational designing of small molecule ring systems with optimal side chain functionalization can preferentially select biomolecular targets in cellular systems.[64, 146] Small molecules having structural features to aid membrane permeability and cellular detection while exhibiting inherent fluorescence, is highly desirable for the selective recognition of specific G4s. Furthermore, small molecules must be synthesized in high overall yields using simple synthetic protocols, and they must exhibit cytotoxicity towards cancer cells over normal somatic cells.[147] Our previous reports with carbazole derivatization have used bis-triazolyl substitutions with a larger molecule size to target G-quadruplexes.[148, 149] In this study, carbazoles (**MC 1-7**) and dibenzofurans (**MD 8-14**) were designed and synthesized with mono-triazolyl functionalization to enable differential recognition of specific G4 topology over other G4s and *ds*DNA abundant within the genome (Figure 3.1). In addition, a less explored dibenzofuran scaffold was studied as a potential G4 interacting chemical moiety. The rationale for selecting mono-triazolyl side chains over the previously reported bis-triazolyl ligands from our lab,[142] is our attempt to address poor permeability and reduce molecular size to allow ease of entry to the ligands through the cellular membrane.

In this study, biophysical investigations were carried out to study the stabilization potential of **MC** and **MD** series of small molecules for different G-quadruplexes and duplex DNA (*ds*DNA). Initial screening was performed to identify the best G4 stabilizer in the 14 small molecular library using a FRET-based melting assay. Fluorescence spectroscopy and Isothermal titration calorimetry (ITC) indicated a triazolyl carbazole **MC-4** as the selective binder of *c-KIT 1* G4. Biological studies established that **MC-4** is cell-permeable and downregulates *c-KIT* gene expression. We also highlight that inspite of containing the same side chain, the scaffold selection is a major recognition and stabilizing factor for the molecules as evidenced by the carbazole **MC-4**, which has selectivity for the *c-KIT 1* G, whereas its dibenzofuran counterpart has very little binding for the same G-quadruplex.

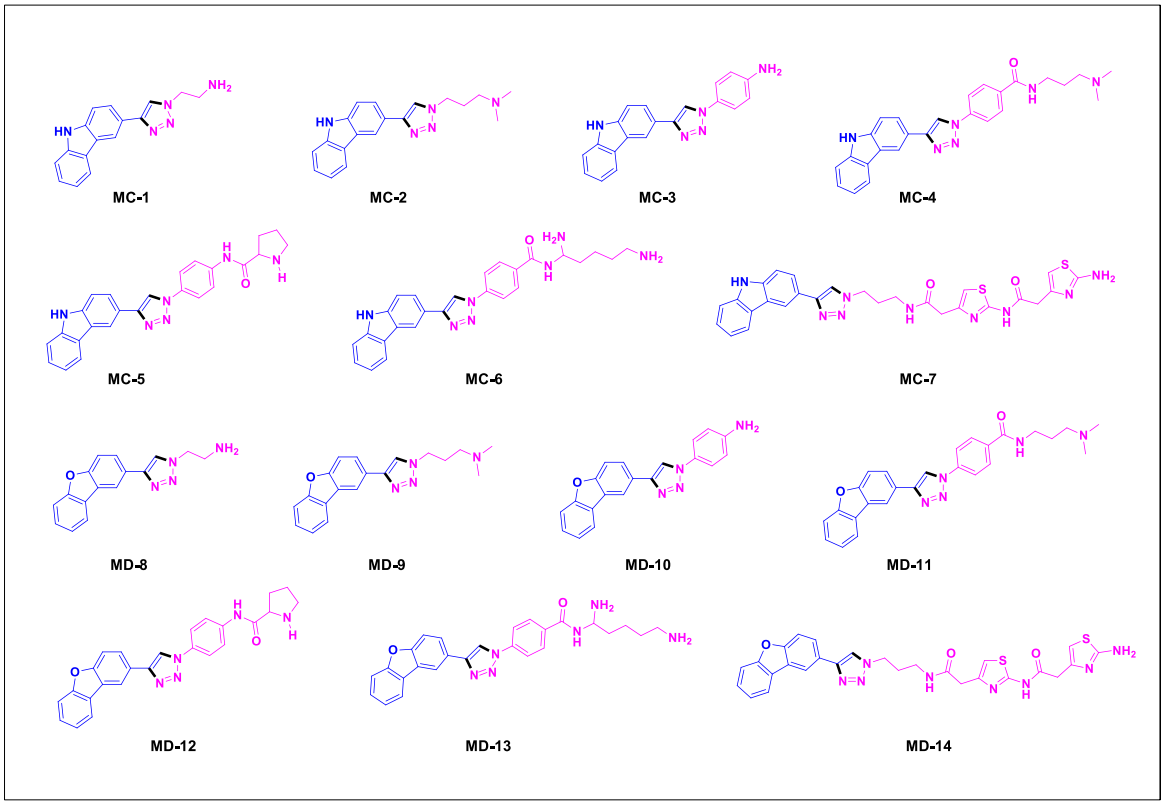


Figure 3.1. Chemical structures of mono-triazolyl carbazole (**MC1-7**) and mono-triazolyl dibenzofuran (**MD8-14**) derivatives with respective side chain modifications.

Results And Discussion

Ligand mediated stabilization of G-quadruplexes

FRET Melting Assay

The FRET melting assay[150-153] was employed to evaluate how effectively the synthesized small molecules could stabilize various various dual labeled (5' FAM and 3' TAMRA) G-quadruplexes present in the oncogenic promoter *(c-MYC, BCL-2, c-KIT 1 c-KIT 2, VEGF*) and telomeric (*h-TELO*) regions. A double-stranded DNA (*ds*DNA) sequence was taken as a control (Figure 3.2). Each ligand of the M-series (**MC** and **MD**) was tested at 1 μM ligand concentration. Among the MC series, only **MC-4** demonstrated a marked difference on the T_m (melting temperature) of the G-quadruplexes, with the most notable change observed with *c-KIT 1* (ΔT_m = 20.1 °C). **MC-4** stabilized other G4s to a lesser extent [ΔT_m °C): *c-MYC* = 3.2, *c-KIT 2* = 9.9, *BCL-2* = 2.2, *VEGF* = 3.5, *h-TELO* = 4.5, and *ds*DNA = 0.7] (Figure 3.3).

The other **MC** ligands (**MC-1**, **MC-2**, **MC-3**, **MC-5**, **MC-6**, and **MC-7**) did not exhibit an appreciable change in ΔT_m for the studied G4s. Similarly, the **MD** ligands (**MD-8-14**) showed no significant stabilization (≤5.0 °C) for any of the examined G4s. Among the dibenzofurans, only **MD-13** and **MD-10** exhibited moderate ΔT_m values of 5.3 and 5.7 °C for *c-KIT 1* and *VEGF* G4 respectively. **MD-11** contains the same side chain as **MC-4** and only differs in the hetero-aromatic region with a dibenzofuran ring instead of a carbazole ring. Interestingly, despite sharing the same side chain as **MC-4**, **MD-11** did not show any stabilization (ΔT_m < 3.0 °C) for the G4s, suggesting that the central chemical scaffold is the most significant contributor to G4 interaction in this case, which may be owing to the protonable nitrogen atom. In contrast, the side chain plays a secondary role in promoting a specific G-quadruplex recognition. In addition, the **MC** and **MD** ligands did not increase the melting temperature (T_m) of the control *ds*DNA.

As **MC-4** emerged as a molecule of interest in the initial screening experiment, **MC-4** was selected to study the stabilization profiles of G-quadruplexes in presence of this small-molecule. Furthermore, **MD-11** ligand containing the dibenzofuran moeity and the same side chain as **MC-4** did not show substantial stabilization with *c-KIT 1* G4, we have

taken it for further studies as a negative ligand to show the differential binding and activity profiles of the different heteroaromatic scaffolds in regards to G4 binding.

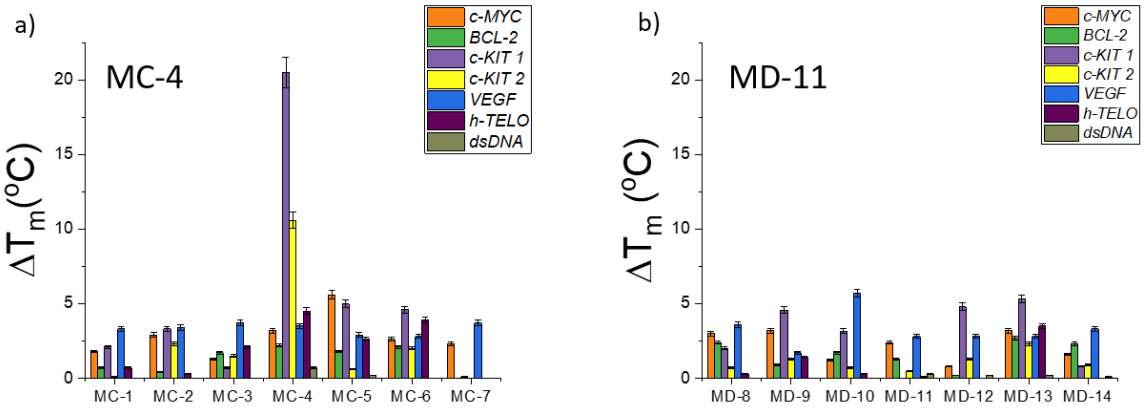


Figure 3.2. FRET melting assay carried out with pre-folded G-quadruplex sequences (*c-MYC, BCL-2,* c-*KIT 1, c-KIT 2, VEGF, h-TELO* and *ds*DNA) at the concentration of 0.2 µM and the ligand concentration of 1 µM with ligands, a) **MC1-7** and b) **MD8-14**. All experiments were carried out in 60 mM Potassium cacodylate buffer (pH-7.4).

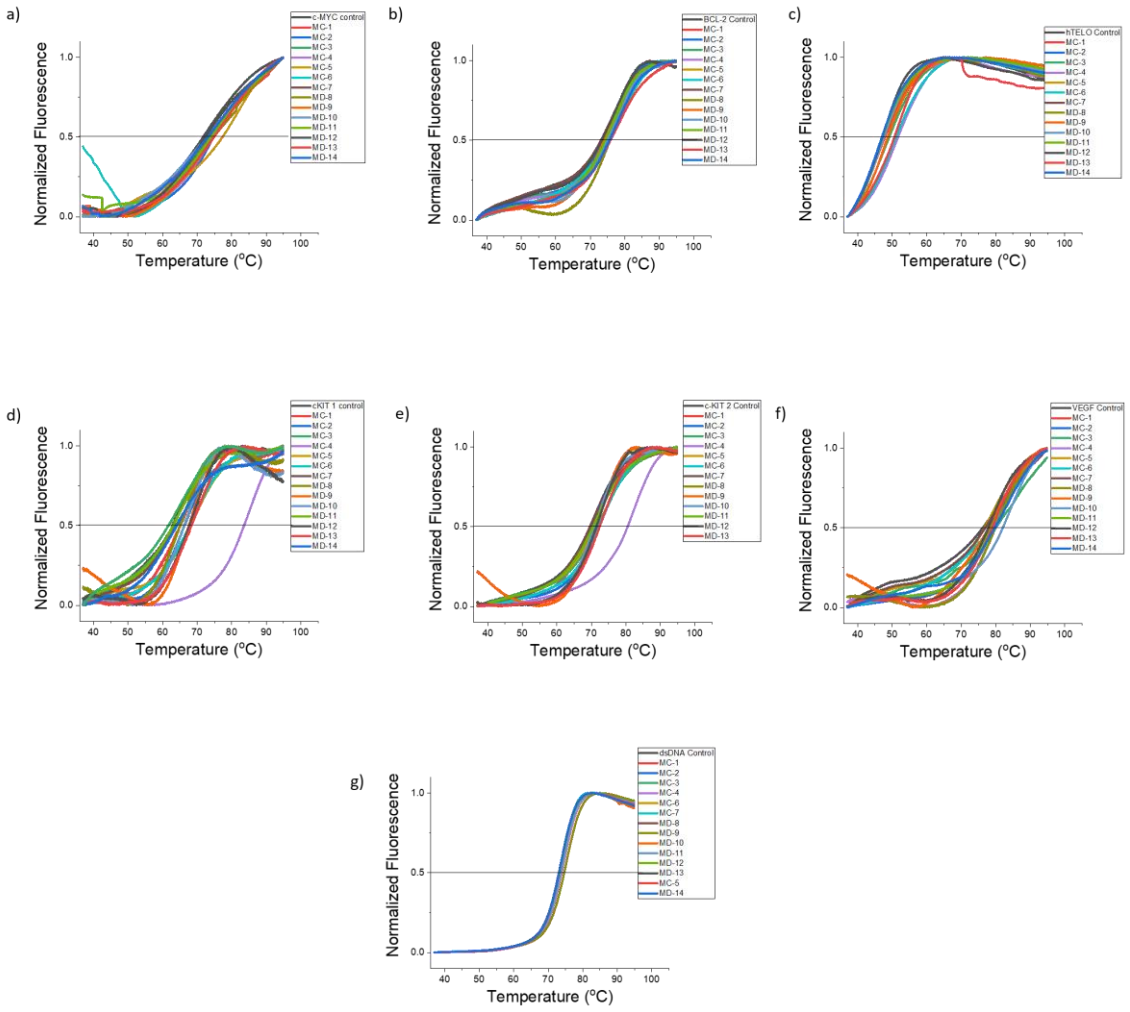


Figure 3.3. FRET melting curves of M-series **(MC-1-7 and MD8-14)** ligands with G-quadruplexes, a) *c-MYC*, b) *BCL-2*, c) *h-TELO*, d) *c-KIT 1*, e) *c-KIT 2*, f) *VEGF* and g) *ds*DNA. The experiment was performed with dual labeled G-quadruplexes (200nM) and the ligands of **M**-series (1 µM) in 60mM Potassium cacodylate buffer (pH-7.4).

Dose Dependent FRET Melting assay

FRET melting assay was performed at different ligand concentrations with the lead carbazole ligand **MC-4** and its dibenzofuran counterpart **MD-11** (Figure 3.4-3.5). Increasing concentrations of **MC-4** and **MD-11** (1 µM, 3 µM, 5 µM, 7 µM and 10 µM) were added to each G4 to determine dose- dependent stabilization of G4s. Interestingly,

MC-4 exhibited the highest stabilization of *c-KIT 1* in a dose-dependent manner. The ΔT_m values of *c-KIT 1* in the presence of **MC-4** at 1 μM, 3 μM, 5 μM, 7 μM and 10 μM were observed to be 20.1 °C, 22.6 °C, 25.5 °C, 26.5 °C, and 26.5 °C, respectively**.**

In comparison, the stabilization of other G4s by **MC-4** was notably lesser with17.3 °C,12.1 °C, 18.1 °C, 11.2 °C, 20.3 °C for *c-MYC , BCL-2*, *c-MYC*, *c-KIT 2, VEGF*, and *h-TELO* respectively, at the highest concentration of 10 μM. In addition, *ds*DNA displayed only a small degree of stabilization at 10 μM of **MC-4**. The results suggest that **MC-4** has a high affinity for the *c-KIT 1* compared to other G4s and *ds*DNA. In contrast, **MD-11** with the same side chain showed no significant alteration in ΔT_m with any of the G4s. A slight increase in T_m was observed for *c-KIT 1* with ΔT_m of 0.0 °C, 1.6 °C, 2.9 °C, 3.3 °C and 5.1°C at 1 μM, 3 μM, 5 μM, 7 μM and 10 μM of **MD-11** respectively. Thus, carbazole derivative **MC-4** selectively stabilizes *c-KIT 1* G4 over other G4s.

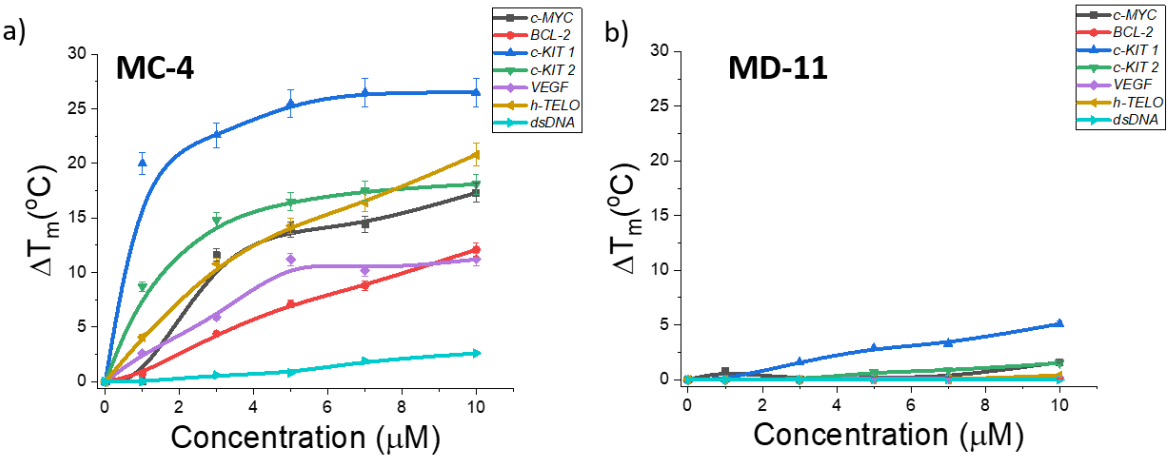


Figure 3.4. Dose dependent FRET melting assay of G-quadruplexes (*c-MYC, BCL-2,* c-*KIT 1, c-KIT 2, VEGF, h-TELO* and *ds*DNA) at concentration of 0.2 μM with increasing doses of a) Carbazole derivative, **MC-4** at concentrations of 1 μM, 3 μM, 5 μM, 7 μM and 10 μM, and b) Dibenzofuran derivative, **MD-11** at concentrations of 1 μM, 3 μM, 5 μM, 7 μM and 10 μM. All experiments were carried out in 60 mM Potassium cacodylate buffer (pH-7.4).

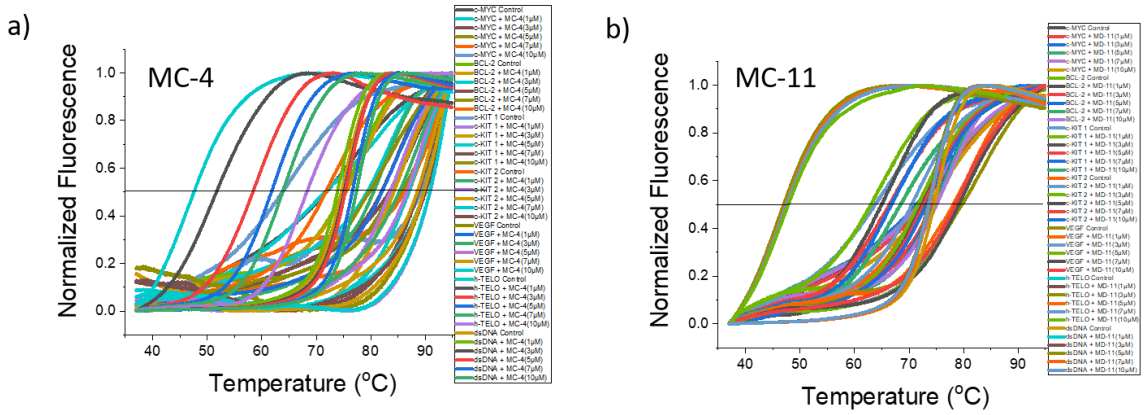


Figure 3.5. FRET melting curves of M-series **(MC-1-7 and MD-8-14)** ligands at different doses (1, 3, 5, 7, 10 µM) with 200 nM of dual-labeled G-quadruplexes (*c-MYC*, *BCL-2*, *h-TELO*, *c-KIT 1*, *c-KIT 2*, *VEGF* and *ds*DNA). The experiment was performed in 60 mM Potassium cacodylate buffer (pH-7.4).

Competition FRET melting assay

Next, a competitive FRET melting assay was performed in the presence of an excess unlabeled DNA competitor to assess the selectivity of **MC-4** for the *c-KIT 1* G4. Calf thymus (CT) DNA was used as a DNA competitor for the studies. The competitor CTDNA was added at several fold molar equivalents (0 eq., 1 eq., 5 eq., 10 eq., 20 eq., and 50 eq and 100 eq.) excess of the fluorophore labeled *c-KIT 1* DNA in the presence of **MC-4** and **MD-11** (1 µM) (Figure 3.6). The ΔT_m of *c-KIT 1* (20.1°C) in the presence of **MC-4** did not exhibit any significant (≤3.5 °C) alteration with the competitor DNA up to 50 equivalent molar excess. Thus, it was found that *CT*DNA is not a substrate for **MC-4** even at very high concentrations and **MC-4** showed high selectivity for *c-KIT 1* G4 over other DNA substrates. In addition, the negative ligand, **MD-11** did not show any significant alteration of T_m (≤3.5 °C of *c-KIT 1* G4 in presenec and absence of the CT DNA.

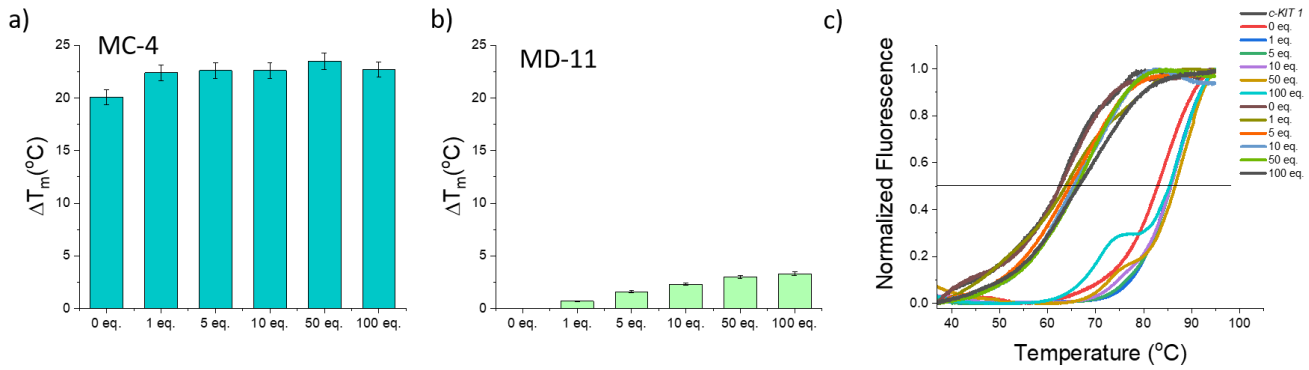


Figure 3.6. Competition FRET melting assay of *c-KIT 1* G-quadruplex (0.2 µM) with the DNA competitor (CT DNA) at excess of 0 eq., 1 eq., 5 eq., 10 eq., 20 eq., and 50 eq and 100 eq. Compared to G-quadruplex DNA concentration.The ligand concentration was 1µM of a) Carbazole derivative, **MC-4**, and b) Dibenzofuran derivative, **MD-11**, c) The melting curves of the competition melting FRET assay performed with *c-KIT 1* G4, CT DNA and ligands **MC-4** and **MD-11**. All experiments were carried out in 60 mM Potassium cacodylate buffer (pH-7.4).

From the initial FRET studies, we considered the lead G4-binder for *c-KIT 1* G-quadruplex to be **MC-4** as it exhibited higher stabilization potential and differential recognition of G4s. Further analyses were conducted with **MC-4** to elaborate on the nature of the *c-KIT 1*/**MC-4** interactions and study its effect on cellular systems.

Selective interaction of triazolyl MC-4 with *c-KIT 1*

Fluorescence Spectroscopy

Fluorescence spectroscopy was performed to determine the binding affinity of the lead carbazole **MC-4** for each G4 employed in this study. **MC-4** was titrated with increasing concentrations of parallel (*c-MYC*, *c-KIT 1, c-KIT-2*, *VEGF*), antiparallel (*BCL-2*), and hybrid (*h-TELO*) G4 structures. **MC-4** was excited at absorbance maxima 285 nm (Figure 3.7) and the emission was obtained from 300 nm to 550 nm (Figure 3.8). The fluorescence emission maxima of **MC-4** were obtained at 341 nm and 356 nm. Fluorimetric titration of **MC-4** (1.0 µM) with increasing concentration of the *c-KIT 1* G4 showed a 5.2-

fold increase (K_d = 1.4 μM) in fluorescence intensity, which is 3.2 times and 2.8 times higher than, *c-KIT 2* (1.6 fold change, K_d = 2.8 μM) and *VEGF* (2.1 fold change, K_d = 2.9 μM) respectively. Negligible changesin fluorescence intensity of **MC-4** were observed for *BCL-2* (K_d = 4.8 μM), *h-TELO* (K_d = 2.3 μM), *c-MYC* (K_d = 2.7 μM) and *ds*DNA (K_d = 4.0 μM). A low dissociation constant (K_d) of 1.4 μM indicated a high affinity for the binding interaction of **MC-4** with *c-KIT 1* compared to other G4s and *ds*DNA (Figure 3.8h).

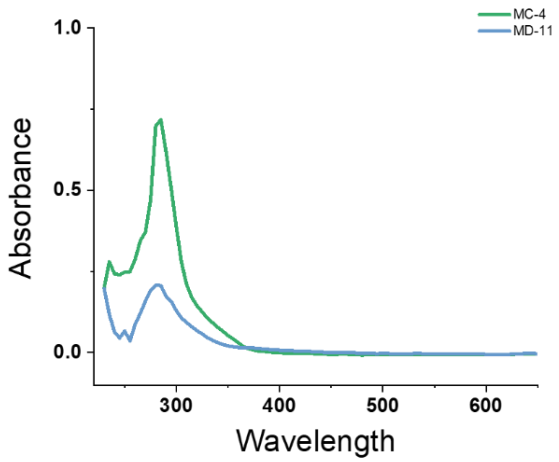


Figure 3.7. The absorbance maxima of **MC-4** and **MD-11** (100 μM) as obtained from UV/Vis spectroscopy.

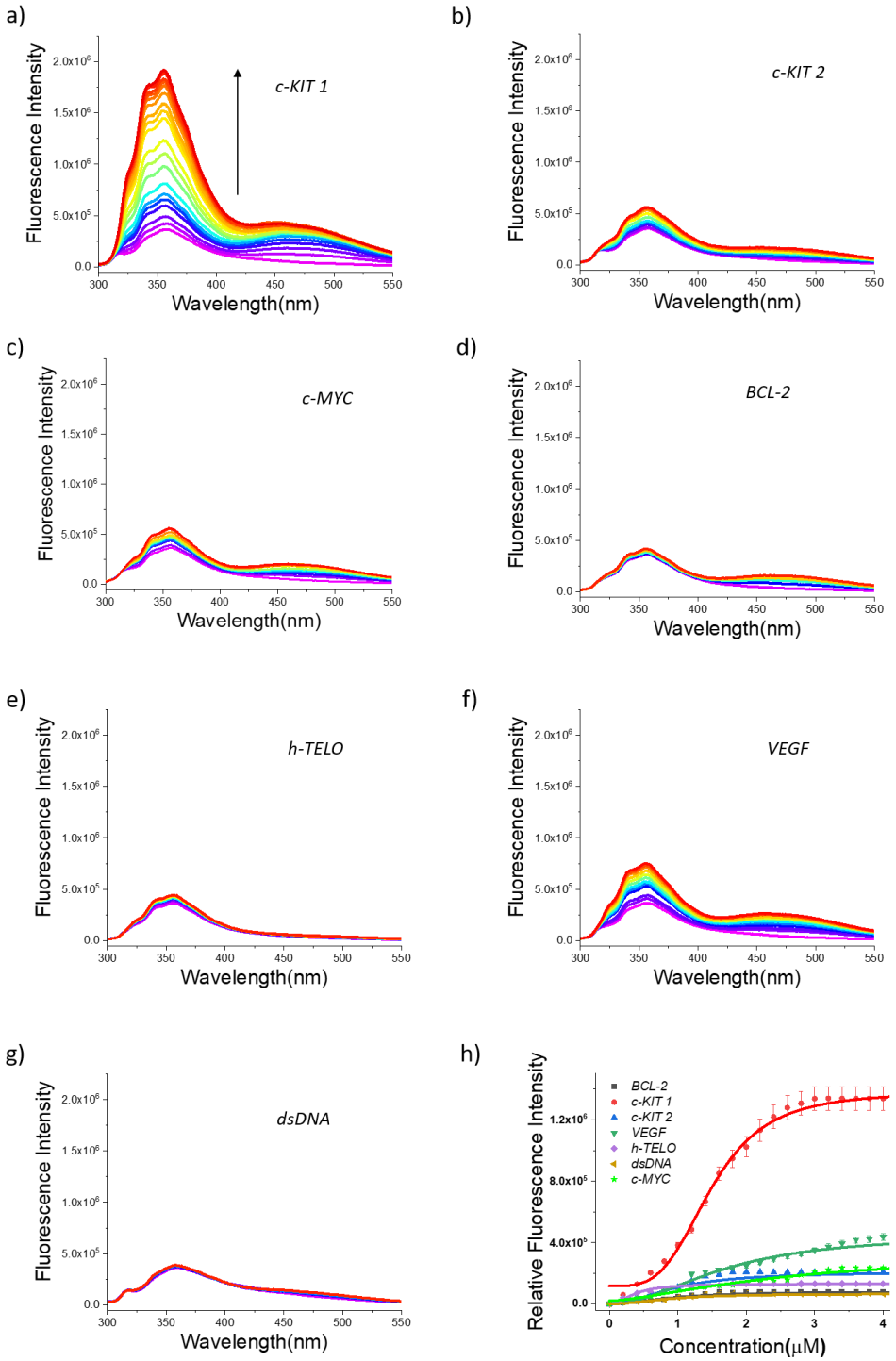


Figure 3.8. Fluorescence spectroscopy of MC-4 titrated with increasing concentrations of different G-quadruplexes, a) *c-KIT 1,b) c-KIT 2,* c) *c-MYC,* d) *BCL-2,* e) *h-TELO*, f) *VEGF,* and g) *ds*DNA; h) Hill plots of the fluorescence titration of MC-4 with different G-quadruplexes. All experiments were carried out in 100mM Tris 100mM KCl buffer (pH-7.4).

Circular Dichroism Spectroscopy

Structural investigations of the folded *c-KIT 1* G4 with the lead ligand **MC-4** were carried out using circular dichroism (CD) spectrophotometry to monitor structural alterations. Different G4 structures display unique CD spectral signatures according to the differences in their topology: parallel (maxima at ~262 nm, minima at~240 nm), antiparallel (maxima at ~290, minima at ~260 min), or mixed (maxima at ~290 and ~262 max, minima at ~240). The CD spectra of *c-KIT 1* exhibited characteristic features of parallel G4 with maxima at 262 nm and minima at 240 nm. Furthermore, CD titration of *c-KIT 1* G4 was performed with excess equivalents of **MC-4** (0.5, 1, 2, 3, 4, and 5 equivalents) to assess the change in the folded structure of *c-KIT 1* (Figure 3.9). It was observed that the significant CD peaks (at 240 nm and 262 nm), were not disrupted even in the presence of excess **MC-4**, suggesting that the carbazole derivative **MC-4** does not purturb the secondary G4 structure of the folded *c-KIT 1* G4.

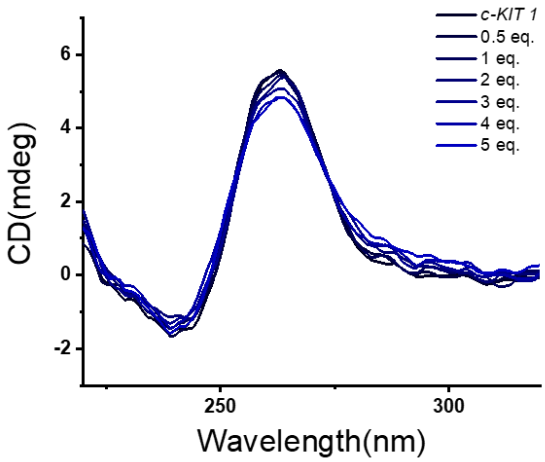


Figure 3.9. Circular dichroism spectra exhibiting the signature peaks of the parallel topology of *c-KIT 1* topology titrated with increasing concentrations (0.5, 1, 2, 3, 4, and 5 equivalents) of **MC-4**. All experiments were carried out in 100 mM Tris 100 mM KCl buffer (pH-7.4).

Isothermal titration calorimetry

Isothermal titration calorimetry (ITC) was performed to probe the thermodynamic binding of the lead compound **MC-4** with different G4s (*c-MYC, BCL-2, c-KIT 1, c-KIT 2, VEGF, h-TELO*, and *ds*DNA). Thermograms were obtained for each G4s titrated with increasing concentrations of **MC-4** added using 30 uniform injections. Each thermograms was obtained at 25 °C. The heat released on the addition of aliquots of **MC-4** gradually decreased to an **MC-4**/DNA molar ratio by titrating increasing concentrations of the ligand **MC-4** with G4 or *ds*DNA. It was observed that **MC-4** exhibited a K_d of 1.6 μM with *c-KIT 1*, which is in agreement with the previously observed K_d value obtained from fluorometric titration. Furthermore, the **MC-4**/*c-KIT 1* interaction was exothermic with a stoichiometry of 1.0 (Figure 3.10).

In contrast, the thermograms obtained for the reaction of **MC-4** with other G4s (Figure 3.11) exhibited high K_d values (*c-MYC* = 6.2 μM, *BCL-2* = 10.8 μM, *c-KIT 2* = 20.9 μM, *VEGF* = 15.9 μM) or unsatisfactory interactions, where K_d values could not be determined (*h-TELO* and *ds*DNA). These findings further corroborate the rest of the biophysical assessments of **MC-4**, implying that this ligand can selectively recognize and interact with the specific G4 topology of *c-KIT 1* with high affinity.

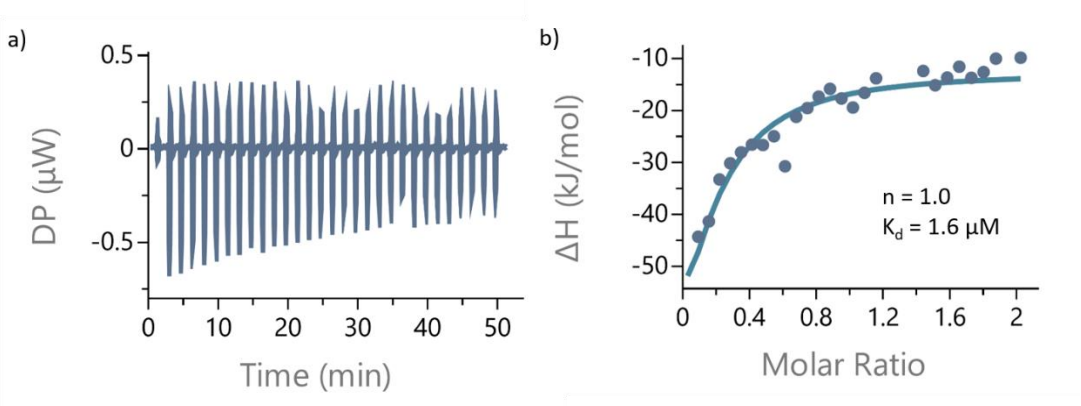


Figure 3.10. a-b) ITC heat curve and binding isotherm of the interaction of **MC-4** with *c-KIT 1* G-quadruplex. The experiment were performed in 100mM Tris 100mM KCl buffer (pH 7.4).

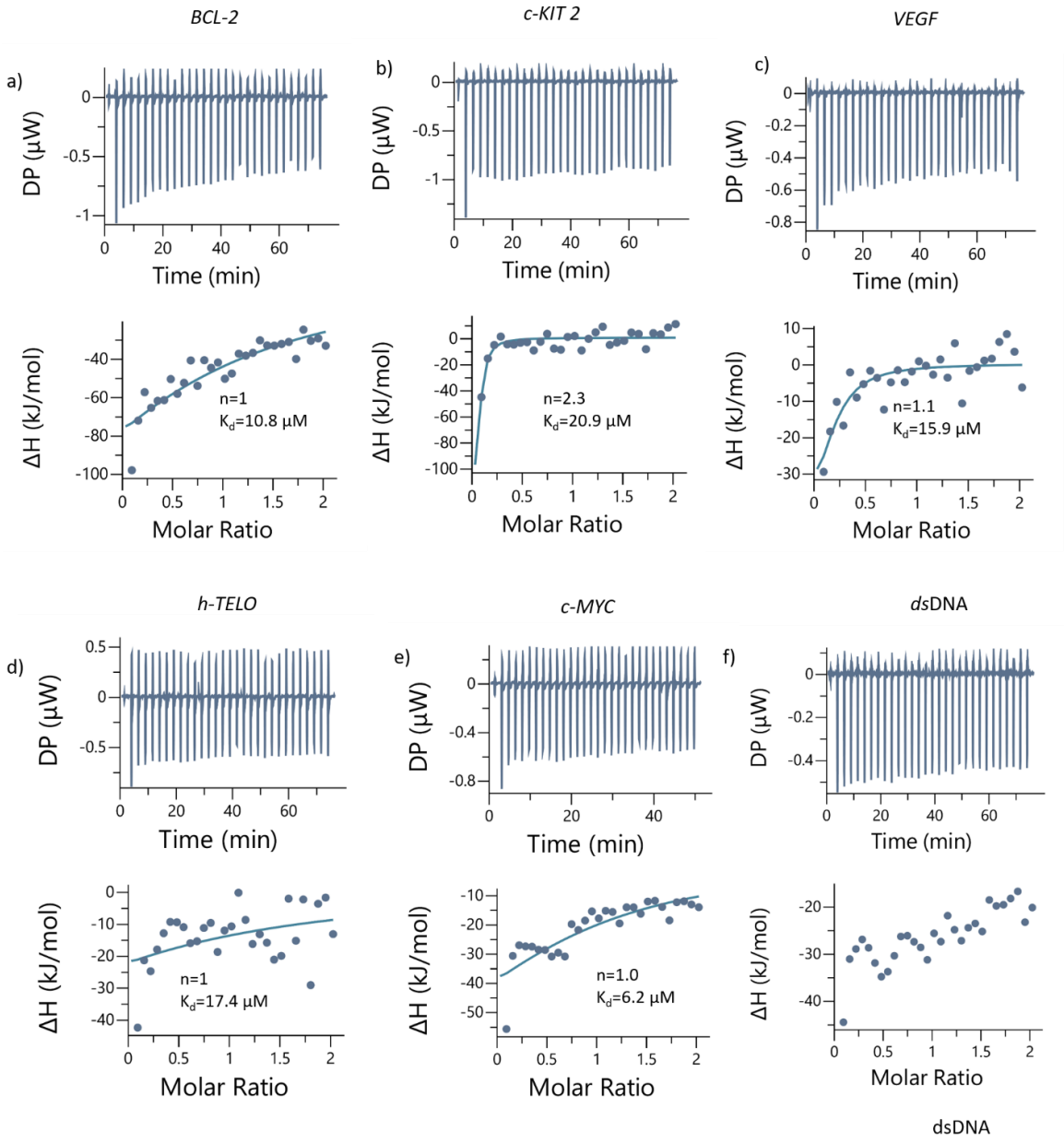


Figure 3.11. ITC heat curves and binding isotherms of the interaction of **MC-4** with a) *BCL-2*, b) *c-KIT 2*, c) *VEGF*, d) *h-TELO*, e) *c-MYC* and f) *dsDNA*. All experiments were performed in 100mM Tris 100mM KCl buffer (pH 7.4).

Investigation of the mode of interaction of MC-4 with G4s

Molecular Docking

Structure-activity relationship (SAR) of the lead **MC-4** with different G-quadruplexes was performed using autodock 4.2 to ascertain the mode of interaction of the **MC-4** with G-quadruplexes *c-KIT 1, c-KIT 2, c-MYC, BCL-2, VEGF*, and *h-TELO*. It was observed that **MC-4** interacted with *c-KIT 1* G4 with a minimum energy (ΔG) of -6.8 kcal/mol, suggesting a strong and energetically favorable interaction (Figure 3.12 a-d). It was found to interact with the 5' G-tetrad composed of G2, G6, G10, and G13 of the *c-KIT 1* G4 (Table 3.1). The flexible amine side chain enabled stronger and more specific interaction with the phosphate backbone of the *c-KIT 1* G4. In contrast, the interaction of **MC-4** with G4s other than *c-KIT 1* was energetically less favorable with ΔG values of (*c-KIT 2* = -5.8 kcal/mol, *BCL-2* = -5.3 kcal/mol, *c-MYC*= -5.6 µM, *VEGF* = -2.5 kcal/mol and *h-TELO* = -1.5 kcal/mol) (Figure 3.13a-e).

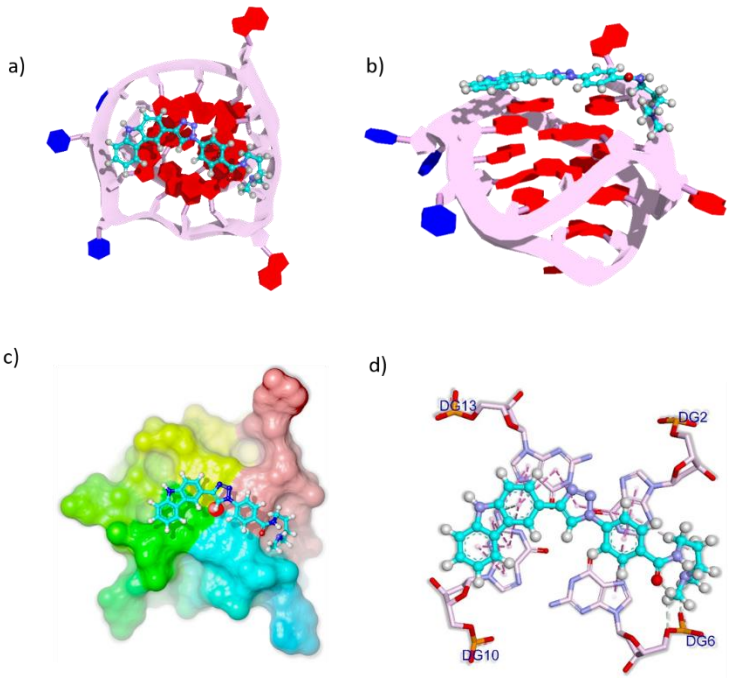


Figure 3.12. Molecular docking studies of **MC-4** (blue) and *c-KIT 1* G4 interaction (PDB ID: 4WO2) as observed from different views, a) Top view, b) Side view, c) Space filled model of *c-KIT 1* G-quadruplex and d) The site of DNA-ligand interaction showing the 5' tetrad guanines where the **MC-4** molecules bind via π-π stacking. Bonds are in dots. **MC4** surrounding base pairs are in two-letter code represented in blue.

Therefore, **MC-4** consisting of a heteroaromatic carbazole scaffold, promotes stacking interaction with the 5' G-tetrad of the *c-KIT 1* G4. The extended side chain structure of **MC-4** further improves the binding affinity *via* interaction with the phosphate backbone of the G4 and the loops. The mono-functionalized smaller side chain enables optimal interaction with the G4, which is often difficult to achieve with larger molecular weight molecules. These studies suggest that **MC-4** stacks upon the 5'terminal G-tetrad of *c-KIT 1*, and its side-chain interacts with the loop region of the G4, resulting in selective recognition of *c-KIT 1* G4 topology over other G4s.

Table 3.1. Bond distances and types of interaction between **MC-4** and *c-KIT 1* G4 DNA.

Molecule Name MC-4	**Distance (Å)**	**Bond category**	**Bond type**
MC4:H25 - :MC4:O	2.46745	Hydrogen Bond	Carbon Hydrogen Bond
MC4:H24 - B:DG6:O5'	2.60689	Hydrogen Bond	Carbon Hydrogen Bond
MC4:H21 - B:DG6:OP1	2.74422	Hydrogen Bond	Carbon Hydrogen Bond
B:DG2 - :MC4	3.82565	Hydrophobic	Pi-Pi Stacked
B:DG10 - :MC4	3.97165	Hydrophobic	Pi-Pi Stacked
MC4 - B:DG10	4.02241	Hydrophobic	Pi-Pi Stacked

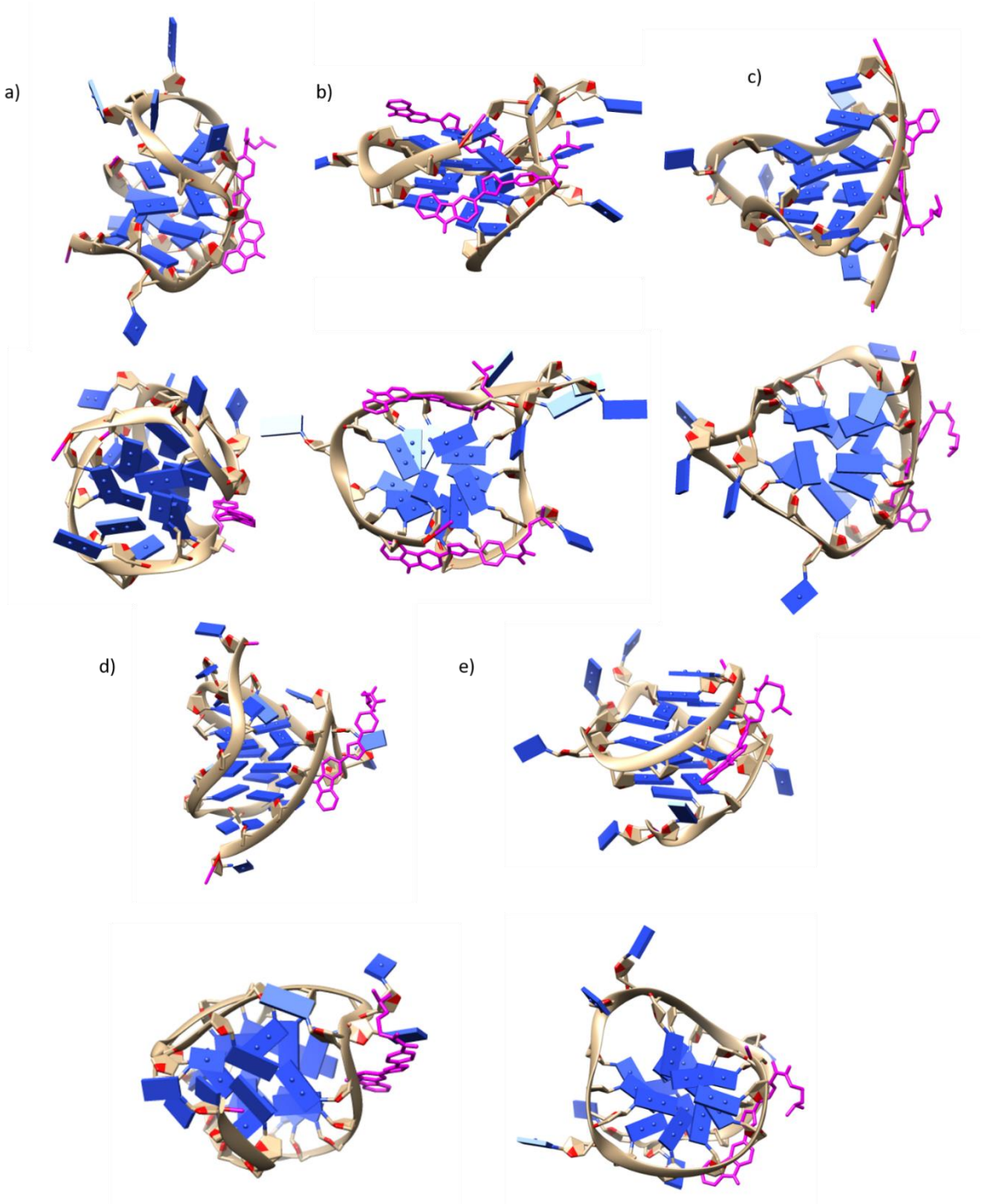


Figure 3.13. Molecular docking studies of **MC-4** (pink) with a) *BCL-2* (-5.32 kcal /mol), b) *c-KIT 2* (-5.8 kcal /mol), c) *VEGF* (-2.52 kcal /mol), d) *h-TELO* G4 (-1.54 kcal /mol) and e) *c-MYC* (-5.6 kcal /mol)

Inhibitory potency of the M-series in cancer cells

Cytotoxicity assay

The biophysical studies carried out with the **M-** series suggest that only **MC-4** selectively interacts and binds with *c-KIT 1* with high affinity. The effect of this interaction was next studied in the cellular context. The cell cytotoxicity of **MC-4** ligand was assessed towards *c-KIT*-positive cancer cell line (K562) which shows overexpression of *c-KIT*. A cancer cell line without expression/overexpression of *c-KIT* was also assessed (MCF-7) as well as a non-cancerous cell line (HEK 293) to determine the cytotoxity of **MC-4** and **MD-11** towards these different cell lines. Myelogenous leukemia cells (K562), and breast cancer cells (MCF-7) and were treated with increasing concentrations of **MC-4** and **MC-11**. It was observed that **MC-4** displays cytotoxicity towards K562 but not MCF-7 with IC_{50} values of 6.2 µM, and 19.8 µM respectively (Figure 3.14a).

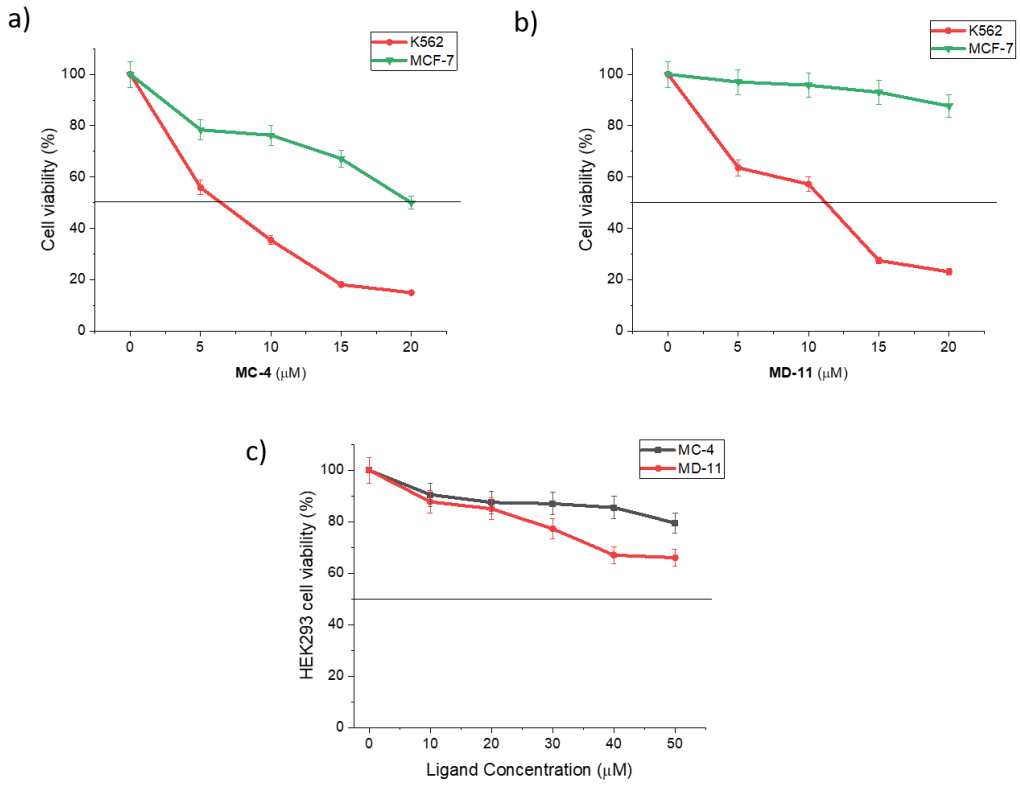


Figure 3.14. Cell cytotoxicity analysis of a) **MC-4**, b) **MD-11** in different cancer cell lines after 24 h, c) Cell cytotoxicity analysis of **MC-4** and **MD-11** in HEK293 cell line after 24h.

In addition, the ligand **MC-4** was non-toxic to normal HEK293 cells with an IC_{50} value of >50 μM (Figure 3.14c). This is in line with our previous observations which show interaction of **MC-4** with *c-KIT 1* G-quadruplex. In contrast, the dibenzofuran scaffold containing counterpart **MD-11** displayed higher IC_{50} values (K562 = 11.7 μM, μM MCF-7 = >19 μM, and HEK293 = >50 μM) (Figure 3.14b). The low IC_{50} value of **MC-4** exclusively in *c-KIT* overexpressing K562 cancer cells line suggest that **MC-4** can effectively penetrate the cell and nuclear membranes to stabilize *c-KIT* G-quadruplex in cancer cells while inducing a cytotoxic effect only in leukemia cellline which highly overexpresses *c-KIT* protein without inducing cytotoxicity in *c-KIT* negative cancer cells.

Amplification of cellular G4s by MC-4

As **MC-4** exhibited an anti-proliferative effect on cancer cells, its ability to penetrate the cellular membrane was visualized using cell imaging investigations. In addition, **MD-11** treated K562 cells were also assessed. K562 cells were treated with 3 μM of **MC-4** and 5.5 μM of **MD-11** and observed under a confocal microscope. Both **MC-4** and **MD-11** showed an inherent fluorescence in the green filter range characteristic of carbazole and dibenzofuran compounds. It was observed that both **MC-4** and **MD-11** are cell permeable and localizes in the cells.

BG4, a monoclonal single-chain antibody developed through phage display, exhibits exceptional affinity and specificity for G4 structures and has become a valuable tool in numerous studies aiming to probe the existence of G4 structures within the genome. In mammalian cells, immunostaining with BG4 has revealed the presence of G-quadruplex structures under various conditions. Notably, live cells treated with the G4-binding small molecules prior to immunostaining exhibit a significant increase in the number of BG4 foci (Figure 3.15). Thus, ligand treated K562 cells were observed under the confocal microscope to evaluate the presence and prevalence of G-quadruplexes in this study.

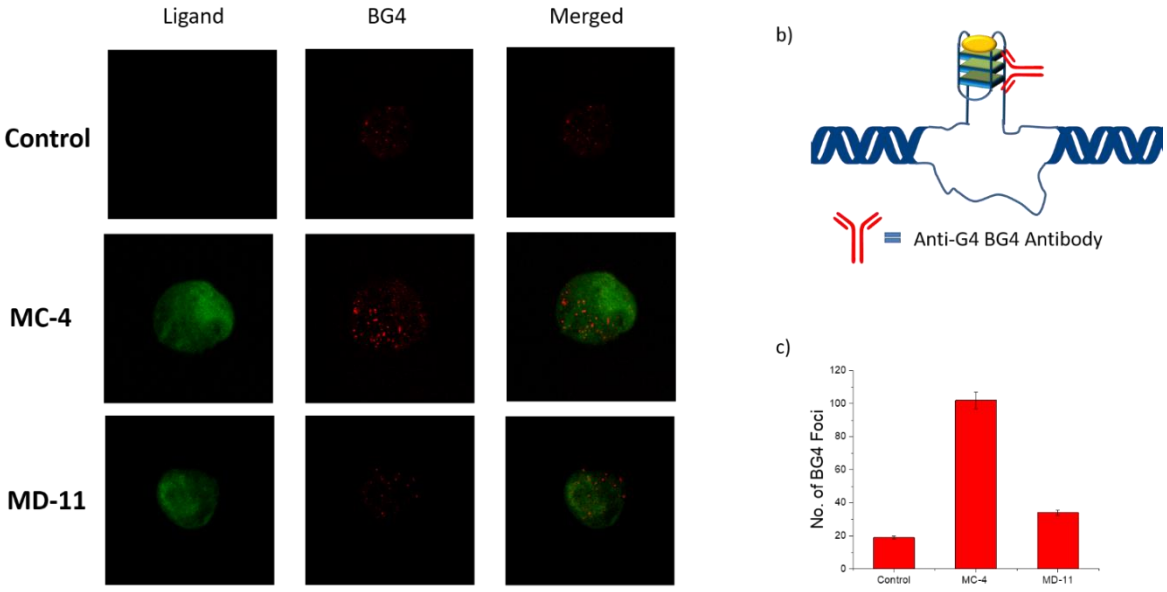


Figure 3.15. The stabilization of cellular G-quadruplexes by specific G4 –interacting small molecules results in detection of a greater number of G4s visualised by the G4- specific antibody BG4.

MC-4-treated cells showed a more significant number of G-quadruplexes visualized using the G-quadruplex specific antibody (BG4). The number of BG4 foci increased by 5.4 fold and 1.8 fold in the **MC-4** and **MD-11** treated cell compared to the number of foci observed in untreated (Figure 3.16a-c and Figure 3.17) and the **MD-11** treated cells respectively. This direct evidence suggests that the treatment of **MC-4** leads to a greater number of G-quadruplex stabilization, resulting in their detection in significant numbers inside the cells.

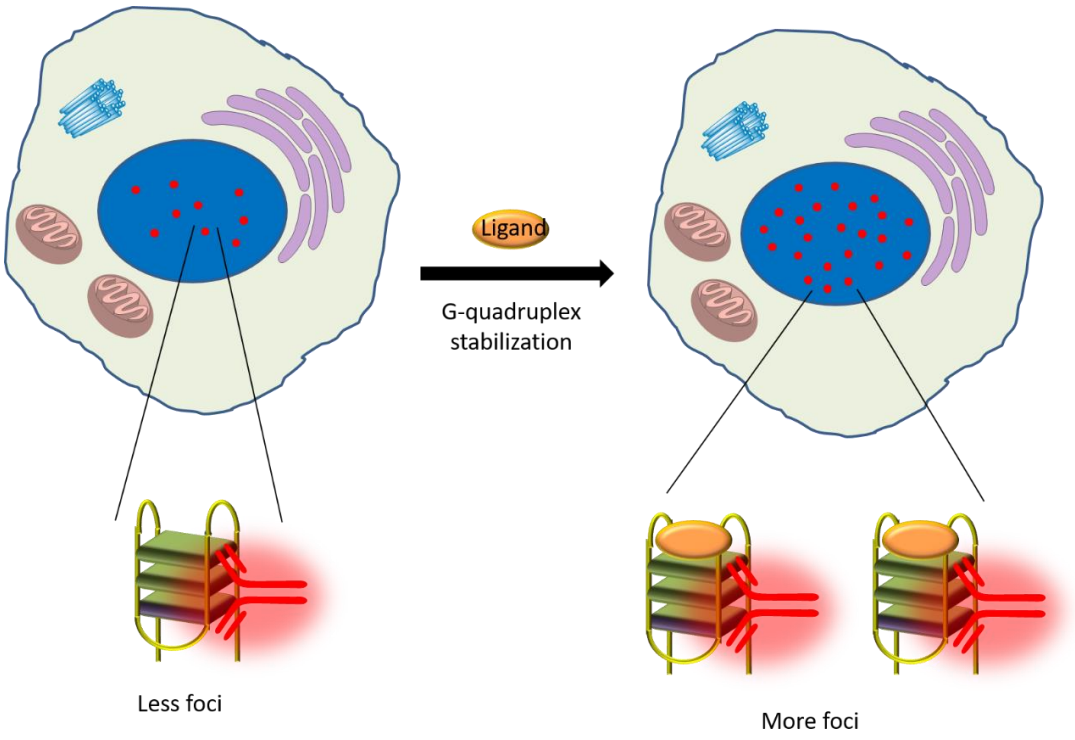


Figure 3.16. a) Cell images of K562 with and without treatment of **MC-4** and **MD-11**. The cell has been stained with **MC-4** and **MD-11** (green) and the G4 specific antibody is visualized as foci (red) present within the cell. b) Schematic representation of **MC-4** stabilized G-quadruplex and BG4 antibody interaction. c) Graphical representation of the number of BG4 foci in the nucleus of untreated and treated K562 cells.

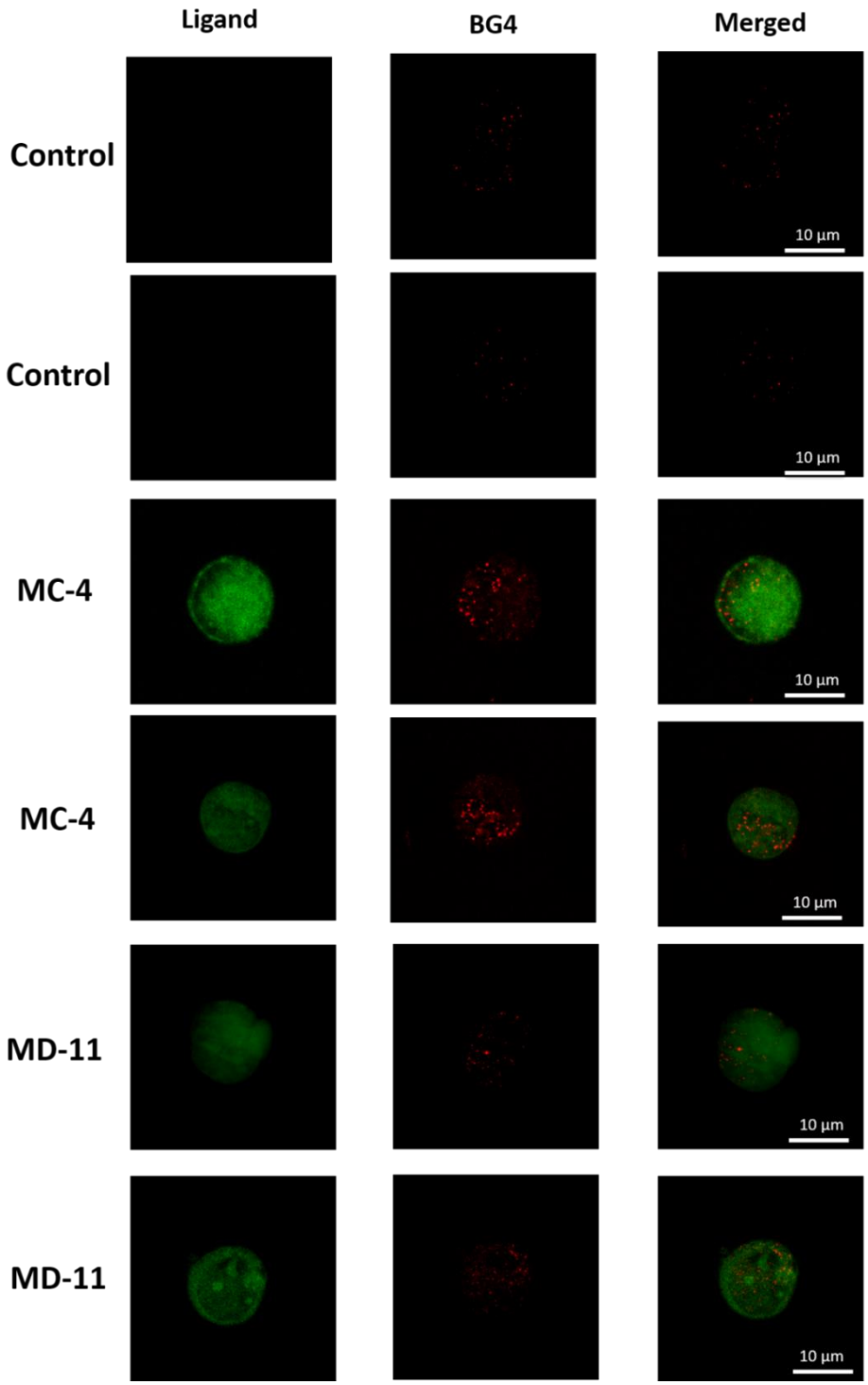


Figure 3.17. Fluorescence microscrope images of **MC-4** (3 µM) and **MD-11** (5.5 µM) treated K562 cell line showing localization of **MC-4** (green) and **MD-11** (green) in the cells. The red puncti of BG4 antibody indicate presence of G-quadruplex structures.

MD-4 induced gene regulation of *c-KIT*

mRNA quantification

To gain more conclusive insights into the effect of the ligands in the context of cellular machinery and gene regulation, gene expression studies were undertaken. qRT-PCR was performed to examine the effect of **MC-4** on *c-KIT* transcriptional regulation (Figure 3.18 and Table 3.2). In addition **MD-11** was taken as a negative ligand and the well-explored **TMPyP4** was taken as a positive control to obtain a broader perspective for the gene regulation studies. K562 cells were treated with two different doses of **MC-4** (3 µM and 6 µM), **MD-11** (5.5 µM and 11 µM) and **TMPyP4** (2 µM and 4 µM) for 24 hrs, and the expression of each gene *(cKIT, c-MYC* and *BCL-2)* were normalized with the expression of housekeeping gene *18s*. **MC-4** led to the reduction of the *c-KIT* oncogenic expression by 28% and 73% at 3 µM and 6 µM concentrations, respectively. In addition, treatments with the negative ligand **MD-11** induced a downregulation of >1% and 3%, whereas the positive control **TMPyP4** induced a downregulation of 30% and 48% respectively.

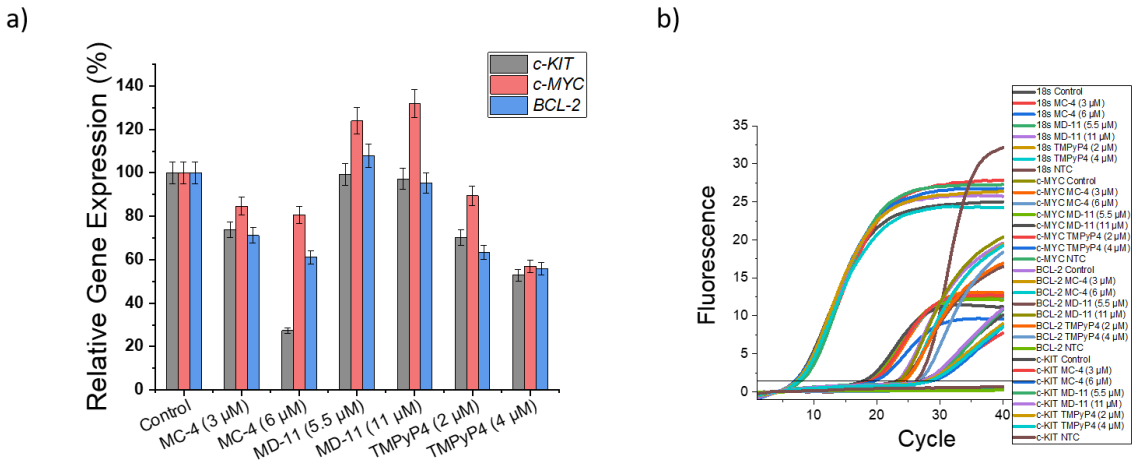


Figure 3.18. a) Determination of gene regulation at the level of mRNA in K562 cells treated with different doses of **MC-4** (3 and 6 µM), **MD-11**(5.5 and 11 µM) and **TMPyP4** (2 and 4 µM) using qRT-PCR to measure the expression of *c-KIT*, *c-MYC* and *BCL-2* genes; b) Fluorescence sigmoidal curves of real-time PCR measuring the amplification of SYBR Green.

Additional genes of *c-MYC* and *BCL*-2 were also studied and showed a lesser decrease in gene expression with downregulation of 15% and 19% in *c-MYC* gene and a decrease of 29% and 39% in *BCL-2* after treatment with **MC-4**. A significant decrease of gene expression of *c-KIT* associated with anti-apoptotic *BCL-2* gene suggests a relation between the stabilization of *c-KIT* G-quadruplex by **MC-4** and apoptosis. However, *c-MYC* showed an upregulation of 23% and 31% and *BCL-2* showed 23% upregulation and 5% downregulation respectively after treatment with **MD-11**. Expectedly, the G4 binder **TMPyP4** (2 and 4 μM) showed a dose dependent downregulation of 19 and 43% in *c-MYC* and a downregulation of 57% and 45% in *BCL-2* gene expression. This suggests that the carbazole derivative **MC-4** can efficiently reduce the transcription of *c- KIT* oncogene and lead to a downregulation of up to 73%.

Table 3.2. Cycle threshold (***C*T**) values:

	18s	***c-MYC***	***BCL-2***	***c-KIT***
Control	7.18	18.89	23.54	27.57
MC-4 (3 μM)	7.65	19.60	24.50	28.48
MC-4 (6 μM)	7.18	19.20	24.25	29.44
MD-11 (5.5 μM)	7.60	19.00	23.85	28.00
MD-11 (11 μM)	7.47	18.78	23.90	27.90
TMPyP4 (2 μM)	7.48	19.35	24.50	28.38
TMPyP4 (4 μM)	7.70	20.22	24.90	29.01

Promoter Luciferase assay

For further validation into the binding interactions with G-quadruplexes and their effect on gene expression reporter luciferase and dual-luciferase reporter assays were performed with different genes. Plasmid constructs containing wild type promoter sequences of the genes (*c-KIT*, *c-MYC*, *BCL-2*) with G4 forming sequences located upstream of the renilla (*c-KIT* construct) and firefly (*c-MYC* and *BCL-2* constructs) coding genes were used for the study. Transfected K562 cells were further treated with different

doses of **MC-4**, **MD-11** and **TMPyP4**. The luciferase expression was obtained for the ligand treated transfected cells and normalized against the RLTK expression or estimated protein (Figure 3.19a-b). **MC-4** treated cells exhibited *c-KIT* downregulation in a dose-dependent manner with 22% and 69 % decrease at 3 μM and 6 μM respectively. In contrast, **MD-11** showed only 3.3% and 7.6% of decrease at 5.5 μM and 11 μM concentrations and the positive control treated with **TMPyP4** showed a decrease of 20.12% and 34.08% at the concentration of 2 and 4 μM respectively. In addition, much smaller degree of changes were observed in *c-MYC* and *BCL-2* constructs transfected cells treated with **MC-4** (*c-MYC* = 6.0 and 12.3%, *BCL-2* = 7.8 and 15.8% and **MD-11** (*c-MYC* = 3.2 and 10.1%, and *BCL-2* = 9.5 and 14%) at both their respective dose concentrations. **TMPyP4** treated cells all showed downregulation (*c-MYC* =34.4 and 45.3%, and *BCL-2* = 20.7 and 39.7%) suggesting that the G4 binding **TMPyP4** downregulates the expression of all promoter luciferase vectors by stabilizing the G-quaduplexes in the promoter regions of the oncogenes. These observations further corroborates that ligand **MC-4** induces downregulation of *c-KIT* gene expression in cancer cells whereas it does not induce any significant changes in theexpression of *c-MYC*, and *BCL-2* luciferase constructs by interacting with their respective G-quadruplexes.

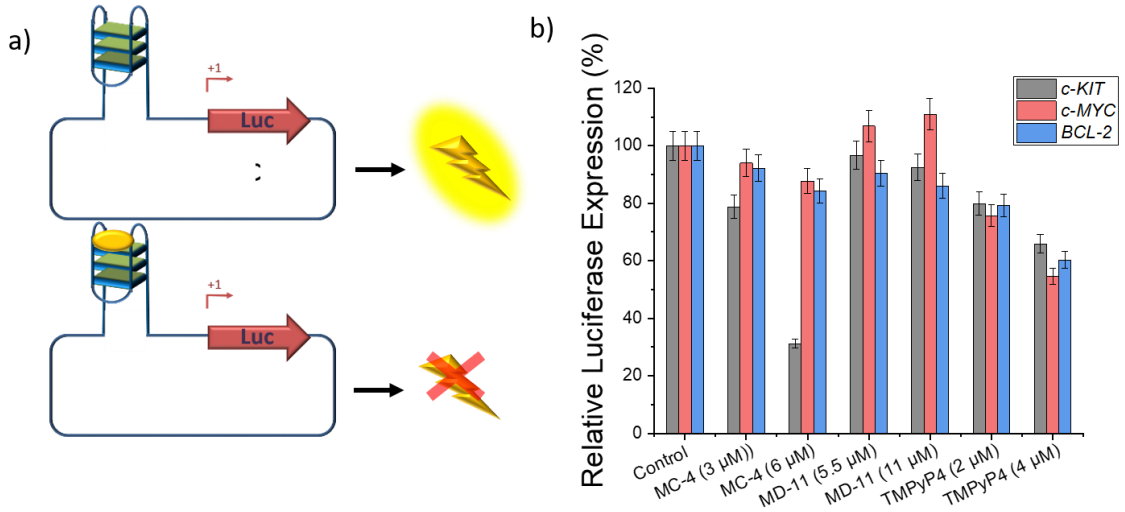


Figure 3.19. a) Schematic representation of the promoter luciferase assay using plasmid constructs containing the G-quadruplex forming sequence in the promoter region which

can be stabilized with small molecules to decrease the luciferase expression; b) The relative luciferase expression of *c-KIT c-MYC and BCL-2* G4 containing luciferase-expressing plasmids transfected in K562 cells treated with different doses of **MC-4**, **MD-11** and **TMPyP4**.

Western Blotting assay

Western Blotting assay was performed to confirm the results of *cKIT* gene inhibition by the lead ligand. **MC-4** treated K562 leukemia cells showed a dose dependent downregulation of *cKIT* with 23.3 % and 65.5 % expression at 3 and 6 µM (Figure 3.20). Negative control ligand, **MD-11** treated cell showed very small changes in the expression of *c-KIT 1* with a small downregulation of 7.8% at 5.5 µM dose and an upregulation of 15.8 on treatment with a higher dose (11 µM). Expectedly, **TMPyP4** showed a downregulation of *c-KIT* (~10 and 20%) suggesting that it inhibits the protein expression of *c-KIT* gene but at a much lesser extent .

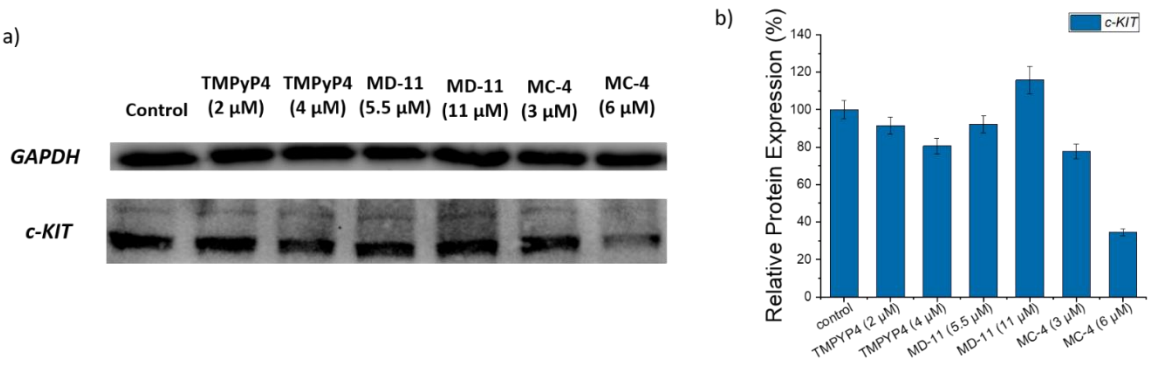


Figure 3.20. a) Protein expression of housekeeping gene *GAPDH* and gene of interest *c-KIT* in K562 cells treated with different doses of **MC-4**, **MD-11** and **TMPyP4** over a period of 24h; b) Bar diagram of relative protein expression of *c-KIT* gene in K562 cells after treatment with respective doses of **MC-4**, **MD-11** and **TMPyP4**.

MC-4 triggers cell cycle arrest, DNA damage and apoptosis in cancer cells

Cell cycle assay

Flow cytometric studies were conducted to investigate the relationship between the inhibitory effects of **MC-4** and the possible arrest of cancer cells in any phase of the cell cycle. **MC-4** (3 μM and 6 μM) and **MD-11** (5.5 μM and 11 μM) treated K562 cells were stained with propidium iodide (PI) to determine the number of cells in each phase of the cell cycle (Figure 3.21). **MC-4**-treated cells exhibited cell cycle arrest in a dose-dependent manner. The cells were arrested at important checkpoints of G0/G1. At 3 μM **MC-4**, an increase in cell population was observed in G0/G1 phase with an arrest of 74.1% of cells compared to 68.6% under control conditions. At a higher dose (6 μM) the cells arrested at G0/G1 increased to 79.7%. This suggests that **MC-4** arrests the cancer cells in G0/G1 phase of the cell cycle. This observation supports current literature, where G4 binding ligands are observed to arrest cells at G1 phase of the cell cycle.[154, 155] In addition, the **MD-11** treated cells did not show any significant changes in the cell populations in any cycle phase.

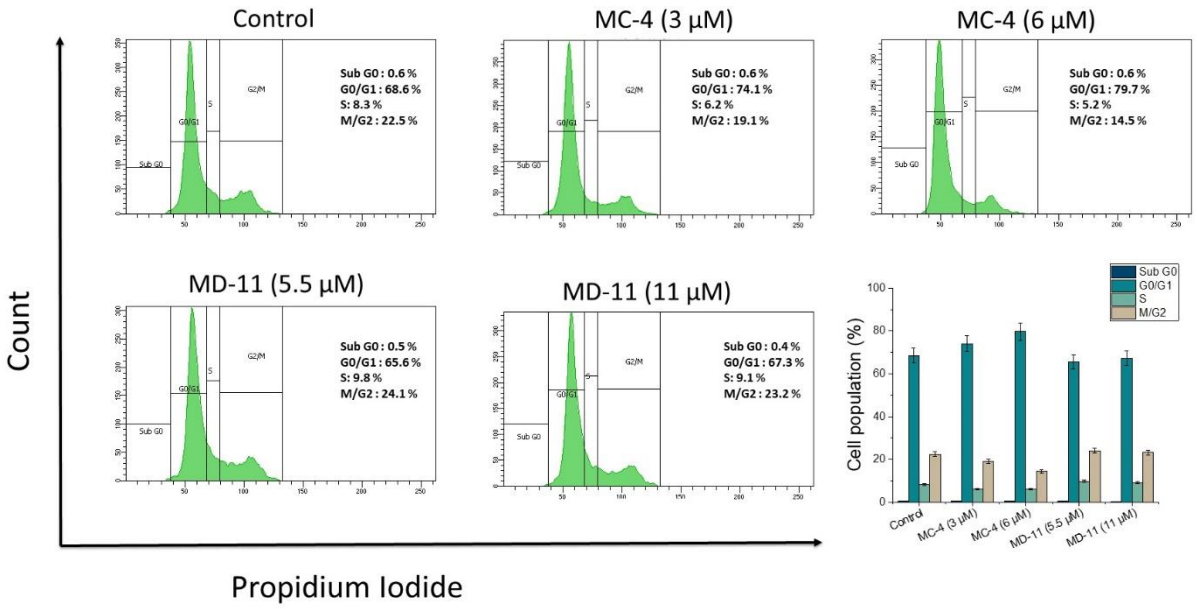


Figure 3.21. a) Flow cytometric analysis of cell cycle parameters after treatment of K562 cells with **MC-4** (3.0 and 6.0 μM) and **MD-11** (5.5 and 11 μM) and bar diagram representing the percentage of cells in each cell cycle phase after treatment with M-series of ligands.

DNA Damage assay

Additional flow cytometric studies to evaluate ligand-induced DNA damage were conducted using Alexa-fluor 645 tagged γH2AX antibodies (Figure 3.22). Flow cytometric detection of γH2AX in cells is a reliable method for quantitation DNA damage. Analysis of **MC-4** (3 and 6 μM) treated K562 cells showed DNA damage in 38.6% and 67.1 % of the cell population. In contrast, **MD-11** (5.5 and 11 μM) treated cells did not show an appreciable DNA damage (27.6 and 30.3%).

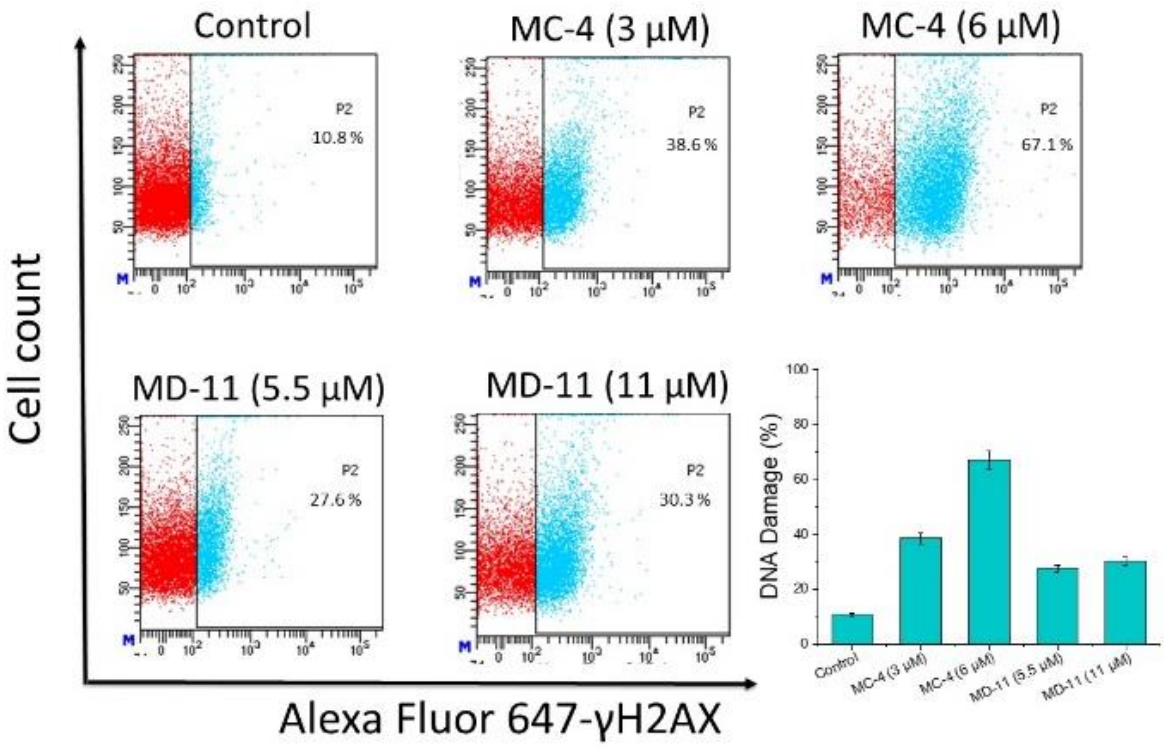


Figure 3.22. a) Flow cytometric analysis of DNA damage after treatment of K562 cells with **MC-4** (3 and 6 μM) and **MD-11** (5.5 and 11 μM); b) bar diagram representing the percentage of cells in each cell cycle phase.

Apoptosis assay

Flow cytometry was carried out to investigate the mode of death induced by in cancer cells the carbazole and dibenzofuran derivatives, **MC-4** and **MD-11** respectively (Figure 3.23). K562 cells treated with 3 μM and 6 μM of **MC-4** and 5.5 μM and 11 μM of **MD-11** were analyzed upon Annexin V and PI dual staining. Dot plot analysis indicated.that **MC**-4 induced dose dependent apoptosis in K562 cells in a dose-dependent manner, with increasing rates of apoptotic cells observed at both doses of ligand concentrations. At 3 μM concentration, **MC-4** induces 35 % of cell death whereas, at the second dose of 6 μM, an increase of apoptotic cells to 63% was observed with a more significant shift from early to late apoptosis. On the other hand, **MD-11** induced apoptosis in 1 and 11% of the cell population at 5.5 and 11 μM concentrations respectively.

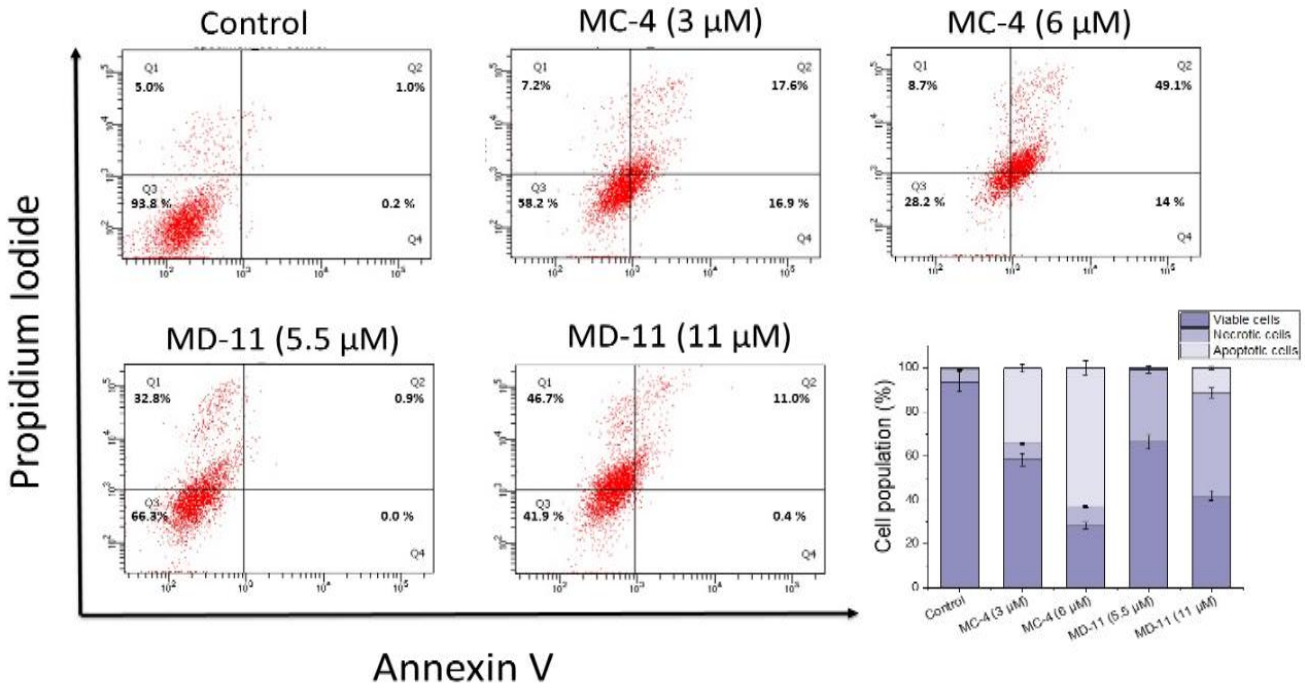


Figure 3.23. Flow cytometric analysis of the mode of cancer cell death after treatment with **MC-4** (3.0 and 6.0 μM) and **MD-11** (5.5 and 11 μM); and bar diagram representing the percentage of apoptotic, necrotic and viable cells in control and ligand treated K562 cells.

Overall, the flow cytometric studies revealed that **MC-4** containing the carbazole is capable of arresting the K562 leukemia cells in G0/G1 phase of the cell cycle. It also triggers DNA damage and leads to apoptosis in cells likely by interacting and stabilizing the *c-KIT 1* G4.

Conclusion

In summary, we have synthesized a library of 14 small molecules consisting of carbazole (**MC**) and dibenzofuran (**MD**) scaffolds with a series of identical side chain modifications The aim was to explore the role of fused heteroaromatic ring systems and side chain modifications in G4 recognition and selective stabilization. Our findings revealed that both the heteroaromatic scaffold and side chain are crucial for recognizing and binding to specific G4 structures, with predominant binding being driven by the aromatic scaffold. However, the initial G4 binding by the **M**-series makes way for more selective recognition with emphasis on the role of the side chain modification. Despite containing the same carbazole moiety, among the **MC** molecules, only **MC-4** emerged as a lead ligand capable of interacting with *c-KIT 1* G4 (Figure 3.24). In contrast, the **MD** molecules did not show any ability to stabilize the studied G4s. Although **MC-4** and **MD-11** contain the same carboxamide side chain and differ only in the heteroaromatic ring system, the comparative biophysical studies revealed **MC-4** to be the lead ligand for *c-KIT 1* G4 over the other molecules of the **M**-series. The biophysical analyses further corroborate the cytotoxic profiles of both compounds towards *c-KIT* positive and negative cancer cells, suggesting one heteroaromatic scaffold's ability to selectively stabilize a G4 over the other scaffold.

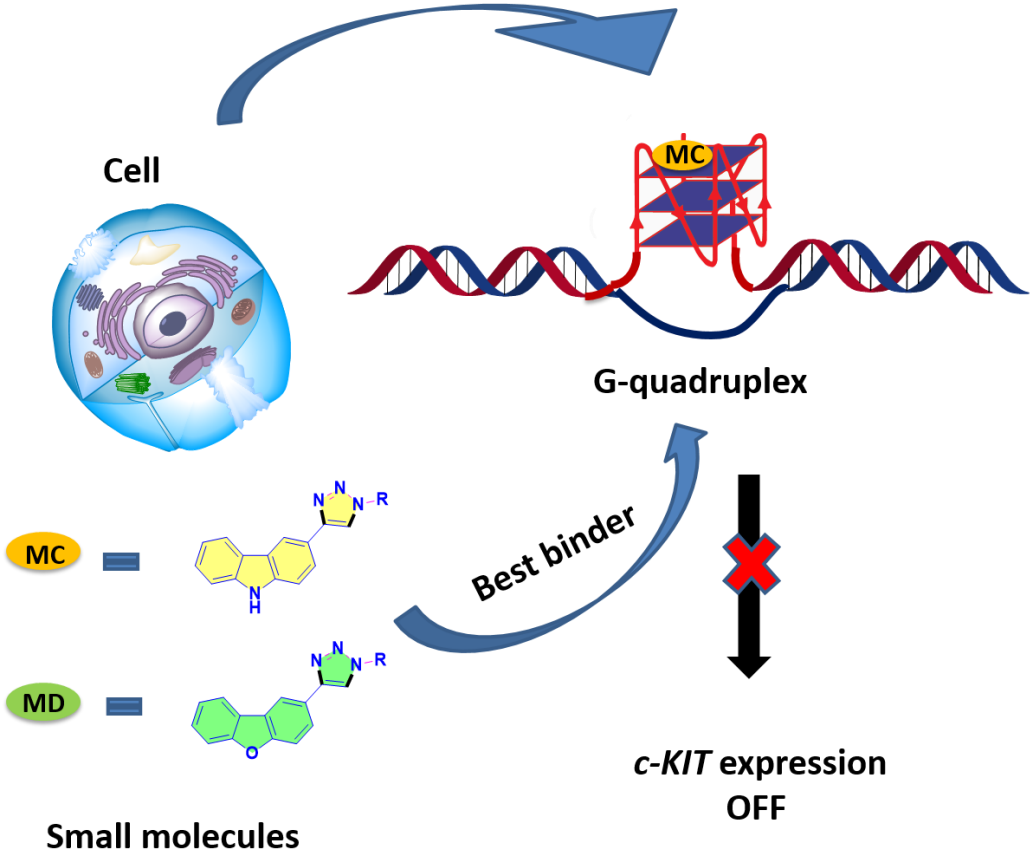


Figure 3.24. Schematic representation of the binding of the best binder, **MC-4** which is a carbazole derivative containing a carboxamide side chain and stabilizes the *c-KIT* G-quadruplex and induces downregulation of c-KIT gene expression.

As observed using biophysical assays of FRET, fluorescence spectroscopy, and ITC, **MC-4** selectively interacts *with c-KIT 1* G4 with a low K_d value and high stabilization compared to other G4s and *ds*DNA. In addition, confocal microscopy and flow cytometric studies show that MC-4 can easily enter cancer cells, arrest the cells in G0/G1 phase, trigger DNA damage and induce apoptosis. Furthermore, **MC-4** treated K562 cells contain more

number of G-quadruplex structures as visualized by the G4-specific antibody BG4. This indicated the presence of a higher number of G-quadruplexes compared to untreated cells. **MC-4** can modulate gene expression inside cancer cells, and decrease the transcription and translation of the *c-KIT* gene studied by mRNA quantification and promoter luciferase assay. Therefore, the modulation of *c-KIT* oncogene is achieved by the selective interaction of **MC-4** with the *c-KIT 1* G-quadruplex, whereby the aromatic planar scaffold of the ligand promotes stacking on 5' G-tetrad of the G4 and the positively charged carboxamide side chain enhances electrostatic interaction with the phosphate backbone of the G4. The findings of this study further subtantiates a better binding of G4 by carbazole containing scaffold compared to debenzofuran containg ones. This enhanced interaction maybe due to the presence of a nitrogen atom in the heterocyclic scaffold of carbazole which is protonated at physiological conditions. The findings also indicate that optimization of side chain for carbazole compounds can lead to selective binding with cellular G4s based on differential recognition of specific G4 topology. Fine-tuning side chains of protonable scaffolds further taps into the possibility of rationally designing synthetic carbazole derivatives to enhance targeting of biomolecular G4s which demonstrate variable topologies. The study also paves the way for future functionalization of carbazoles to achieve a higher degree of selectivity for different DNA conformations to modulate the expression of cancer and other faulty gene expressions.

Chapter 4

Structural and functional characterization of Novel DNA and RNA G-quadruplexes in the *mTOR* gene

Introduction

The mammalian target of rapamycin (*mTOR*), is an evolutionarily conserved serine/threonine kinase that integrates signals from growth factors, nutrients, and stresses to regulate multiple processes like mRNA translation, cell-cycle progression, autophagy, and cell survival (Figure 4.1).[156, 157 158-160] mTOR was initially discovered in the early 1990s during investigations into the mechanism of action of rapamycin (sirolimus), a macrolide that was originally found as an antifungal agent and was later recognized as having immunosuppressive and anticancer properties.[161, 162] The *mTOR* signaling pathway has been involved in multiple diseases, including cancer,[163, 164] neurodegeneration and diabetes,[165] emphasizing the importance of identifying and understanding the function of the components of the *mTOR* gene and the *mTOR* signaling network.

The mTOR signaling pathways are critical for maintaining cell homeostasis through regulation of various biological functions, such as cell growth, metabolism, survival, and immune response. These functions are mainly regulated by mTOR, primarily through the formation of two distinct complexes, namely, mTOR complex 1 (mTORC1) and mTOR complex 2 (mTORC2).[159, 166-168] This signaling pathway is commonly activated in tumors and it controls cancer cell metabolism by altering the expression and/or activity of a number of key metabolic enzymes.[169, 170] The compelling evidences linking activated *mTOR* signaling to cancer has generated significant interest in developing *mTOR* inhibitors for targeted cancer therapy or chemotherapy.[171, 172] Rapamycin and its analogs (rapalogs) have been reported as specific allosteric inhibitors of *mTOR* and represent first-generation *mTOR* inhibitors.[173, 174] Rapalogs like everolimus/RAD001 and temsirolimus/CCI-779 are approved by the Food and Drug Administration (FDA) for treatment of certain cancers such as metastatic renal cell carcinoma, pancreatic neuroendocrine tumors, and postmenopausal hormone receptor-positive advanced breast cancer.[175] Apart from mTOR, the other best studied targets for inhibiting *mTOR* signalling pathway are the cell growth effectors S6K1 and 4E-BP1.[176, 177] Great efforts have been made to develop second-generation ATP competitive *mTOR* kinase inhibitors (TORKinibs) such as INK128, Torin 1, and AZD8055, which suppress mTORC1 and mTORC2 activity, as potential cancer therapeutic agents.

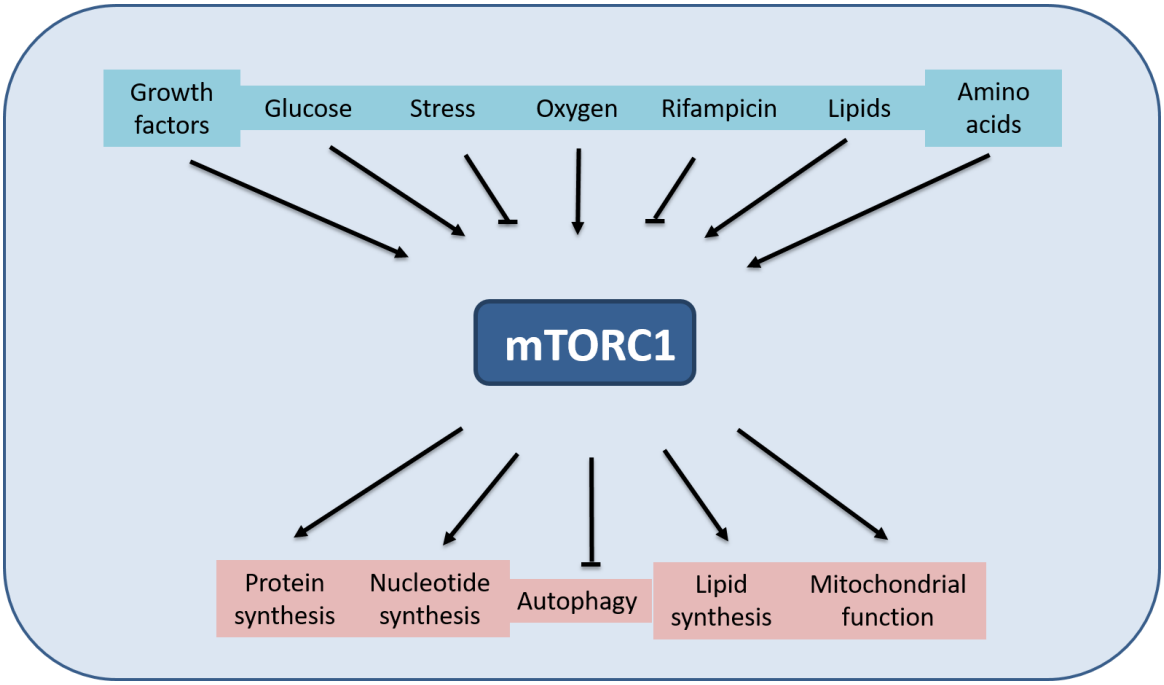


Figure 4.1. Schematic representation of the cellular stimuli that mediated the function of mTOR complex 1 to regulate different cellular processes.

However, rapalogs, have been shown to have little efficacy and use of second-generation TORKinibs might be limited by more severe toxicities.[171, 178] Given the prominent role of *mTOR* in cancer drug resistance and the yet limited clinical success with *mTOR* inhibitors, we aim to identify novel therapeutic options for cancer cells having mTOR-dependent drug resistance. However, most of the therapeutic studies have been focused on the targeting of *mTOR* protein itself or in conjunction with its complex forming proteins.[179-186] Regulation of mTOR gene expression can be considered as an attractive alternative approach for *mTOR* inhibition. However, the number of studies carried out to control *mTOR* gene expression at the DNA and RNA level remains scarce. Secondary structures of DNA are implicated to have high tunability in gene expression *via* modulation of DNA conformation, leading to changes in transcription and translational machineries of the cell.[187] Based on these facts, we envisage to investigate the presence of any DNA/RNA

secondary structures in the promoter or untranslated domains of *mTOR* gene, which may present an attractive alternative approach for *mTOR* inhibition.

G-quadruplexes (G4) are widely present throughout the human genome especially in the promoters,[187, 188] telomeres,[88, 189, 190] UTRs,[139, 191] intronic regions of genes involved in development, survival and proliferation.[192, 193] These four-stranded nucleic acid structures are also increasingly being recognized as regulatory elements for replication, transcription, and other important regulatory roles.[194] Computational studies estimate that over 3,00,000 sequences in the human genome may potentially fold into a G4 structure.[13, 19, 195] Owing to the G4s' abundant presence and significant biological role in the genome, the G-quadruplex is considered an excellent target for chemical intervention of gene expression using small molecules.[49, 196-198] Bioinformatic studies show that *mTOR* has several stretches of GC-rich sequences in the promoter/UTR region and thus presents a promising domain for the search of potential G-quadruplex folding sequences. As the putative sequence rule ($G_{\geq 3}N_{1-7}G_{\geq 3}N_{1-7}G_{\geq 3}N_{1-7}G_{\geq 3}$)[195] are almost absent in the *mTOR* promoter and 5'-UTR regions, the G-quadruplexes have not been identified yet in *mTOR* gene. Therefore, analysis of a primary sequence of mTOR gene (NCBI accession number: NG_033239.1) on chromosome 1 along with a 1500 bp stretch present at the upstream region, 5' UTR and coding region of the human mTOR gene was undertaken to detect Putative G-quadruplex Sequences (PQSs) that could potentially tune the transcriptional machinery of *mTOR* gene expression. The presence of multiple G-rich tracts in the regulatory regions were observed. In this report, the presence of a potential G-quadruplex forming sequence in the upstream region of mTOR mRNA and its functional implications have been discussed in detail.

Results and Discussion

4.1. G-rich sequence of *mTOR* DNA and 5'-UTR region of *mTOR* mRNA folds into stable intramolecular G-quadruplex structure under influence of K^+

Narrowing down the search for G4s in *mTOR* gene

Owing to multiple reports citing the G-quadruplex occurrence in promoter regions and UTRs of multiple genes being relevant in the biological functions and their regulation, a search to identify potential PQSs in the genetic sequence of mTOR gene (NCBI accession number: KJ399980.1 and NG_033239.1) was conducted. The *mTOR* gene, promoter and mRNA sequence were obtained from NCBI and exon/intron sequence was assigned from Ensembl. The QGRS mapper[199-201] software was used to identify the PQSs in the *mTOR* gene. Though the entire genetic sequence of *mTOR* showed a total of 649 probable G-quadruplex foldable regions, the investigation was limited to the most crucial regions near the transcription start site (-1000 to +500bp). G-quadruplexes are generally form with a minimum of three G-tetrads, though two G-tetrads folded G4s have recently been evidenced as well.[202] In accordance with these observations, the G4-forming sequences with both three-tetrad and two-tetrad probabilities were given equal consideration during the study. PQS searching studies have explored five potential DNA stretches (P1-P5) in the upstream and downstream region of transcriptional start site (-1500 to +800) of *mTOR* gene which have a high GC content and a potentially a crucial role in regulating the gene expression machinery (Figure 4.2). Each PQS contained flanking regions consisting of three nucleotides at both 3' and 5' ends for enhanced stability of proposed G-quadruplexes.[203] These PQSs were named P1-P5 for the current study. Each sequence presented high frequency of G-tract repeats accompanied with the potential loop regions which may play critical role in G-tetrad formation and aid the structure to adopt a stable G-quadruplex conformation. The length of the identified PQSs lie between 29-35 nucleotide long and contain 4-6 G-tracts.

The P1 PQS was found in the upstream promoter region and its G-tracts contain four arrays of two guanosines (G_2). P2-P5 PQSs were identified in the +1 to +400 region of *mTOR* gene. The P2 sequence is comprised of six arrays of G-tracts: four containing >3

guanosines (G_{3-5}) and the other two contain only two guanosines (G_2). The P3 and P5 contain six and four G_2 stretches, respectively and the P4 sequence has two arrays of G_2 and G_3 stretches. The linkers between G-stretches differ in both length and sequence. Interestingly, when the 5'-UTR region of *mTOR* mRNA was searched for potential POSs, we could only detect the presence of P2 and P3 PQS which may be the effect of post-transcriptional modifications.

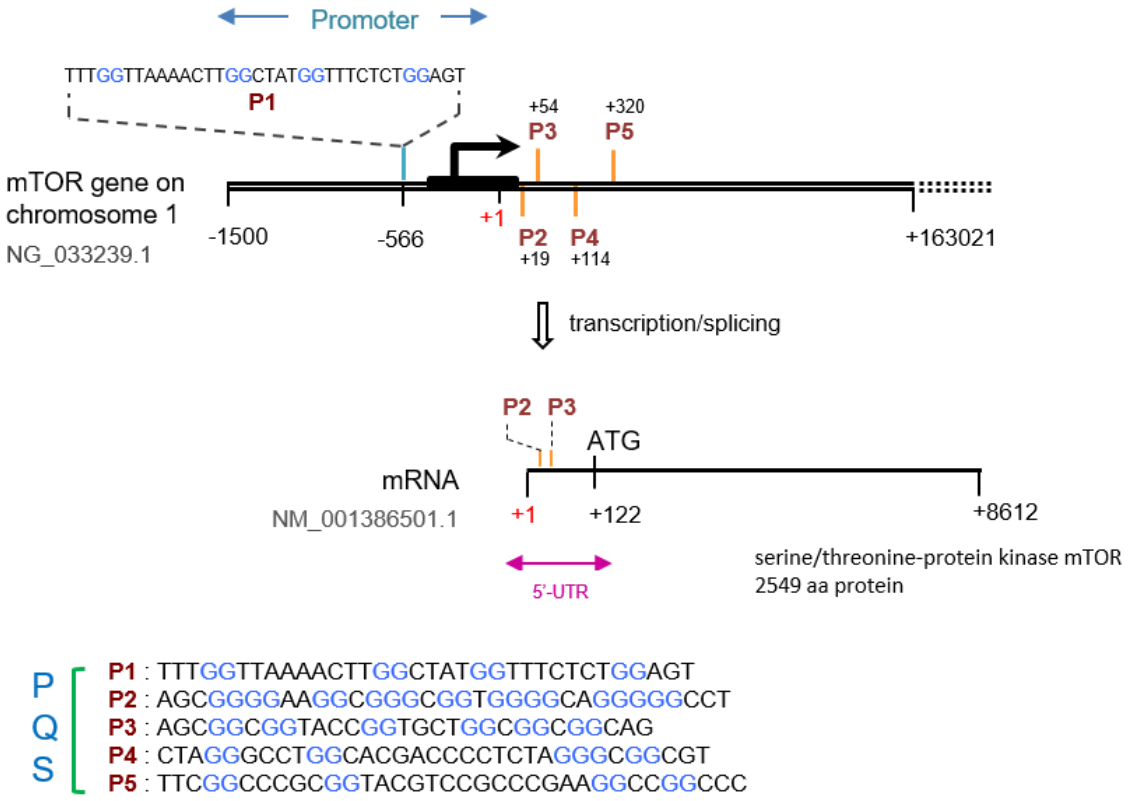


Figure 4.2. Schematic representation of the putative G-quadruplex sequences (PQS) in the DNA (P1-P5) and the 5' UTR of the mRNA (P2-P3) of *mTOR* gene.

Circular dichroism spectroscopy

Depending on the guanine glycosidic bond angle conformation of either *syn* or *anti*, the G-quadruplex structures can be categorized as parallel (a positive and negative peak at 260 and 240 nm respectively), anti-parallel (a positive and negative peak at 290 nm and 260 nm respectively) or hybrid (two positive peaks near 260 and 290 nm and a negative peak at 240 nm) conformations. To assess the folding patterns of these G-rich sequences P1-P5, Circular Dichroism (CD) spectroscopy, a widely used method for characterization of G-quadruplexes was employed. Circular dichroism signatures were obtained for each of the PQS, P1-P5 in a buffer solution containing no ion or K^+, Na^+ and Li^+ ions at physiological pH of 7.4 (Figure 4.3a-e). Among these PQSs, P2 exhibited an ion dependent secondary structure formation whereas P1, P3, P4 and P5 did not show appreciable differences in their CD profiles in presence of different ions. P2 containing six stretches G-tracts (G_2 and $G_{3\text{-}5}$) showed characteristic CD spectra of a G-quadruplex in the presence of K^+; a prominent maxima at 262 nm, a smaller maxima at 286 nm and a negative minima at 239 nm. This characteristic peak of P2 was highly diminished in presence of Na^+ and Li^+ as well as in the absence of ions suggesting that the G_4-structure is only stable in the presence of K^+ ion.

In case of P1, the observed peaks without any ion or in the presence of each monovalent cation showed features of random coil conformation, with a positive CD band near 275 nm and a negative band near 245 nm (Figure 4.3a). P3, P4, and P5 exhibited broad positive bands with comparatively lower intensity in the range of 260-280 nm in presence of different ions (Figure 4.3c-e), suggesting that these sequences could not adopt G4 secondary structure in the presence of a metal ion. P2 PQS emerged as the only potential G4 forming sequence capable of forming a stable G-quadruplex structure in the presence of K^+. The CD signature of P2 (positive maxima at 262 and 286 nm and a negative minima at 239 nm) indicated that the P2 sequence could potentially fold into a mixed parallel/anti-parallel G-quadruplex with characteristic peaks similar to the *BCl-2* promoter[204, 205] and *vegfr-2* promoter G-quadruplexes.[200] Interestingly, the CD spectra of P2 in presence of K^+ showed a much more distinct CD signature compared to the P2 spectrum in the presence of Na^+ or without any ion (Figure 4.3b). In addition, it was observed that P2 lost the G4 structural conformation in presence of Li^+ ion, an ion known to destabilize the G-

quadruplex structure.[206, 207] The preliminary results suggest that P2 DNA sequence may preferentially adopt a hybrid G_4 structure in presence K^+ over Na^+ and the structure is disrupted in presence of Li^+.

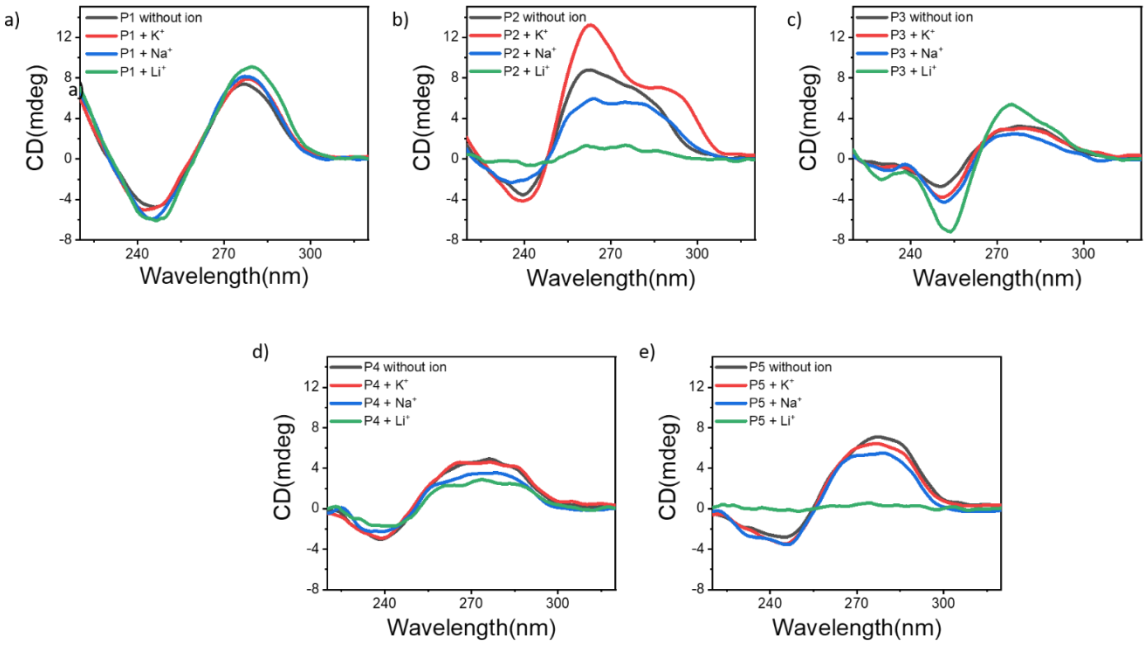


Figure 4.3. Circular dichroism spectra of putative G4 sequences (5 µM) without ions and in presence of different ions (K^+, Na^+ Li^+) where ,a) P1, b) P2, c) P3, d) P4,e) P5. All experiments were performed in respective buffers of 100 mM Tris with 100 mM HCl, 100 mM Tris with 100 mM KCl, 100 mM Tris with 100 mM NaCl and 100 mM Tris with 100 mM LiCl. The ph oaf all the buffers was maintained at 7.4.

Circular dichroism melting assay

Additionally, CD melting experiments were conducted to monitor the thermal denaturation patterns of P1-P5 sequences in a buffer containing 100 mM K^+ ion at a strand concentration of 5 µM (Figure 4.4a-e). G4s reportedly possess higher stability than duplex DNA/random coil DNA owing to the presence of Hoogsteen hydrogen bonds. This additional stability results in higher T_m value of the G4 compared to non-structured DNA. The PQSs were subjected to gradual heating from 20 °C to 90 °C and the CD spectrum was recorded for

every 10 °C rise in temperature. The CD melting experiments demonstrated that a significant T_m value was successfully acquired only for P2 which is at 66.6 °C while the melting profiles of the rest of the PQSs could not be acquired satisfactorily (Figure 4.4f). This result prompted us to investigate the G-quadruplex forming potential of P2 in more detail, especially in the mRNA counterpart (rP2) as it lies in the 5' UTR of the *mTOR* mRNA.

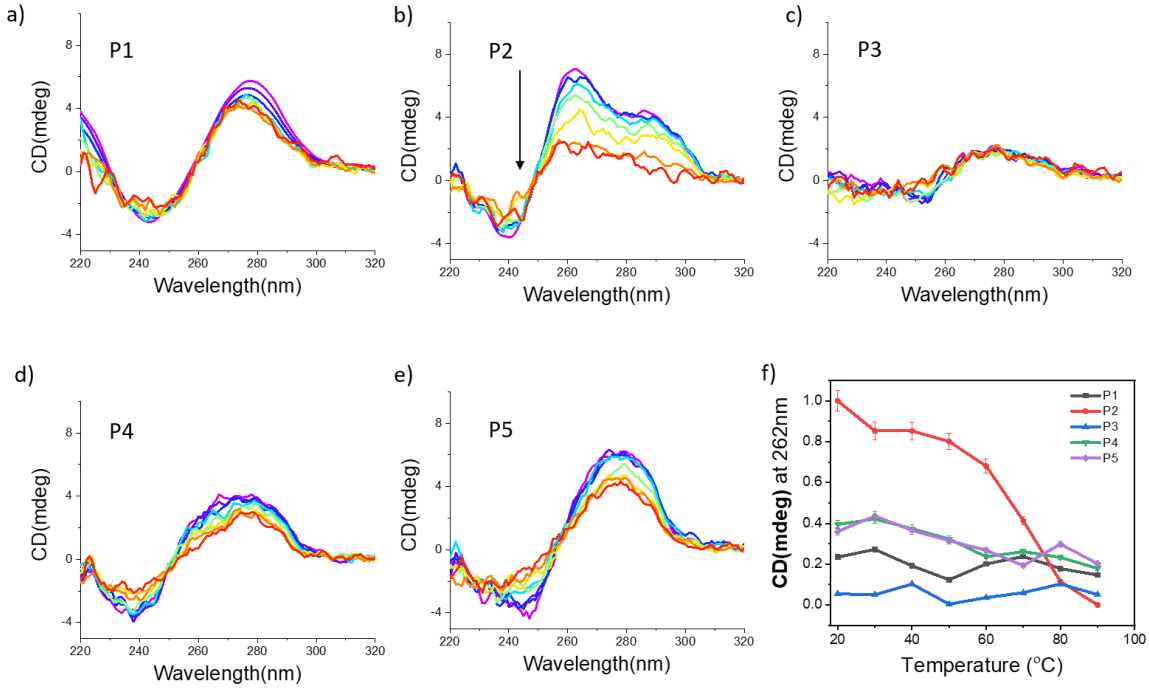


Figure 4.4. CD melting spectra of DNA sequences exhibiting the change in the CD spectra from 10-100 °C recorded at intervals of 10 °C, a) P1, b) P2, c) P3, d) P4,e) P5. All experiments were performed in 100 mM Tris with 100 mM KCl buffer (pH- 7.4).

As the P2 and P3 sequence is also found in the 5' UTR of *mTOR* mRNA, both of their RNA counterparts (rP2 and rP3) was also investigated in presence of different metal ions (K^+, Na^+ and Li^+) using CD spectroscopy(Figure 4.5). Interestingly, a similar formation of G-quadruplex was observed for rP2 with a prominent positive peak at 263 and a smaller merged peak at 290nm and a negative peak at 239nm. This rP2 G4 also

showed a maximum stability in K^+, closely followed by Na^+ and without ion. The rP2 structure showed least stability in presence of Li^+ ion which probably destabilizes the structure. The rP3 RNA sequence failed to show any ion-dependent changes in the CD spectra suggesting that it does not form any ion-dependent secondary conformation such as G-quadruplex. Interestingly, both the DNA and RNA sequence of P2 and rP2 respectively are capable of forming stable G4 structures which exhibit highest stability in presence of K^+.

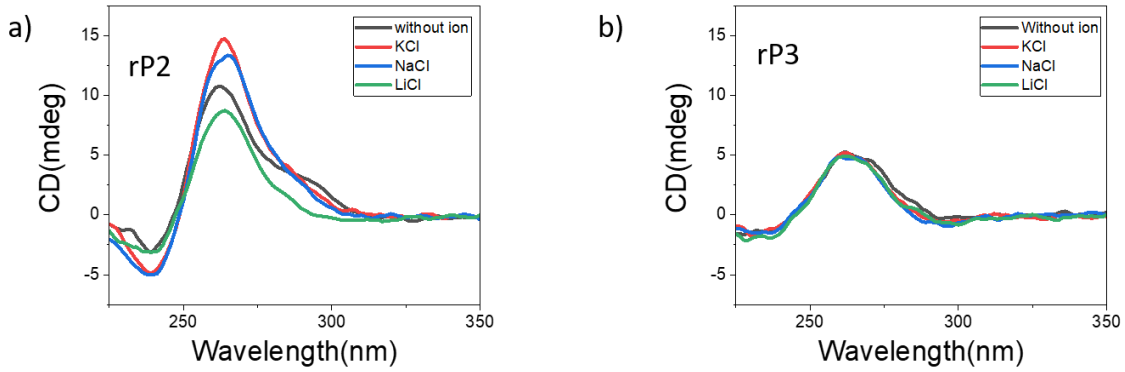


Figure 4.5. Circular dichroism spectra of putative RNA G4 sequences (5 µM) without ions and in presence of different ions (K^+, Na^+ Li^+) where , a) rP2, b) rP3. All experiments were performed in respective buffers of 100 mM Tris with 100 mM HCl, 100 mM Tris with 100 mM KCl, 100 mM Tris with 100 mM NaCl and 100 mM Tris with 100 mM LiCl. The ph oaf all the buffers was maintained at 7.4.

In accordance to the promising rsults obtained from the CD spectroscopy investigation sof P2, detailed CD studies for rP2 was also carried out. The CD spectra of rP2 at different strand concentrations (5 µM, 10 µM, 25 µM, 50 µM and 100 µM) was investigated to monitor the effect of strand concentrations on G4-stability (Figure 4.6a). A direct relation between strand concentration and the G4-characteristic peak intensity was observed; the G4-forming potential of rP2 was increased with increased strand concentration in the presence of K^+. Interestingly, the positive maxima at 262 nm became more prominent as

compared to the other positive maxima at 286 nm at 100 µM strand concentration. Thermal denaturation profile was also analyzed at different rP2 strand concentrations in order to validate the stability of the G4 structure formed by P2 (Figure 4.6c-d). The T_m was found to be 68±2 °C at strand concentrations of 5 µM, 10 µM and 100 µM respectively (Figure 4.6e). This suggests that at different strand concentrations, rP2 maintains its overall T_m value further indicating that P2 could acquire a higher-order G4 DNA structure that are stable in the presence of K^+ ion.

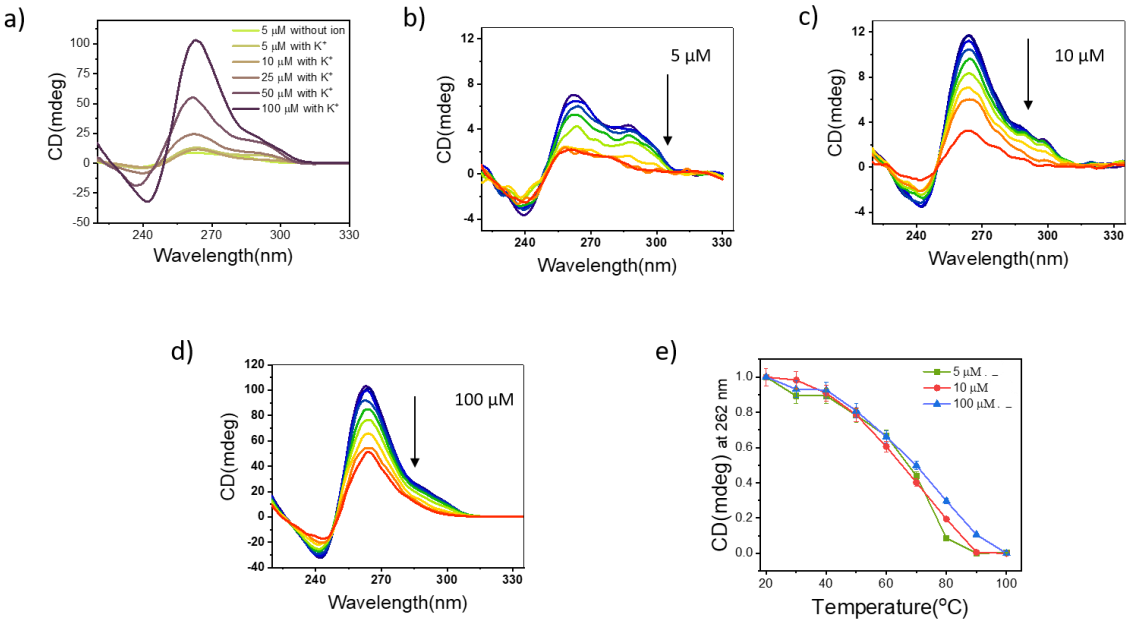


Figure 4.6. Circular dichroism spectra of putative RNA G4, rP2 at different conditions; a) rP2 at different strand concentrations in presence of K^+ , b-d) CD melting spectra of rP2 at 5 µM, 10 µM and 100 µM strand concentration. The spectra from 20-100 °C were recorded at intervals of 10 °C, e) Normalized melting graph of rP2 at different stand concentrations. All experiments were performed in 100 mM Tris with 100 mM KCl (pH-7.4).

Electrophoretic Mobility Shift assay

Further investigations to visualize the formation of secondary structures in the studied DNA (P1-P5) and RNA (rP2 and rP3) sequences were conducted. Electrophoretic mobility shift assay (EMSA) is a reliable method to study the movement of folded DNA and RNA structures compared to unfolded structures.EMSA of the structures showed that only P2 and rP2 are able to form secondary structures of nucleic acid especially in presence of K^+ (Figure 4.7). These observations are in agreement with the CD spectroscopy results suggesting that P2 and rP2 are putative G-quadruplex sequences which are capable of forming G-quadruplexes.

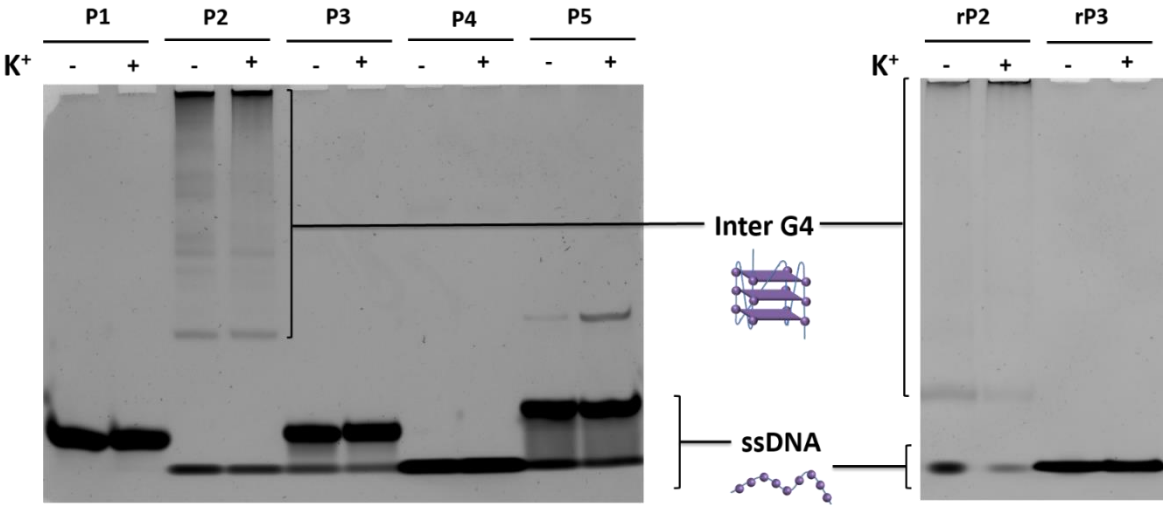


Figure 4.7. Electrophoretic mobility shift assay demonstrating the formation and movement of folded nucleic acid structures, P2 DNA (left panel) and rP2 RNA (right panel). The nucleic acidconcentration was 10 µM and the movement of the oligonucleotides were observed in 15% PAGE. The oligonucleotide samples were prepared in 100 mM Tris with 100 mM KCl (pH-7.4).

NMR Spectroscopy

For further insights into the possibility of P2 and rP2 folding into a G4, 1D H^1 NMR was performed with the P2 and rP2 sequences in the presence of the K^+ containing buffer (pH 7.4). Both G4 sequences demonstrated characteristic imino protons from 9.5 to 12.0 ppm which correspond to G-tetrad formation. The NMR spectra of P2 PQS showed characteristic features of Hoogsteen bond formation further suggesting that P2 may potentially form a potential G-quadruplex structure (Figure 4.8). Characteristic peaks indicative of Hoogsteen bond formation was observed for P2 and rP2 PQS in the region of 10.2 to 11.7. These evidences further suggest that P2 and rP2 may be induced to form the G-tetrad stacked G-quadruplex structure in the presence of K^+ ion.

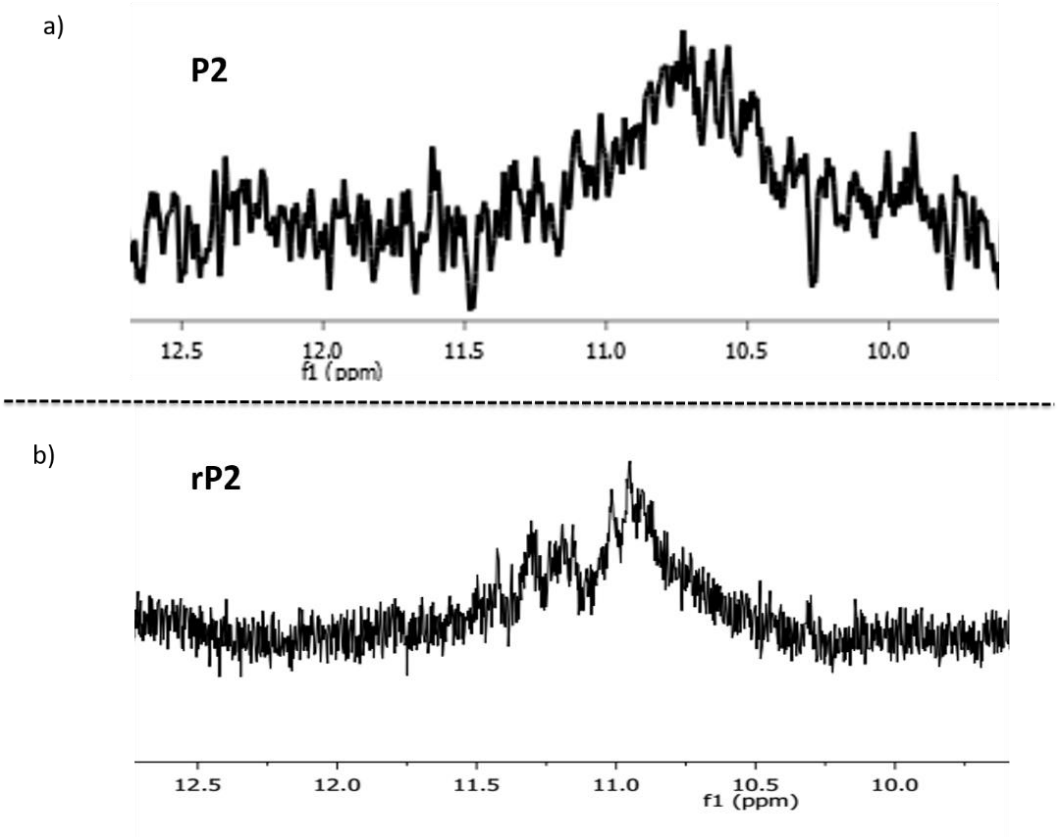


Figure 4.8. Nuclear magnetic resonance spectra of a) P2 (500 µM) and c) rP2 (500µM) in 100 mM Tris with 100 mM KCl (pH-7.4).

The implications of an RNA G4 in the 5'UTR region of *mTOR* were studied in greater detail. The rP2 oligonucleotide with a sequence of r($G_5N_2G_2N_1G_3N_1G_2N_1G_4N_2G_5$) can fold into a stable G-quadruplex structure. However, among these six G-tracts of rP2, only four could be involved in G-quadruplex formation. In the next set of experiments, the aim was to identify which G-tracts of rP2 are responsible for the formation of G-quadruplex structure.

In the biological context, the functional implications of a G-quadruplex in the non-template DNA strand is lesser than in the 5'UTR. This is due to the direct relation between 5'UTR G4 and translation of the protein. Thus further investigations in this study was focused majorly on the rP2.

4.2. Investigation to decipher the contribution of individual G-tracts in the stability of the rP2 G-quadruplex in vitro and in gene expression

Mutation studies

To evaluate the role of each of the G-tracts comprising the rP2 sequence as a possible contributor of G-quadruplex formation, different mutations were introduced into each G-tract of the rP2 PQS (Figure 4.9). Initially, the CD signatures of the partial rP2 mutants (rP2.1a to rP2.6a) were analyzed (Figure 4.10a-h), where G→A mutations were introduced separately in each G-tract.

a)

Name	Sequence	Location of mutation in G-tract	Nature of mutation
rP2 WT	AGCGGGGAAGGCGGGCGGUGGGGCAGGGGGCC	-	None
rP2.1a	AGCGAAGAAGGCGGGCGGUGGGGCAGGGGGCC	1	Partial
rP2.2a	AGCGGGGAAGACGGGCGGUGGGGCAGGGGGCC	2	Partial
rP2.3a	AGCGGGGAAGGCGAGCGGUGGGGCAGGGGGCC	3	Partial
rP2.4a	AGCGGGGAAGGCGGGCGAUGGGGCAGGGGGCC	4	Partial
rP2.5a	AGCGGGGAAGGCGGGCGGUGAAGCAGGGGGCC	5	Partial
rP2.6a	AGCGGGGAAGGCGGGCGGUGGGGCAGAAAGCC	6	Partial
rP2.7a	AGCGAAGAAGACGAGCGAUGAAGCAGAAAGCC	1,2,3,4,5,6	Partial

b)

Name	Sequence	Location of mutation in G-tract	Nature of mutation
rP2 WT	AGCGGGGAAGGCGGGCGGUGGGGCAGGGGGCC	-	None
rP2.1b	AGCAAAAAAGGCGGGCGGUGGGGCAGGGGGCC	1	Full
rP2.2b	AGCGGGGAAAACGGGCGGUGGGGCAGGGGGCC	2	Full
rP2.3b	AGCGGGGAAGGCAAACGGUGGGGCAGGGGGCC	3	Full
rP2.4b	AGCGGGGAAGGCGGGCAAUGGGGCAGGGGGCC	4	Full
rP2.5b	AGCGGGGAAGGCGGGCGGUAAAACAGGGGGCC	5	Full
rP2.6b	AGCGGGGAAGGCGGGCGGUGGGGCAAAAAACC	6	Full
rP2.7b	AGCAAAAAAAACAAACAAUAAAACAAAAAACC	1,2,3,4,5,6	Full

Figure 4.9. List of mutations introduced in the rP2 Wild Type (WT) sequence to study changes in folding dynamics.

In rP2.1a, rP2.3a, rP2.5a and rP2.6a sequences, the central positions of the 1st, 3rd, 5th and 6th G_{3-5} tracts were mutated and in sequences rP2.2a and rP2.4a, the second position of the 2nd and 4th G_2 tracts were mutated into A, respectively. The mutations are designed in such a way that the respective mutated G-rich tract in the oligonucleotide sequence could not be involved in G-quadruplex formation. The CD spectra of the mutated sequences revealed that the peak intensity at 262 nm was significantly decreased for rP2.5a (mutation in tract 5: GGGG→GAAG) and rP2.6a (mutation in tract 6: GGGGG→GAAAG), indicating their sure involvement in G-quadruplex formation (Figure 4.10). The sequence rP2.2a (mutation in tract 2: GG→GA) showed comparatively less decrease in the peak intensity (at 262 nm). For rP2.1a (mutation in tract 1: GGGG→GAAG) and rP2.4a (mutation in tract 3: GGG→GAG), no significant change in the peak signal was observed; indicating that the formed G-quadruplex may not involve tract 1 and 4. But interestingly with rP2.3a (mutation in tract 4: GG→GA), the signal at 262 nm was decreased but a new peak was generated at 295 nm, illustrating the formation of another G-quadruplex conformation having major positive peak at 295 nm and a minor positive peak at 262 nm along with the negative peak at 240 nm. This G-quadruplex of another conformation may be formed by the other non-mutated G-tracts which become available for G4 formation upon mutation in 4th G-tract.

The rP2.7a sequence containing partial mutations in all six G-tracts led to no specified CD signal indicating the absence of any G-quadruplex conformation due to lack of G-rich tract. Overall, these CD profiling of partial mutated sequences (rP2.1a-rP2.6a) indicate that tracts 5 and 6 may have critical role in G-quadruplex formation but the distortion and decrease in the CD signals could not confirm the involvement of tract 1-4 in G4 formation. It seems that when each of the G-tracts at position 1-4 has been mutated, the other G-tracts compensate for the missing G-tract and may thus be involved in G-quadruplex formation.

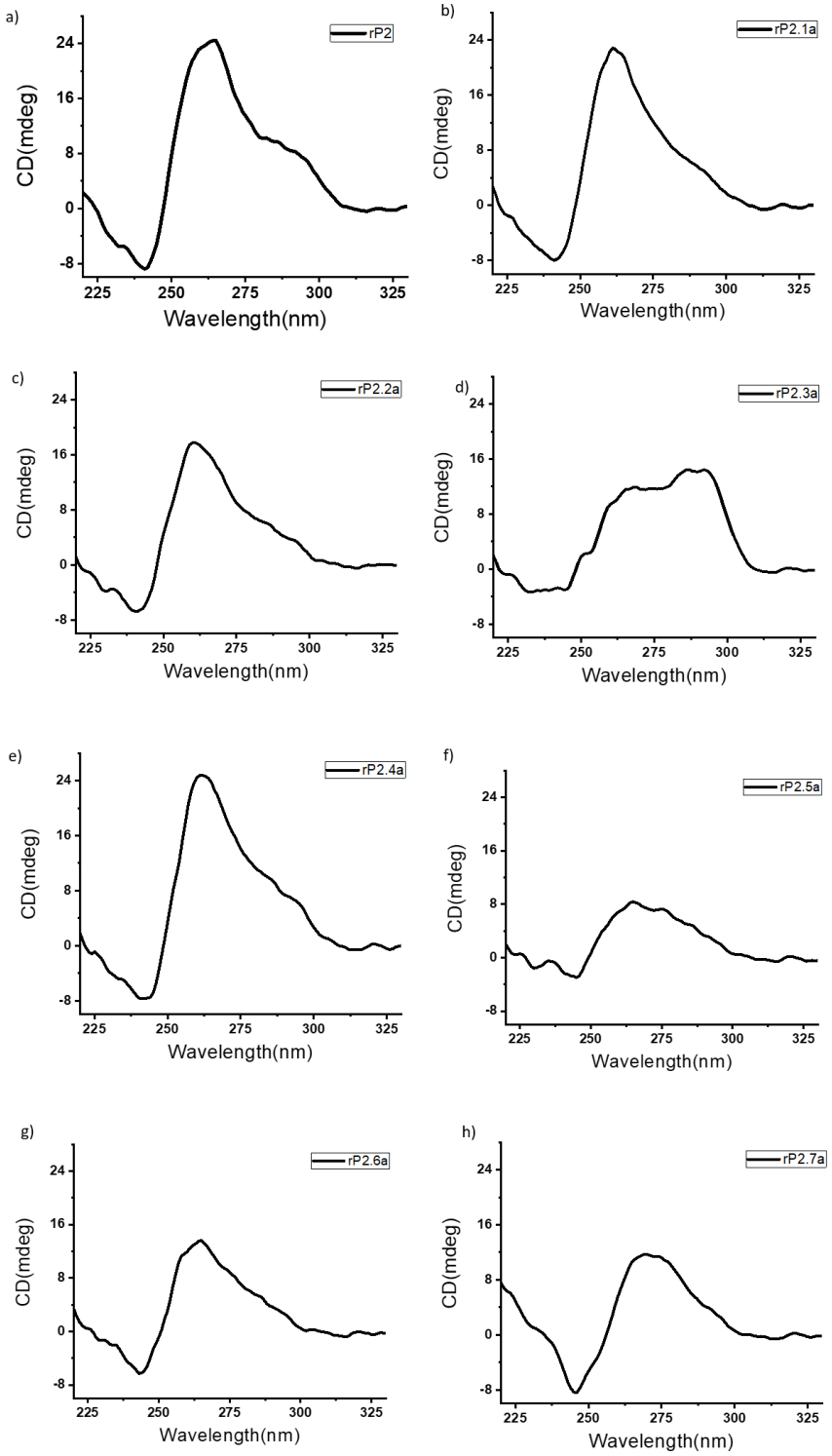


Figure 4.10. a-h) CD spectra of rP2 partial mutants (rP2.1a-rP2.7a) containing G to A mutations in a part of the G-tracts. All experiments were carried out in 100 mM Tris with 100 mM KCl buffer (pH-7.4).

In order to fully confirm the involvement of the G-tracts in G4 formation, we next evaluated the effect of introducing complete or full tract mutations in each G-tract of the rP2 PQS (rP2.1b to rP2.6b) (Figure 4.11a-h). The introduction of full tract mutations in G-tract 5 (sequence rP2.5b) and 6 (sequence rP2.6b) resulted in significantly decreased peak intensities at 262 nm while mutant P2.1b exhibited a complete disruption of the signature CD peak at 262 nm (Figure 4.11f). This is quite interesting as we have previously observed that the rP2.1a sequence having partial mutations in tract 1 did not lead to any significant change in the G-quadruplex peak intensity. These observations thus validated the involvement of tract 1, 5 and 6 in P2 G4 formation as introduction of full G→A mutations in these G-tracts critically hamper the G4 formation. However, the rP2.2b and rP2.4b sequences (having complete G→A mutations in tract 2 and 3, respectively) did not lead to any significant disruption or decrease in the G-quadruplex peak intensity, suggesting that tract 2 and 4 may not be involved in G4 formation. In case of rP2.3b, the peak intensity at 262 nm was decreased but to a lesser extent in comparison to rP2.1b, P2.5b and rP2.6b (Figure 4.11f). Therefore, it can be inferred but tract 3 may contribute to G-tetrad formation to generate P2 PQS G-quadruplex structure.

The concomitant complete mutation of all G-tracts (rP2.7b), as expected, led to an absolute absence of a specified folded structure like our previous observation in its partially mutated counterpart (rP2.7a). Together these results suggest that tracts 3, 5 and 6 may play a critical role in the P2 G-quadruplex formation, loss of which disrupts the formation of a G-quadruplex structure, while tract 1 is also highly critical for the G-quadruplex formation. We hypothesize that the loss of the internal-guanosines in tract 1 (as in rP2.1a) may lead to an alternate participatory role of the G-nucleotides from the adjacent tract 2, thereby allowing the G-quadruplex to form with elements of both tract 1 and tract 2 as a compensatory mechanism to still fold into the G-quadruplex structure with a decreased peak intensity. In addition to this, we further observed that the mutation of tract 1 and 6 leads to significant disruption of the CD signature peaks at 262 nm, whereas in case of full tract mutations of tract 3 and tract 5, a lesser degree of CD peak disruption is observed. This may be due to a compensatory action of the adjacent G-tracts 4 and 5 which allows a lesser perturbation if one of the tracts are mutated. Based on these observations, we can conclude that tracts 1, 3, 5 and 6 are primarily involved in G4 formation of the rP2 PQS.

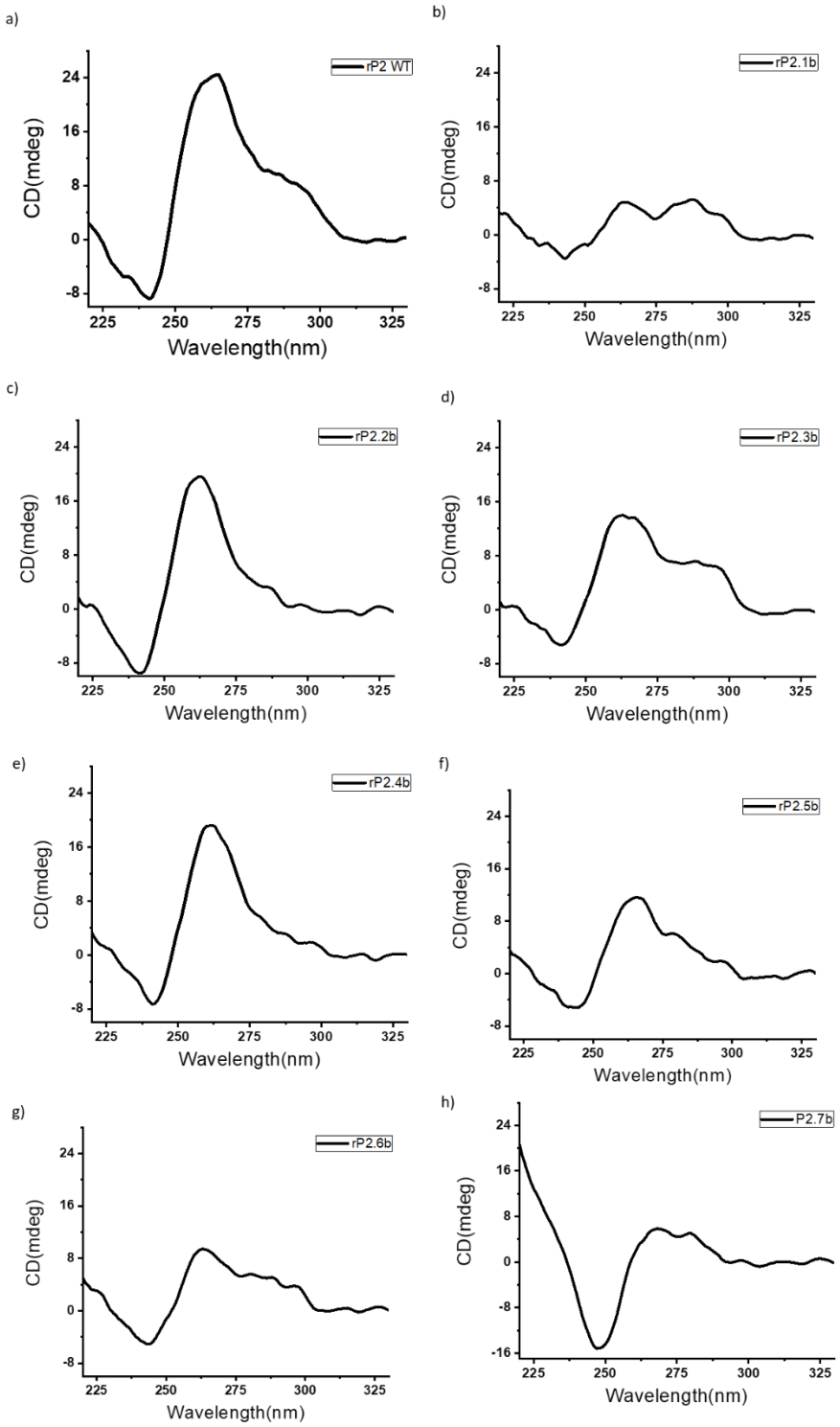


Figure 4.11. a-h) CD spectra of rP2 full mutants (rP2.1b-rP2.7b) containing G to A mutations in a part of the G-tracts. All experiments were carried out in 100 mM Tris with 100 mM KCl buffer (pH-7.4).

DMS Footprinting

Furthermore, dimethyl sulphate (DMS) footprinting assay was performed to ascertain which G-tracts of rP2 are involved in the G-quadruplex formation (Figure 4.12). The N7atom of a guanine residue in a G-quadruplex is protected from methylation by DMS and subsequent chemical cleavage by piperidine. These Gs involved in G4 formation can be identified by the pattern of cleavage bands using the DMS footprinting method. 5'-FAM labeled rP2 was methylated with DMS followed by chemical cleavage by Piperidine. The resulting samples with and without K^+ displayed distinct cleavage bands suggesting that the G4 contributing G-tracts are 1, 3, 5 and 6. The cleavage bands for tracts 2 and 4 showed similarly high intensities in both samples with and without K^+, suggesting that they are not involved in the formation of the G-quadruplex. These results are in agreement with the previous observations where the CD spectroscopy consisting of multiple mutations were used to investigate the tracts involved in G4. Interestingly, the cleavage bands for the Gs in G-tracts 4 and 5 were very less in intensity even in absence of K^+ which were further discriminable on addition of K^+ to show that only three Gs in each G-tract are involved in the formation of rP2 G-quadruplex. The DMS footprinting results also suggest that even in the absence of K^+, the rP2 structure is partially stable as observed in the CD spectroscopy experiments.

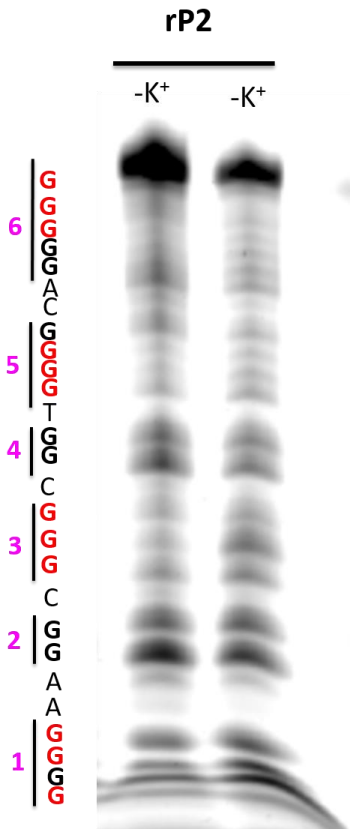


Figure 4.12. Dimethyl sulphate (DMS) footprinting assay of rP2 (100 µM) in absence and presence of K^+. The guanines (G) are shown in red indicate the probability of being involved in the G-quadruplex formation.

4.3. The binding of TMPyP4 with P2 G4 interferes with replication of *mTOR* DNA

UV/Vis Spectroscropy

Small molecules have emerged as G-quadruplex stabilizing ligands to modulate expression of critical genes. To determine the capability of G-quadruplex binding agent to stabilize our sequence of interest, we carried out a number of interaction profile studies with the cationic porphyrin 5, 10, 15, 20-tetra-(N-methyl-4- pyridyl) porphine, TMPyP4 in association with our target sequences, P2. Firstly, the UV-Vis absorption spectroscopy-

based titration of TMPyP4 with incremental additions of P2 was carried out (Figure 4.13). The absorbance peak of TMPyP4 was observed at 421 nm and was followed by incremental additions of pre-folded P2 sequence. A hypochromic shift accompanied with bathochromic red shift of approx. 15 nm was observed upon titration of TMPyP4 with P2. The dissociation constant was also calculated in the micromolar range where the K_d value of 2.1 µM was observed.

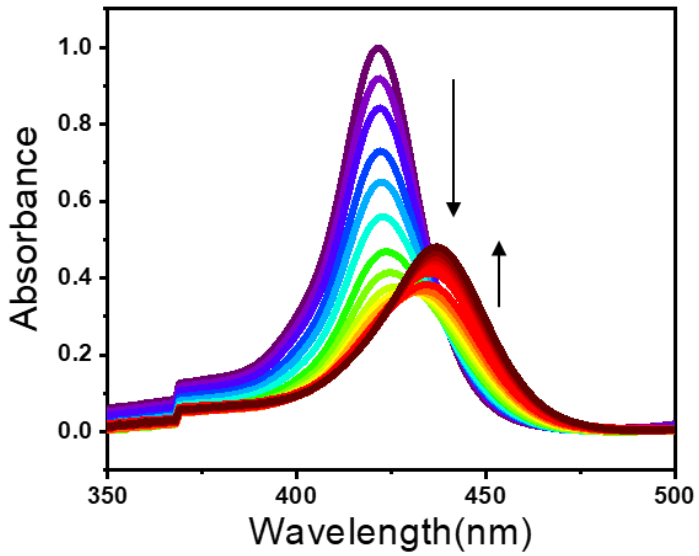


Figure 4.13. UV/Vis spectroscopy study demonstrating the titration of TMPyP4 (0.5µM) with P2 DNA G4 sequence. The experiments were carried out in 100 mM Tris with 100 mM KCl buffer (pH-7.4).

Fluorescence spectroscopy

Furthermore, fluorescence spectroscopic investigation was undertaken to study the interaction of TMPyP4 and P2 sequence (Figure 4.14). The fluorescence measurement of TMPyP4 was initially observed and found to form a broad peak with λ_{max} at 681nm. Upon incremental additions of pre-folded P2, a gradual quenching of TMPyP4 emission spectra was observed. The quenching led to the breakdown of the initial broad peak into two sharp peaks at 660nm and 721nm. This result is consistent with previous observations of

TMPyP4 interaction with folded G-quadruplexes suggesting that P2 PQS is able to form a folded secondary conformation of G-quadruplex. Thus, both UV-Vis and Fluorescence spectroscopic studies suggest that P2 can form a G-quadruplex structure capable of interacting with the porphyrin structure of TMPyP4.

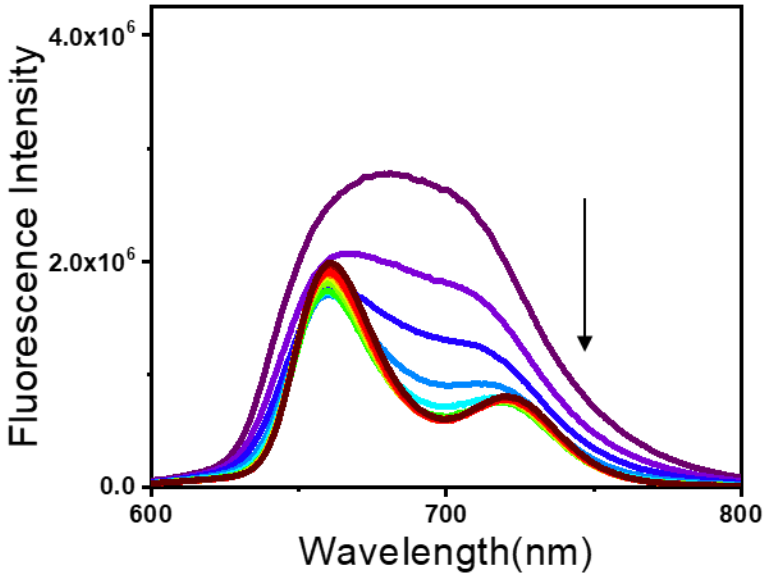


Figure 4.14. Fluorescence spectroscopy study demonstrating the titration of TMPyP4 (0.5μM) with P2 DNA G4 sequence. The experiments were carried out in 100 mM Tris with 100 mM KCl buffer (pH-7.4).

FRET melting assay

Förster resonance energy transfer (FRET) melting assay was performed with dual fluorophore tagged (5' FAM and 3' TAMRA) P2 sequence in presence of incremental concentrations of TMPyP4 (0.5, 1, 3, 5, 7, 10 μM) (Figure 4.15). It was observed that TMPyP4 increased the ΔT_m of our target sequence by 19.1 °C at ligand concentration of 0.5 μM. Higher concentrations of TMPyP4 resulted in further stabilization of P2 PQS with observed ΔT_m of 21.9 °C, 22.8 °C, 23.1 °C, 25.8 °C, and 27.1 °C at 1 μM, 3 μM, 5 μM, 7 μM, 10 μM respectively.

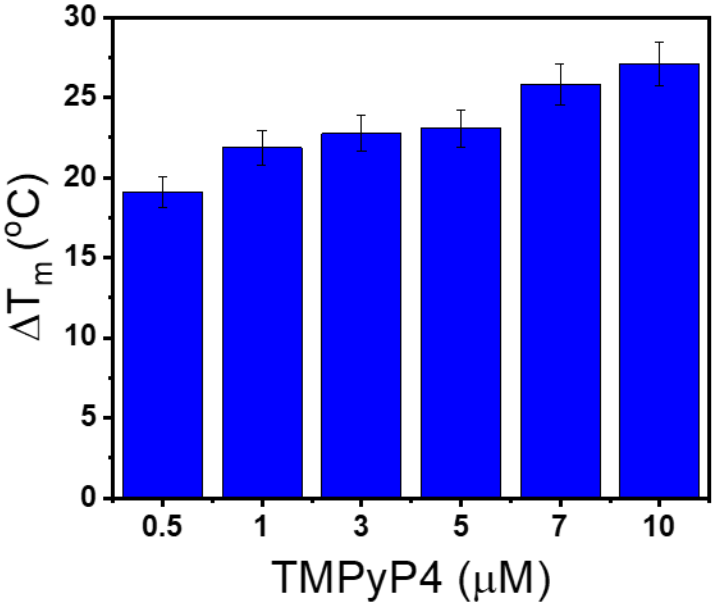


Figure 4.15. FRET melting assay of dual fluorophore tagged (5'FAM and 3' TAMRA) *mTOR* P2 sequence (0.2 μM). with increasing concentrations of TMPyP4 (0.5, 1,3,5,7 and 10 μM). All experiments were carried out in 60 mM Potassium Cacodylate buffer (pH-7.4).

Isothermal titration calorimetry

Isothermal titration calorimetry (ITC) was performed with P2 sequence in presence of TMPyP4 to ascertain the thermodynamic stability of the DNA ligand complex (Figure 4.16). The pre-annealed P2 PQS was titrated with incremental concentration of TMPyP4 through 30 regular injections in regular intervals of time and the isotherm was obtained for the DNA-ligand complex. The stoichiometry of binding was found to be 1.0 and a K_d of 2.3 μM was observed for the DNA-porphyrin interaction. The K_d found is in agreement with the UV/Vis calculation.

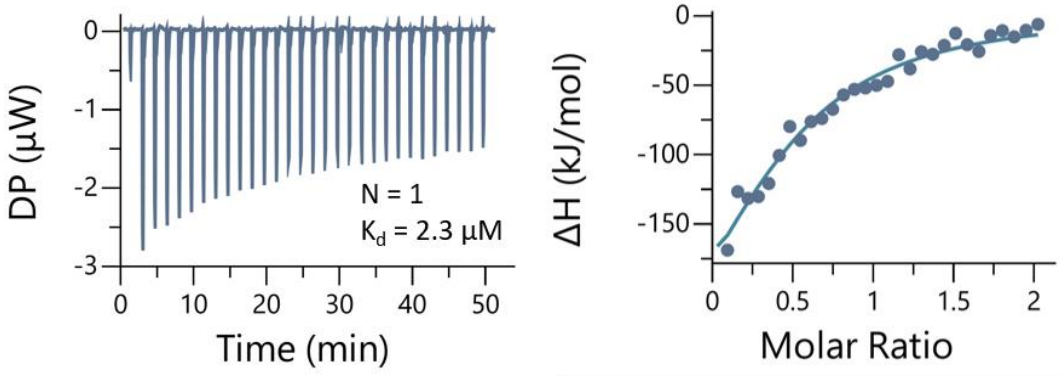


Figure 4.16. ITC heat curve and isotherm of the titration of P2 DNA G4 sequence (5 µM) with increasing concentrations of TMPyP4. The experiment was carried out in 100 mM Tris with 100 mM KCl buffer (pH-7.4).

Polymerase stop assay

The potential regulatory role of the P2 G-quadruplex in the gene expression of *mTOR* is a valuable aspect to explore. Initially, a Taq-DNA polymerase stop assay was performed to observe whether the DNA sequence of interest could fold into a G4 structure in the presence of a G4-binder such as TMPyP4 and act as a roadblock to Taq Polymerase, stalling replication and leading to a decreased amplification efficiency (Figure 4.17a). With increasing concentrations of TMPyP4, we observed pausing of DNA Pol extension and decreased amounts of full-length products. This indicated that stabilization of the P2 G4 structure by TMPyP4 could inhibit DNA elongation by Taq-DNA polymerase. In contrast, Taq-DNA polymerase stop assay performed with increasing concentration of K^+ (Figure 4.17b) did not exhibit a dose dependent stalling in the elongation process by Taq-DNA polymerase, suggesting that the K^+ presence is the only determinant in P2 G4 formation and increasing concentrations of K^+ has no significant effect on the G4 stabilization.

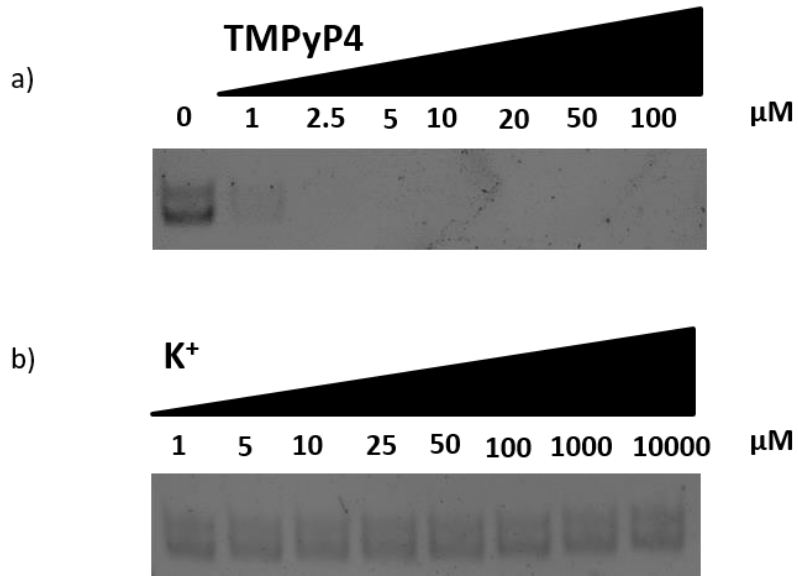


Figure 4.17. Polymerase stop assay of P2 sequence in presence of increasing concentrations of a) TMPyP4 and b) K^+.

4.5. The binding of TmPyP4 with rP2 G4 interferes with translational machinery of *mTOR* gene

The P2 sequence was observed to be stabilized by the G4 binder TMPyP4 and hinder the activity of polymerase to arrest the replication process. In this line of investigation, rP2 was also studied to moniter the any interaction of TMPyP4 with rP2 RNA G4. Initially, the UV-Vis absorption spectroscopy-based titration of TMPyP4 with incremental additions of rP2 was carried out using UV/Vis spectroscopy and Fluorescence spectroscopy (Figure 4.18a-b). Similar results were observed as with P2 and TMPyP4 interaction. In case of UV/Vis spectroscopy, the characteristic decrease in absorbance intensity was followed by hypochromic shift accompanied with bathochromic red shift. The shift was found to be 22nm. In addition, fluorescence-based titration of TMPyP4 with rP2 showed decrease in fluorescence and the breaking of the broad emission peak into two separate peaks, characteristic of TMPyP4 binding with a G4 sequence. Both these

spectroscopic studies yielded positive results for the interaction of rP2 with TMPyP4 suggesting that the G4 binding ligand interacts and binds with the rP2 sequence with high affinity.

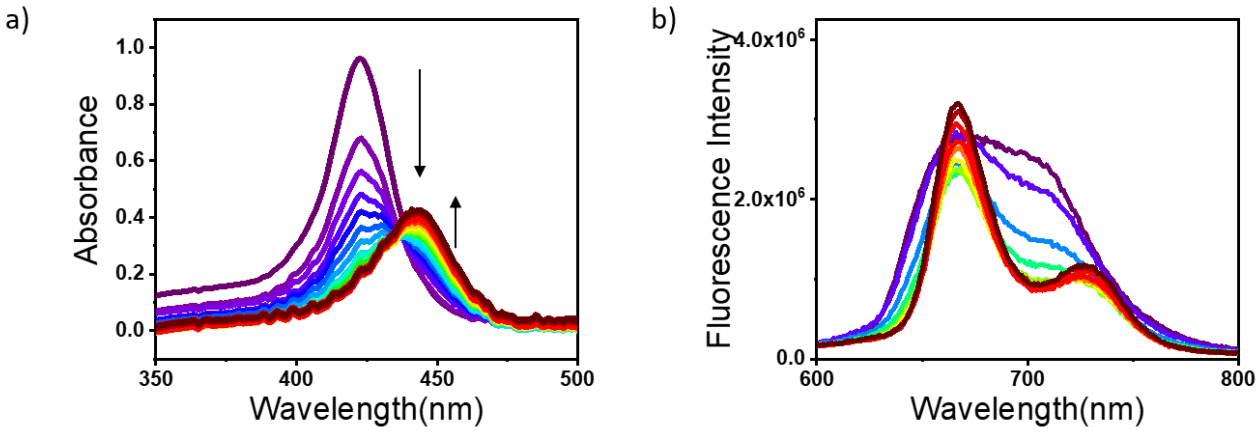


Figure 4.18. Spectroscopy studies demonstrating the titration of TMPyP4 (0.5µM) with rP2 DNA G4 sequence using the techniques of a) UV/Vis spectroscopy and b) Fluorescence spectroscopy. All experiments were carried out in 100 mM Tris with 100 mM KCl buffer (pH-7.4).

Isothermal Titration Calorimetry

Further investigation to study the binding of TMPyP4 with rP2 was carried out. The rP2 RNA G-quadruplex was titrated with TMPyP4 over 30 injections of equal volume and the change in the thermal profile was monitored (Figure 4.19). The K_d value for the interaction was found to be 1.1 µM with a stoichiometry of 1. This suggests that TMPyP4 bind rP2 with high affinity and when compared to the interaction of TMPyP4 with P2, it was found that TMPyP4 stabilizes rP2 with a higher affinity.

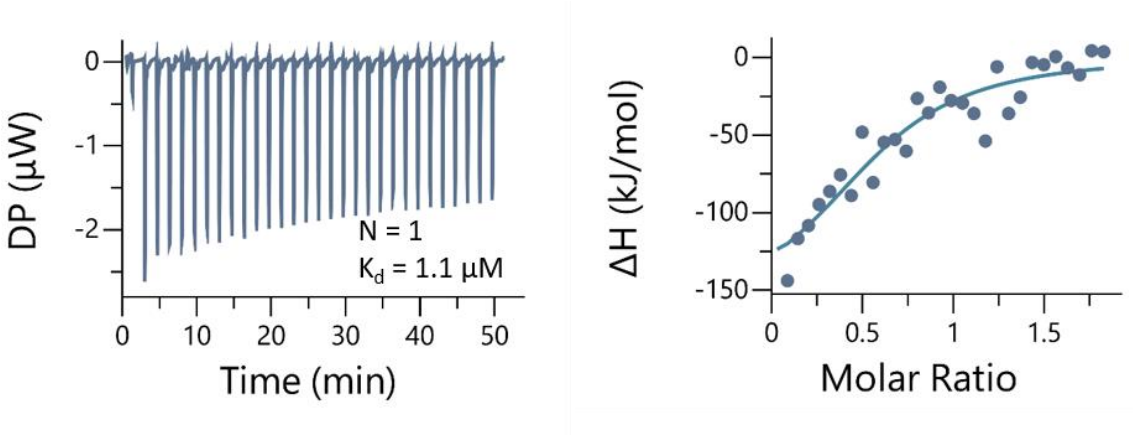


Figure 4.19. ITC heat curve and isotherm of the titration of rP2 RNA G4 sequence (5 µM) with increasing concentrations of TMPyP4. The experiment was carried out in 100 mM Tris with 100 mM KCl buffer (pH-7.4).

Reporter Luciferase Assay

The regulatory influence of rP2 G4 formation on the efficiency of translation was evaluated using luciferase reporter assay. A luciferase expressing vector (pGL4.73) containing cloned WT UTR sequence of 125 bp (containing rP2) upstream of the renilla luciferase gene was used to measure the level of luciferase expression (Figure 4.20). A similar cloned Mutant UTR (containing Mut rP2) sequence containing mutations (G➔A) incapable of folding into a G4 was used to study the role of G4 formation in the expression of the gene. The respective plasmids were transfected into HeLa cells and subjected to treatment with TMPyP4 to show the effect of ligand interaction where necessary.

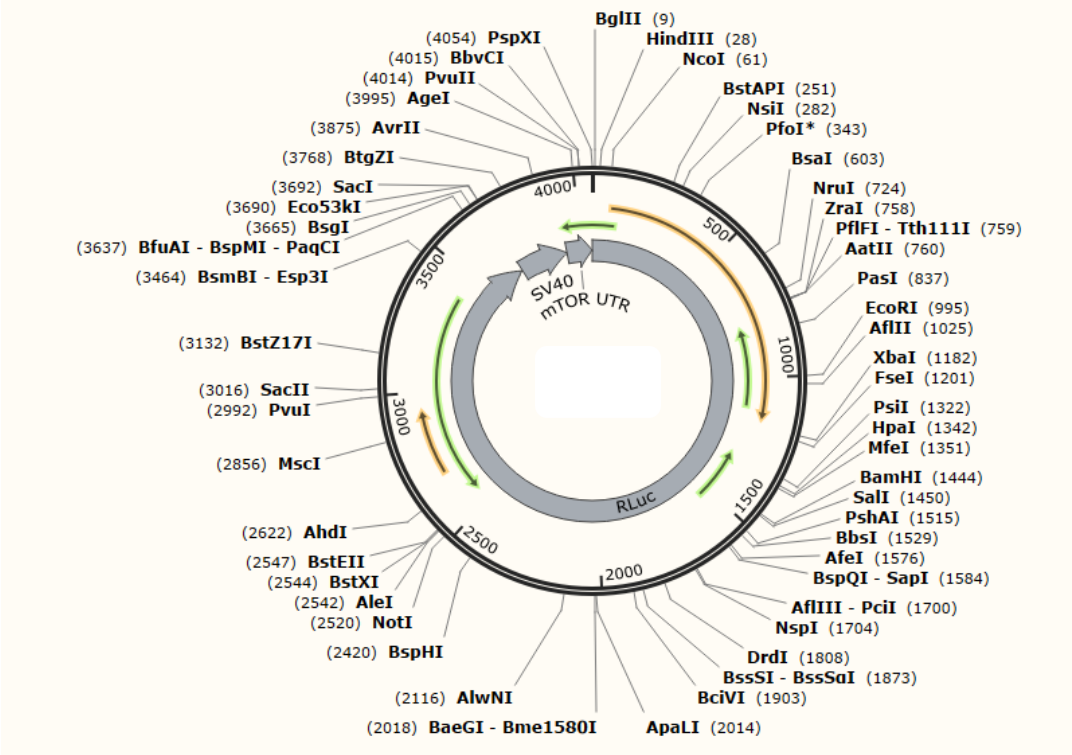


Figure 4.20. Schematic map of the renilla luciferase (rLuc) expressing plasmid contruct (pGL4.72) containing either wild type sequence (P2/rP2) or the mutant sequence (G to A mutation) in the UTR region of the RLuc gene. SV40 was used as the promoter.

The isolated cells lysates were measured for luciferase expression and normalized against their estimated protein respectively. The luciferase expression of WT rP2 containing plasmid was taken as 100% and the rest of the samples were evaluated accordingly (Figure 4.21). The renilla luciferase expression of RLTK plasmid was taken as the positive control whereas untransfected cells without the luciferase expressing plasmid were taken as the negative control for the experiment. The luciferase expression of Mut rP2 containing plasmid showed ~42% increase in translation efficiency relative to the WT system suggesting that the G4 formation in the 5'UTR region suppresses translation. In addition, when the rP2 and Mut rP2 systems were treated with TMPyP4 (4 uM) and BRACO-19 (2 uM), the translation efficiency of rP2 system decreased by ~30% whereas the translation efficiency of the Mut rP2 system continued to express ~40% increased efficiency. These results indicate that the native G4 formation in the 5' UTR region leads

to inhibition of the translation process which is further stabilized in presence of G4 binder TMPyP4. The regulation of translational of genes mediated *via* the 5' UTR is generally intricate and rP2 G4 may be one of the elements that can be further explored to understand this complex mechanism.

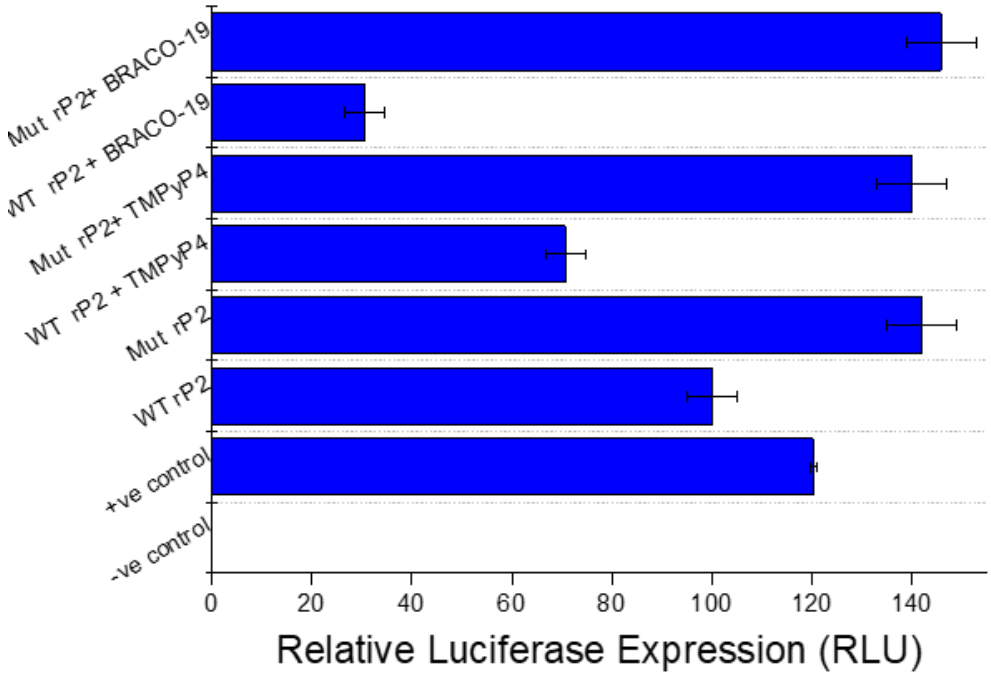


Figure 4.21. The relative luciferase expression of untreated and ligand treated (TMPyP4 and BRACO-19) luciferase expressing plasmid containing the Wild Type (WT) and Mutant (Mut) rP2 inserted in the 5'UTR region.

Western Blotting assay

Western blotting assay was performed to confirm the regulation of *mTOR* translation *via* the rP2 G4 stabilization with G4 binders. Relative protein expression of untreated and **TMPyP4** treated HeLa cells was assessed for *mTOR* gene and its downstream effectors and related proteins responsible for various biological processes including autophagy, cell survival and apoptosis. In addition to **TMPyP4**, another well

studied G4 binder, BRACO-19 was also used as an additional ligand in the study. It was observed that *mTOR* showed a dose-dependent downregulation of 20 and 32% on treatment with **TMPYP4**. A much greater dose dependent response was observed for **BRACO-19** treated HeLa cells which showed a downregulation of 30 and 73%. This suggests that the *mTOR* inhibition is due to the stabilization of the 5'UTR rP2 G4 which leads to downregulation of the translation of *mTOR* and reduced protein expression. This observation is also in agreement with the results of the reporter luciferase assay where the G stabilizers **TMPYP4** and **BRACO-19** induced a similar inhibition in gene expression. Owing to the crucial role of *mTOR* in the *mTOR* signaling pathway, the inhibition of *mTOR* may have a complex and multifaceted impact on various cellular factors. These factors are involved processes like apoptosis, autophagy, and cell survival, and their regulation is intricately connected to *mTOR* signaling. One of the processes directly inhibited by *mTOR*. Inhibition of *mTOR* can have significant effects on various factors involved in autophagy. Therefore, two such factors *ATG7* (Autophagy-related gene 7) and *ATG12* (Autophagy-related gene 12) were investigate to study the impact of rP2 stabilization on their activities.

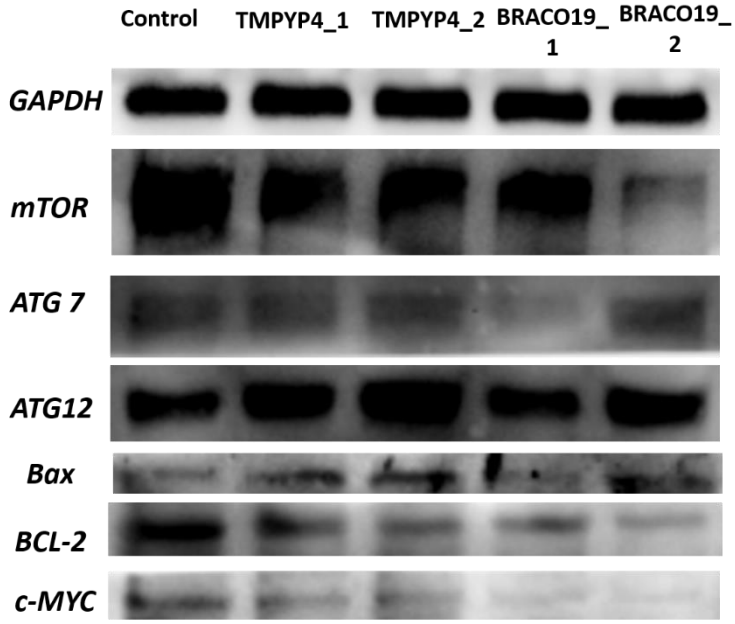


Figure 4.22. Protein expression of *GAPDH, mTOR, ATG7, ATG2, Bax, BCL-2* and *c-MYC* in HeLa cells treated with different doses of TMPyP4 (2 and 4 μM) and BRACO-19(1 and 2 μM).

ATG7 is essential for autophagosome formation as *ATG7* acts as an initiator of autophagosome biogenesis. The observed protein levels of *ATG 7* showed an upregulation of 8 and 15% with **TMPYP4** and downregulation of 38% and upregulation of 40% with **BRACO-19**. *ATG 12*, another important protein crucial for the elongation and closure of the autophagosome membrane also showed an upregulation of on treatment with **TMPyP4** (40 and 85%) and **BRACO-19** (20 and 46%). This suggests that the inhibition of *mTOR* reduces its suppressive effects on autophagy, promoting autophagosome formation and resulting in the autophagy machinery to become more active. *ATG 12*:

Increased autophagy can promote apoptosis by activating pro-apoptotic factors or through interactions with apoptotic signaling pathways. *Ba*x (Bcl-2-associated X protein) and *BCl-2* (B-cell lymphoma 2) are key regulators of apoptosis, or programmed cell death. Bax promotes apoptosis, while *BCl-2* functions as an inhibitor of apoptosis. As the balance between these two proteins is crucial for determining cell fate, their protein expression was also scrutinized. It was observed that the expression of *Bax* increases (TMPyP4=40 and 63% and BRACO-19=18 and 72%) while the expression of *BCL-2* decrease (TMPyP4=86 and 56% and BRACO-19=82 and 59%) in presence of **TMPyP4** and **BRACO-19**. This may be due to the decreased expression of *mTOR* on treatment with G4 stabilizers which trigger apoptosis in cancer cells.

Conclusion

In summary, we have identified a novel G-rich DNA sequence in the 5' UTR region of *mTOR* mRNA which can fold into the G-quadruplex structure as well as play a crucial role in regulating gene expression. In addition to this finding, it is also notable to observe that *mTOR* promoter and UTR does not have many potential G4 sequences and that only one G-rich sequence was observed to fold into the G4 conformation out of all the studied G-rich PQSs. Furthermore, the occurrence of a G4 in the non-template DNA strand prompted and eventually led to the study of the RNA G4 formation in the mRNA of *mTOR* which exercises a higher degree of control on *mTOR* translation. This makes the finding even more unique and therapeutically relevant as the rP2 sequence can have a high degree

of tunability in terms of regulating translation of *mTOR*. A series of biophysical techniques including CD spectroscopy, mobility shift assay, NMR spectroscopy and DMS footprinting was employed to observe the folding pattern of rP2 G4. The contribution of individual G-tracts in the folding of G4 formed by rP2 was also evaluated and resulted in the observation of four non-consecutive G-tracts emerging as highly significant in the formation of the G-quadruplex. The mutation of even a single G-tract involved in G4 formation results in appreciable perturbation of the G4 formation.

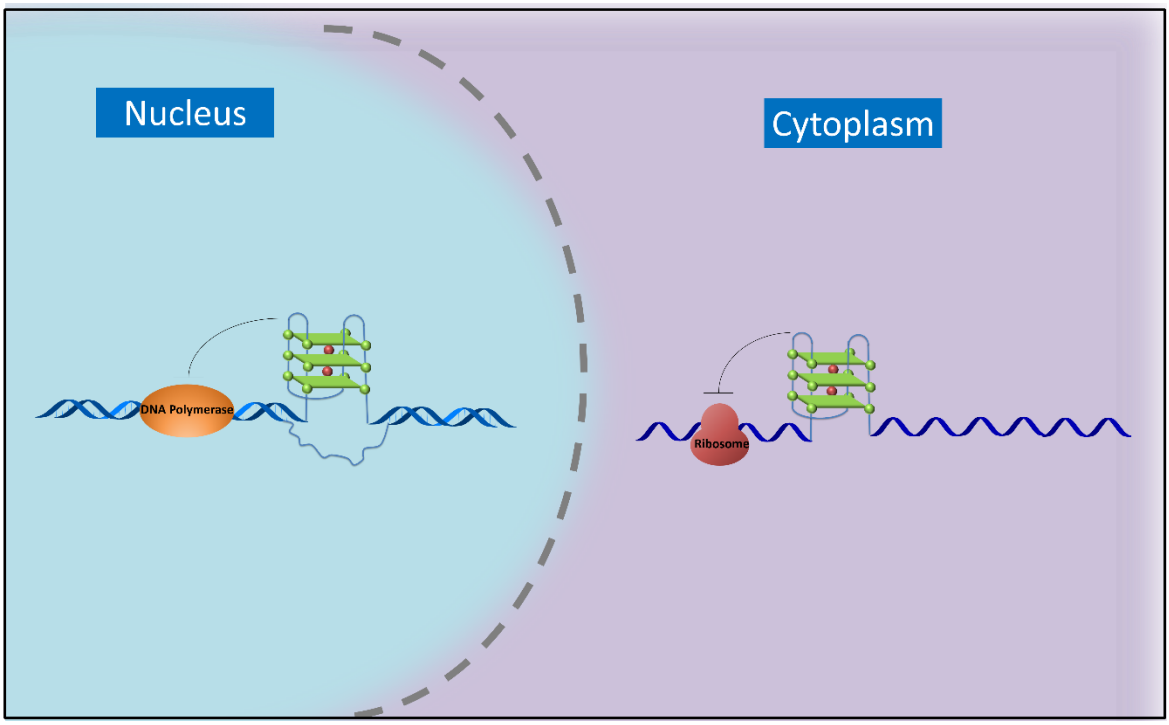


Figure 4.23. Schematic representation of the formation of mTOR G-quadruplexes in the DNA (in nucleus) and RNA (in cytoplasm) with cellular effects on the processes of replication and translation.

The relevance of rP2 in *mTOR* gene expression is further confirmed by the binding studies and gene expression investigations of *mTOR* carried out in the presence of extensively well studied G4 ligands like **TMPyP4** and **BRACO-19**. This raises the possibility that the transcriptional/translational control of critical genes like *mTOR* can be

modulated by the use of G-quadruplex-interactive agents. The lack of any additional G4s near the TSS of *mTOR* gene and the UTR region of the *mTOR* mRNA makes this finding even more significant. The degree of tunability exercised by the rP2 G4 sequence can result in regulation of *mTOR* gene, making this finding highly relevant for future studies. This study aimed to identify and investigate a novel G4 secondary structure present in a highly critical gene of interest involved in multiple fundamental processes of the cell. Thus unique structural scaffold provides a specific nucleic acid target for future design and synthesis of *mTOR* specific nucleic acid targeting ligands. This study also serves to identify diverse targets for alternative therapeutic approaches to control overexpression of *mTOR* gene paving the path for detailed studies for therapeutic intervention in this frontier.

Chapter 5

Material and Methods

All DNA and RNA oligonucleotides were purchased from Sigma-Aldrich. The oligonucleotide (unlabeled and labeled with fluorophores) sequences were purchased from Sigma Aldrich. Cell culture reagents were purchased from Thermo Fisher Scientific and Himedia unless stated otherwise. The XTT reagent, Annexin V-FITC kit for apoptosis assay, propidium iodide and BG4 antibody were purchased from Thermo Fisher Scientific. The luciferase reporter assay kit was purchased from Promega.

5.1. FRET Melting Assay

FRET melting assay was carried out using dual fluorophore labeled (6-FAM at 5'-end and TAMRA at 3'-end) oligonucleotides. The oligonucleotides were annealed at a concentration of 400 nM by heating and holding the temperature at 95 °C for 5 min in 60 mM potassium cacodylate buffer (pH 7.4) followed by gradual cooling to room temperature at a controlled rate of 0.1 °C/min and kept overnight at 4 °C. The working dilutions of the labeled DNA were made by diluting in 60 mM potassium cacodylate buffer (pH 7.4) to a final concentration of 200 nM. The samples were loaded into 96 well plates. Ligands dilutions of ligands were added at respective concentrations (1 µM, 3 µM, 5 µM, 7 µM and 10 µM) to the samples and incubated in dark for 1 h.

Competitive FRET assay was performed using dual labeled *c-KIT 1* (5'FAM-GGGAGGGCGCTGGGAGGAGGG-TAMRA3') DNA and unlabeled double-stranded calf thymus (CT) DNA as the DNA competitor. **MC-4** and **MD-11** were added at 1 µM concentration to pre-folded labeled *c-KIT 1* G4. Several fold excess concentrations of CT DNA was added (0, 1, 5, 10, 50, 100 eq. excess of the labeled G-quadruplex DNA) to the samples compared to the labeled G4 (200 nM) and incubated. The samples were heated up to 94 °C, and the fluorescence of the donor FAM (emission at 530 nm) in a LightCycler 480-II System (Roche). The melting profiles of labeled DNA were processed and analyzed in Origin Pro 2018 software and the Tm values were obtained.

All samples were incubated in the dark for 1 h at 25 °C. The samples were heated up to 94 °C, and the fluorescence of the donor FAM (emission at 530 nm) was monitored at intervals of 1 °C in the range of 37-95 °C in a LightCycler 480-II System (Roche). The melting profiles of labeled DNA were processed and ana-lyzed in Origin Pro 2018 software from

the first derivative curves, and the Tm values were obtained. The following oligonucleotides have been used in the study.

c-KIT 1 5'FAM-GGGAGGGCGCTGGGAGGAGGG-TAMRA3'

BCL-2 5'FAM-AGGGGCGGGCGCGGGAGGAAGGGGGCGGGAGCGGGGCTG-TAMRA3'

mTOR P2 5'FAM- AGCGGGGAAGGCGGGCGGTGGGGCAGGGGGCCT-TAMRA3'

VEGF 5'FAM-CGGGGCGGGCCGGGGGCGGGGT-TAMRA3'

h-TELO 5'FAM-TTGGGTTAGGGTTAGGGTTAGGGA-TAMRA3'

*ds*DNA 5'FAM-CAATCGGATCGAATTCGATCCGATTG-TAMRA3'

5.2.Fluorescence spectroscopy.

Fluorescence spectra were recorded on a HORIBA Jobin Yvon FluoroLog 3 spectrofluorometer in a 1 mm path-length quartz cuvette. All experiments were carried out in 0.5 mL final volume of **ligand** in 100 mM Tris KCl buffer (pH 7.4). The ligands were excited at absorbance maxima (**MC-4** – 281nm, **TMPyP4** - 421 nm) to obtain fluorescence emission spectra . Titrations were carried out using 1 μM ligand concentration of **ligand** with successive additions of different pre-annealed G-quadruplex DNAs and *ds*DNA until saturation. All the binding constants of the fluorescence spectral data have been calculated using the following equation with the help of Origin 2018.

$$F = F_0 + \frac{(F_{max} - F_0)[DNA]}{K_d + [DNA]}$$

c-KIT 1: 5' -GGGAGGGCGCTGGGAGGAGGG-3'

c-KIT 2: 5'-FAM-GGGCGGGCGCGAGGGAGGGG-3'

BCL-2: 5'-AGGGGCGGGCGCGGGAGGAAGGGGGCGGGAGCGGGGCTG-3'

VEGF: 5'-CGGGGCGGGCCGGGGGCGGGGT-3'

h-TELO: 5'-TTGGGTTAGGGTTAGGGTTAGGGA-3'

c-MYC: 5'- TGGGGAGGGTGGGGAGGGTGGGGAAGG-3'

mTOR P2: 5'- AGCGGGGAAGGCGGGCGGTGGGGCAGGGGGCCT-3'

mTOR rP2: 5'- AGCGGGGAAGGCGGGCGGUGGGGCAGGGGGCCU-3'

*ds*DNA: 5'-CAATCGGATCGAATTCGATCCGATTG-3'

5.3.Fluorescence intercalating displacement assay

Fluorescence intercalating displacement assay was performed using thiazole orange (TO) as the intercalator and pre-folded quadruplexes. For each sample, 10 μM of unlabeled G4 forming were diluted to 0.25 μM in 60 mM K-cacodylate buffer (pH 7.4) and TO was added to the final concentration of 0.5 μM. The fluorescene spectrum of mixture of DNA and TO was recorded and then then the TO-DNA mixture titrated with ligands **IQ1** and **IQ2** (20 μM stock in respective buffer) and the fluorescence spectrum was recorded after a 3-min equilibration period in Horiba JobinYvonFluorolog 3 fluorescence spectrometer. The final percentage of TO displacement was calculated from the fluorescence area (FA, 510-700 nm, λex = 501 nm), using,

% of TO displacement= 100-[FA x 100/FA_0]

FA_0 being the fluorescence area from TO bound DNA without ligand addition. The percentages of TO displacement were plotted as a function of the concentration of the ligands. The DNA sequences used in this study are:

h-TELO	5'-TTGGGTTAGGGTTAGGGTTAGGGA-3'
*ds*DNA	5'-CAATCGGATCGAATTCGATCCGATTG-3'

5.4.Circular Dichroism spectroscopy.

CD spectra were recorded on a JASCO J-815 spectrophotometer by using a 1 mm path length quartz cuvette in the region of 200 nm to 350 nm. Aliquots of ligands were added in increments to the pre-annealed G-quadruplex DNA (15 μM) in Tris KCl (100 mM)

buffer at pH 7.4. Each CD spectrum represents an average of three scans, smoothed, and zero corrected. The final analysis of the data was carried out by using Origin 2018. The DNA and RNA sequences of the G-quadruplexes are as follows:

c-KIT 2: 5'-GGGCGGGCGCGAGGGAGGGG-3'

h-TELO: 5'-TTGGGTTAGGGTTAGGGTTAGGGA-3'

mTOR P1: 5'-AGCGGGGAAGGCGGGCGGTGGGGCAGGGGGCCT-3'

mTOR P2: 5'-TTTGGTTAAAACTTGGCTATGGTTTCTCTGGAGT-3'

mTOR P3 : 5'-AGCGGCGGTACCGGTGCTGGCGGCGGCAG-3'

mTOR P4: 5'-CTAGGGCCTGGCACGACCCCTCTAGGGCGGCGT-3'

mTOR P5:- 5'-TTCGGCCCGCGGTACGTCCGCCCGAAGGCCGGCCC-3'

mTOR rP2: 5' AGCGGGGAAGGCGGGCGGUGGGGCAGGGGGCCU-3'

mTOR rP3: 5'-AGCGGCGGUACCGGUGCUGGCGGCGGCAG-3'

5.5.CD Melting assay

CD spectra were recorded on a JASCO J-1500 spectrophotometer equipped with a peltier by using a 1 mm path length quartz cuvette. All scans were taken in the range of 200-350 nm with a scanning speed of 100 nm/min. Respective sample concentrations of PQS and the P2 mutants were prepared in 10 mM Tris.HCl buffer (pH 7.4) in the presence or absence of 100 mM KCl or NaCl or LiCl respectively. Samples were annealed by heating to 95 °C for 5 min followed by slow cooling to room temperature. At least three scans were averaged at 25 °C and the buffer baseline was subtracted for each CD spectrum.

For the determination of T_m values, the samples were prepared at concentrations of 10 µM, 25 µM and 50 µM, 100 µM and annealed as per stated above. The molar ellipticity was measured at 262 nm over a range of 20 °C-90 °C and the melting temperature (T_m) was determined. The data were plotted as ellipticity vs temperature and the T_m was determined at half maximum reading. The data analysis was carried out by using Origin Pro 2018.

c-KIT 2: 5'-GGGCGGGCGCGAGGGAGGGG-3'

h-TELO: 5'-TTGGGTTAGGGTTAGGGTTAGGGA-3'

mTOR P1: 5'-AGCGGGGAAGGCGGGCGGTGGGGCAGGGGGCCT-3'

mTOR P2: 5'-TTTGGTTAAAACTTGGCTATGGTTTCTCTGGAGT-3'

mTOR P3 : 5'-AGCGGCGGTACCGGTGCTGGCGGCGGCAG-3'

mTOR P4: 5'-CTAGGGCCTGGCACGACCCCTCTAGGGCGGCGT-3'

mTOR P5:- 5'-TTCGGCCCGCGGTACGTCCGCCCGAAGGCCGGCCC-3'

mTOR rP2: 5' AGCGGGGAAGGCGGGCGGUGGGGCAGGGGGCCU-3'

5.6.UV/Vis Spectroscopy

Samples of the ligands were prepared by diluting to a concentration of 100 µM in 100 mM Tris.HCl, 100 mM KCl buffer (pH 7.4). The samples were added in 96 well plate and the absorbance maxima was measured using

Absorption spectra were recorded Cary Win 300 ultraviolet-visible spectrophotometer using quartz cuvette of path length 1 mm at 25 °C. Sample of 5 µM TmPyP4 (Sigma, Alrich) was prepared in 10mM Tris.HCl, 100 mM KCl buffer (pH 7.4). The scans were recorded in the range of 200-800 nm. The sample was titrated with increasing concentrations of the pre-annealed PQS until saturation was reached. UV/Vis spectra were recorded after each addition of concentrated PQS. The buffer measurement was subtracted for each experiment. The data was analysed using Origin Pro 2018 and binding co-efficient (K_d) was found.

5.7.Isothermal titration calorimetry

Isothermal titration calorimetry (ITC) studies were carried out on a Microcal PEAQ-ITC microcalorimeter (Malvern, USA). Each prefolded G4 DNA was diluted to a working concentration of 5 µM in 100 mM Tris-HCl, and 100 mM KCl buffer (pH 7.4). DNA to ligand ratio was taken as 1:10, and the titration was carried out at 25 °C with 30 injections

at uniform intervals. The ITC profiles were fitted to a single binding site model. DNA sequences used for ITC studies are as follows.

c-KIT 1	5'-GGGAGGGCGCTGGGAGGAGGG-3'
c-MYC:	*5'*-FAM-TGGGGAGGGTGGGGAGGGTGGGGAAGG-TAMRA- *3*
BCL-2	5'-AGGGGCGGGCGCGGGAGGAAGGGGGCGGGAGCGGGGCTG-3'
VEGF	5'-CGGGGCGGGCCGGGGGCGGGGT-3'
h-TELO	5'-TTGGGTTAGGGTTAGGGTTAGGGA-3'
*ds*DNA	5'-CAATCGGATCGAATTCGATCCGATTG-3'

5.8.Molecular Docking

Molecular docking was performed using the Auto-Dock 4.2 program to study the interactions of the monocarbazole derivative **MC-4** with a series of quadruplexes. The energy-minimized structure of **MC-4** was obtained with Gaussian 03 using density functional theory (DFT) analysis B3LYP/6-31+G(d) level. The structures of the G-quadruplexes were obtained from RCSB-Protein Data Bank and used for docking. 20 docking calculations were performed using the Lamarckian genetic algorithm with the default parameters from Auto-Dock 4.0 program. All Docking calculations were performed without water and counter ions. A maximum of 25 million energy evaluations were applied for the experiment. The docked complex structures were imaged using Chimera 1.11.2 software. The PDB IDs of the G4s used in the study are as follows: *BCL-2* G4 (PDB ID: 2F8U); *VEGF* G4 (PDB ID: 2M27); *c-KIT 1* G4 (PDB ID: 4WO2); *c-KIT 2* G4 (PDB ID: 2KQH), *h-TELO* (PDB ID: 2JPZ) and *c-MYC* (PDB ID: 1XAV).

5.9.NMR Spectroscopy

All NMR experiments were performed using Brüker ADVANCE 600 MHz NMR spectrometer. DNA oligonucleotides were prepared in 100 mM K^+ buffer (pH 7.4) solution to make a final concentration of 500 µM. The DNA and RNA samples were annealed at 95

°C for 5 mins followed by slow cooling to room temperature. 10% D_2O (v/v) was added to the sample and the sample was transferred to NMR tube for measurement. A sample of 10% D_2O (v/v) containing 100 mM K^+ buffer (pH 7.4) solution was taken as a control solution to exclude the effect of the buffer. Tetramethylsilane (TMS) was taken as internal control. 1D ^1H spectra were recorded at 25 °C and in water suppression parameter with 68 scans. The sequences of the oligonucleotide used in the study is as follows:

mTOR P2 5'- AGCGGGGAAGGCGGGCGGTGGGGCAGGGGGCCT-3'

mTOR rP2 5'- AGCGGGGAAGGCGGGCGGUGGGGCAGGGGGCCU-3'

5.10. NMR Titration

The *h-TELO* DNA was purchased from Eurofins MWG Operon in HPSF purity grade and further purified with HPLC, lyophilized, and desalted by using ultrafiltration devices (Vivaspin, MWCO 2 kDa). During the titration the DNA was provided as a 100 μM solution in 25 mM Tris·HCl buffer (pH 7.4) with 100 mM KCl in 5% d 6 -DMSO/95% H 2 O. Small amounts of the ligand stock solution in 100% d 6 -DMSO were added directly into the NMR tube with a final DMSO concentration of less than 7%. 2,2-dimethyl-2-silapentane-5-sulphonate (DSS) was used as internal reference. Gradient-assisted excitation sculpting pulse sequences or jumpreturnEcho were used for water suppression. The sequence of the oligonucleotide used in the study is as follows:

h-TELO 5'-TTGGGTTAGGGTTAGGGTTAGGGA-3'

5.11. Cell cytotoxicity assay

Cervical cancer cells (HeLa), Myelogenous leukemia cells (K562), breast cancer cells (MCF-7) as well as HEK293 non-cancerous human cells, were seeded on 96-well plates and treated with different ligand concentrations (5-20 μM). The cells were cultured in 10% Fetal Bovine Serum (FBS) containing respective medium (HeLa-DMEM, K562 –RPMI, MCF-7-MEM and HEK293- DMEM). After 24 h of treatment with ligands at 37 °C in 5%

CO_2, 25 µL of XTT/PMS solution was added to each well containing respective media and further incubated for 2 h. The XTT/PMS reagent was prepared by mixing 4 mg of XTT in 4 mL of culture media and then by adding 10 µL of 10 mM PMS solution (in PBS). The absorbance of XTT formazan was read at 450 nm on a Multiskan FC microplate spectrophotometer (Thermo Scientific). All the experiments were performed in triplicates. The cytotoxicity was evaluated based on the percentage of cell survival in a dose-dependent manner with respect to the control. The percentage of cell viability was calculated by using the following equation:

$$\% \text{ of cell viability} = \frac{Absorbance\ of\ treated\ cell}{Absorbance\ of\ DMSO\ Control} \times 100$$

5.12. Cell imaging

HeLa cells were cultured in 10 % FBS supplemented DMEM media on coverslips in six-well plates. The cells were treated with 2 µM of **IQ1** and **IQ2** for 24 h inside CO_2 (5%) incubator at 37 °C. The cells were harvested and washed twice with 1x PBS and fixed with chilled 1:1 acetone-methanol (1:1) for 12 min at - 20 °C. Cells were again washed thrice with 1x PBS at intervals of 10 minutes. Cells were stained with DAPI (Invitrogen) and mounted on glass slides. Images were taken in Zeiss microscope.

5.13. Immunofluorescence assay

K562 cells were cultured in RPMI supplemented with 10 % FBS and treated with 3 µM of MC-4 and 5 µM of MD-11 and incubated for 24 h inside CO_2 (5%) incubator at 37 °C. The cells were washed twice with 1x PBS and fixed with chilled 1:1 acetone-methanol (1:1) mixture for 12 min at - 20 °C. Cells were then washed thrice with 1x PBS at intervals of 10 minutes and treated with 0.03 % saponin for 30 min at RT. The cells were washed thrice with 1x PBS at intervals of 10 minutes and treated with 3% BSA. The cells were incubated overnight with BG4 antibody (1:100). Cells were washed thrice with 1x PBS followed by the addition of Alexa fluor 647 (1:1000) tagged secondary antibody for 2 h at RT. Cells were washed thrice with 1x PBS and mounted on slides. Images were taken in

Leica DM18 Stellaris 5 mi-croscope. The foci were analyzed and counted in Image J. An average of foci in different fields was calculated for final number of foci.

5.14. qRT-PCR

HeLa or K562 cells were cultured in 6-well plates and treated with respective ligands. After 24h, the cells were harvested Total RNA was isolated after 24 h ligand treatment using TRIzol reagent (Invitrogen, Life Technologies). RNA was quantified, and 1 µg of RNA was used for cDNA preparation using the Verso cDNA synthesis kit. The relative transcript expression level for genes was measured by quantitative real-time PCR (Roche Light Cycler II) using SYBR Green-based method. ΔCt values were calculated by the difference in threshold cycles (C_T) between test and control samples. *18s* rRNA gene was used as an internal control for normalizing the cDNA concentration of each sample. Primers used for monitoring the gene expressions are as follows;

18s rRNA Forward: 5'-GATTCCGTGGGTGGTGGTGC-3'

18s rRNA Reverse: 5'-AAGAAGTTGGGGGACGCCGA-3'

c-KIT Forward: 5′-CGTGGAAAAGAGAAAACAGTCA-3

c-KIT Reverse: 5′-CACCGTGATGCCAGCTATTA-3′

hTERT Forward: 5'- GCCTGAGCTGTACTTTGTCAA-3'

hTERT Reverse: 5'-AGGCTGCAGAGCAGCGTGGAGAGG-3'

c-MYC Forward: 5′- CTGCGACGAGGAGGAGGACT-3′

c-MYC Reverse: 5′-GGCAGCAGCTCGAATTTCTT-3′

BCL-2 Forward: 5′-ACAACATGGCCCTGTGGATGAC-3′

BCL-2 Reverse: 5′-TTGTTTGGGGCAGGCATGTT-3′

We used the comparative cycle threshold method (*C*T method) for relative quantification of gene expression.

Then the Δ*C*T values are calculated by the following equation:

Δ*C*T = *C*T target - *C*T reference …… (1)

where, *c-MYC, BCL-2 c-KIT and hTERT* are the target genes and *18s* is the reference gene.

Then the ΔΔ*C*T values are calculated by:

ΔΔ*C*T = -(ΔC_T test sample - ΔC_T calibrator sample) …… (2)

In our experiment, the Δ*C*T values of the untreated cells are the calibrator sample and the Δ*C*T values of the treated cells are test samples.

Lastly, the relative mRNA expression of the target genes was calculated by:

$2^{-(\Delta\Delta C_T)}$…… (3)

5.15. Polymerase stop assay

Polymerase stop arrest assay was performed using the P2 PQS template, and a primer sharing reverse complementarity with the terminally located G-tracts. DNA Taq polymerase (Invitrogen) was used to amplify DNA. The experiment was conducted in 25 µl reaction containing 1X PCR buffer, 4.25 mM $MgCl_2$, 2 µM template, 0.33 mM dNTPs, 2.5 units of Taq polymerase. TMPyP4 was added in respective samples in a dose-dependent manner followed by PCR amplification. PCR was performed in Veriti 96 well thermal cycler (applied Biosystems) with the thermal cycle of an initial denaturation at 95 °C for 5 mins, 25 cycles of 95 °C for 30 s, 64 °C for 30 s and a final extension at 72 °C for 1 minute. The PCR amplification products were resolved on 15% PAGE gel stained with ethidium bromide and visualized in Chemi-Doc (Bio-Rad Laboratories).

The sequences of the template and the primer are as follows:

Template sequence: *5'AGCGGGGAAGGCGGGCGGTGGGGCAGGGGGCCT 3'*

Primer sequence: *5'-CAAAGCTAATCGATCGAGGCCCCCTGC-3*

5.16. Luciferase expressing plasmid construct

A 120 bp DNA fragment containing wild type and Mutant *mTOR* sequence was synthesized and inserted in the downstream region of the SV40 promoter just before the coding sequence of the renilla luciferase expressing gene in pGL4.72 plasmid construct (Addgene). The construction and sequencing were done by Biobharati Life Sciences, India.

The inserts containing the P2 WT and Mutant sequences are as follows:

mTOR WT insert:

5'CTCCCGGCTTAGAGGACAGCGGGGAAGGCGGGCGGTGGGGCAGGGGGCCT
GAAGCGGCGGTACCGGTGCTGGCGGCGGCAGCTGAGGCCTTGGCCGAAGCC
GCGCGAACCTCAGGGCAAGCTCGA-3'

mTOR Mutant insert:

5'CTCCCGGCTTAGAGGACAGCGAAGAAAACGAGCATTGAAGCAGAAAGCCT
GAAGCGGCGGTACCGGTGCTGGCGGCGGCAGCTGAGGCCTTGGCCGAAGCC
GCGCGAACCTCAGGGCAAGCTCGA-3'

5.17. Reporter Luciferase assay

0.5 μg of respective plasmid constructs(*c-KIT*, *c-MYC*,and *BCL-2*) (Addgene) containing the promoter region of *c-KIT* G-quadruplex sequence, *c-MYC* G-quadruplex sequence, *BCL-2* G-quadruplex sequences respectively were transfected into K562 cells using Lipofectamine 2000 (Invitrogen). Transfection was performed in serum-free RPMI medium, and after 8 h, media was replaced with complete media and treated with respective doses of the ligands. For *mTOR* renilla luciferase expressing plasmid (0.5 μg) containing the *mTOR* insert in the 5'UTR region and SV40 as the promoter in PGL4.72, the transfection was carried out in HeLa cells cultured in DMEM and supplemented with 10% FBS. After 36 h, cells were harvested, lysed with passive lysis buffer, and luciferase activities of cell lysate were measured using a Dual-Luciferase reporter assay kit (Promega) on SpectraMax iD3 (Molecular Devices). The measurements were normalized against the estimated protein concentrations of each sample

5.18.DMS Footprinting assay

FITC- labeled RNA samples (100 mM, 10 µL) without K^+ and with K^+ were prepared by diluting with MilliQ water and Tris KCL buffer respectively. The samples were heated upto 95° C , temperature was held for 5 min and then the samples were allowed to cool gradually and stored overnight in 4° C. Then the samples were treated with dimethyl sulfate (DMS) in ethanol (10%, 10 mL) for 3 min. The reactions were stopped by addition of 10 µL of DMS stop solution, which consists of sodium acetate buffer (3 M, pH 7.0,)/2-mercaptoethanol /tRNA (100 mg/mL)/MilliQ (25/21/1/9, v/v/v/v). Ethanol (300 µL) was added to the resulting solutions at -80 °C, and these reaction mixtures were incubated for 30 min at -80 °C. The samples were centrifuged at 15,000 rpm for 30 min at 4 °C, and the supernatants were removed. To the DNA pellets were added sodium acetate buffer (3 M, 10 µL), MilliQ (100 µL) and ethanol (250 µL), and centrifugation was repeated under the same conditions as before. The resulting supernatants were removed to leave the DNA pellets. Ethanol in MilliQ (70%, 800 µL) was added to each sample. After centrifugation at 15,000 rpm for 5 min at 4 °C, the supernatants were removed to leave the DNA pellets (this was repeated twice). The resulting DNA pellets were air-dried, incubated with piperidine solution in MilliQ (10%, 100 µL) for 30 min at 95 °C, and then freeze-dried. The DNA pellets were taken up in 5.0 µL of loading buffer, which consists of formamide/EDTA (0.5 M, pH 8.0)/MilliQ (197/10/43, v/v/v), and the resulting samples were heated at 95 °C for 3 min, then cooled rapidly to 4 °C. Finally, 5.0 µL aliquots were loaded on 20% polyacrylamide gel containing 7 M urea and developed at 800 V for 2.5 h, in 1×TBE (Tris-Borate-EDTA) buffer. The gels were visualized in gel doc (Bio-rad).

5.19.Western Blotting assay

HeLa or K562 cells were cultured in 6-well plates and treated with respective ligands. After 24h, the cells were harvested. The cells were washed with 1X PBS, lysed with cell lysis buffer (20 mM Tris, 100 mM NaCl, 1 mM EDTA in 0.5% Tri-ton X-100) for 30 min, centrifuged at 12000g for 15 min and the supernatant was collected. The supernatants were used for protein determination by Folin-Lowry method. Gel electrophoresis was carried out after loading equal amounts of protein (50 µg) in 8 % SDS/PAGE, followed by electrophoretic transfer of proteins onto nitrocellulose membranes. After transfer, the

membranes were blocked with 4 % BSA for 1 h and incubated over-night at 4 °C with respective primary antibodies (1:1000) with gentle shaking. The membranes were washed thrice with 1x TBST and then incubated for 2 h with HRP linked secondary antibodies. Blots were washed thrice with 1x TBST and then once with 1x TBS. The bands were visualized by addition of chemiluminescent substrate (Pico PLUS and West Femto Chemilu-minescent Substrate, Thermo Fisher Scientific). Relative band intensities were determined using ImageJ software. Primary antibodies used are as follows:

Anti *GAPDH* antibody - Mouse origin (Invitrogen)

Anti *hTERT* antibody - Rabbit origin (Santa Cruz Technologies)

Anti *c-KIT* antibody - Rabbit origin (Cell Signaling Technologies)

Anti *mTOR* antibody - Rabbit origin (Abcam)

Anti *ATG 7* antibody - Rabbit origin (Cell Signalling Technologies)

Anti *ATG 12* antibody - Rabbit origin (Cell Signalling Technologies)

Anti *TRF 1* antibody - Rabbit origin (Santa Cruz Technologies)

Anti *c-MYC* antibody - Rabbit origin (Cell Signalling Technologies)

Anti *BCL-2* antibody - Rabbit origin (Invitrogen)

Anti *Bax* antibody - Mouse origin (Invitrogen)

5.20.Cell cycle assay.

HeLa or K562 cells were cultured at a density of 1 × 106 cells per well in six-well plates and allowed to grow for 24 h. In a fresh growth medium, cells were then treated with respective doses of ligands. After treatment with ligands for 24 h, cells were washed with 1XPBS, and trypsinized, and collected by centrifugation at 3000 RPM, washed twice with cold 1XPBS. Fixation of cells was done with 2 mL chilled 70% ethanol and the cells were incubated overnight at 4 °C. The cells were pelleted and washed twice with 1XPBS. Finally, the cells were resuspended in 350 µL 1XPBS containing 0.5 µg/ mL RNase A

(Invitrogen) and incubated for 30 min at 37 °C, followed by the addition of 5 µL of 1 mg/mL propidium iodide (PI) to each sample. Cell distribution data in different cell cycle phases were collected on a flow cytometer (BD Biosciences). A total of 10,000 events were recorded, and cell distribution among cell cycle phases was evaluated using BD FACS Diva 8.0.2 (BD).

5.21.DNA Damage Assay

HeLa cells were cultured at a density of 1×10^6 cells per well in six-well plates and allowed to grow for 24 h. Ligand treatment was done with different doses of ligands and incubated for 24 h. Cells were then centrifuged at 3000 rpm and washed with 0.1% FBS containing PBS (wash buffer) and resuspended in 1% paraformaldehyde (500 µL) and 0.03% saponin (500 µL) for 20 min at room temperature. Cells were washed with 0.1% FBS containing PBS and resuspended in 50 µL of diluted Alexa 647 tagged γ-H2AX antibody solution for approximately 4 h. Samples were run in flow cytometer (BD Biosciences) and a total number of 10,000 events were recorded for each sample. Cytogram analysis was done using the BD FACS Diva 8.0.2 (BD) and data was processed in Origin Pro 2018 software.

5.22.Apoptosis assay.

HeLa or K562 cells were cultured in six-well plates and grown for 24 h. Cells were treated with different concentrations of ligands for 24 h. After 24 h incubation, the cells were washed with 1xPBS, collected by centrifugation, and resuspended in 350 µL of 1x Annexin-V binding buffer (0.01 M HEPES, pH 7.4, 0.14 M NaCl, 2.5 mM $CaCl_2$). For detection of apoptosis, 5 µl 7AAD-annexin V and 2 µL Propidium Iodide (1 mg/mL) were added to the samples and incubated for 30 min in dark at 37 °C. Analyses of cells were performed on a BD FACS Diva 8.0.2 (BD) flow cytometer. A total of 10 000 events were recorded for each sample and final data processing and analysis was done in Origin Pro 2018 software.

5.23.Caspase 3/7

HeLa cells were cultured in six-well plates and grown for 24 h. Cells were treated with different concentrations of ligands for 24 h. After 24 h incubation, the cells were harvested and resuspended in 300 µl of DMEM culture media. 10 µL of 30X FLICA (FAM-DEVD-FMK caspase-3 and -7 reagent, Invitrogen) reagent was added to each sample and incubated in dark for 1h at 37 ºC and 5% CO_2. The tubes were flicked twice twice during incubation to minimize cell settling. After incubation, the cells were washed twice with 1x wash buffer and resuspended in 400 µL of 1x wash buffer. Analysis was performed on flow cytometer (BD Biosciences) with 488 nm excitation and green emission for the FLICA-stained cells. A total of 10 000 events were recorded for each sample and final data processing and analysis was done in Origin Pro 2018 software.

5.24.Caspase 8

HeLa cells were cultured in six-well plates and grown for 24 h. Cells were treated with different concentrations of ligands for 24 h. After 24 h incubation, the cells were harvested and resuspended in 300 µl of DMEM culture media. 10 µL of 30X FLICA (FAM-LETD-FMK caspase-8 reagent, Invitrogen) reagent was added to each sample and incubated in dark for 1h at 37ºC and 5% CO_2. The tubes were flicked twice twice during incubation to minimize cell settling. After incubation, the cells were washed twice with 1x wash buffer and resuspended in 400 µL of 1x wash buffer. Analysis was performed on flow cytometer (BD Biosciences) with 488 nm excitation and green emission for the FLICA-stained cells. A total of 10 000 events were recorded for each sample and final data processing and analysis was done in Origin Pro 2018 software.

5.25.ROS detection assay

HeLa cells were cultured in six-well plates and grown for 24 h. Cells were treated with different concentrations of ligands for 24 h. The cells were harvested and resuspended in serum-free media containing 10 µM of DCFDA reagent (Invitrogen) followed by

incubation in dark for 30 min at 37 ºC. Positive control sample was treted with 25 µM of H_2O_2. Increase in fluorescence intensity was examine in flow cytometer (BD Biosciences) monitored at 530nm with 488 nm excitation. A total of 10 000 events were recorded for each sample.

5.26.Telomerase Repeat Amplification Protocol (Q-TRAP) assay

Telomerase enzyme activity was measured in response to the ligand treatment of ligands at different doses using quantitative telomerase repeat amplification protocol (Q-TRAP) assay with TRAPeze Telomerase Detection Kit (Merck). HeLa cells were cultured in DMEM containing 10% FBS. The cells were harvested and washed with 1X DEPC-PBS followed by lysis with CHAPS lysis buffer and incubate on ice for 30 mins .The cells were centrifugation at 14000rpm for 20 mins at 4 °C and the protein was isolated. Telomerase based extension was performed using 5µg of the isolated cell extract in presence of 1X TRAP buffer, 150 µg TS primer, 10 mM dNTP and respective concentration of the ligands. Protein extract subjected to heat-inactivation was used as negative control. The PCR extension was performed followed by removal of excess ligands by DNA product extraction using with phenol/chloroform/isoamyl alcohol (50:48:2). The telomerase extended DNA products were treated with 2.5 M ammonium acetate and 3 volumes of 100% EtOH and kept at – 80 °C overnight. The DNA pellet was washed with 75% EtOH and dissolved in PCR grade water. Finally, Q-PCR amplification was performed with SYBR Green. The relative telomerase activity (RTA) of the each sample was measured and processed in Origin Pro 2018 software with respect to the standard curve of the cell extract using the equation:

RTA of the sample = 10[(Ct of the sample – m)/slope]

Summary

This thesis aimed to analyze and identify the G-quadruplexes which could be modulated using rationally synthesized small molecules capable of selective recognition, interaction, and regulation of gene expression in cancer cells. Moreover, the possibility of novel and unreported G-quadruplexes in the *mTOR* DNA and mRNA has been carried out and studied in detail to decipher their roles in the functioning of the *mTOR* gene and *mTOR* signaling pathway.

Chapter 1 provides a detailed outline of the G-quadruplex secondary structure found in DNA and RNA. It also delves into the regulation of this unique conformation using well reported small molecules to address treatment strategies for disease physiology.

Chapter 2 addresses the need to inhibit telomerase activity in cancer cells. The stabilization of the telomeric G-quadruplex is envisioned as an alternative therapeutic strategy to induce apoptosis in immortalized cancer cells. A library of rationally designed and synthesized **IQ** ligands were screened for their efficacy in binding to and consequently stabilizing the telomeric G-quadruplex known as *h-TELO*. Two small molecules, **IQ1** and **IQ2** emerged as lead ligands with **IQ2** showing a much higher affinity for *h-TELO* G. A series of biophysical methods were employed to study this interaction. Further biological investigations of these ligands showed that both these ligands have cytotoxic effects on cancer cells, activate the 3/7 caspase cascade, arrest cancer cells in G2/M phase of the cell cycle, trigger DNA damage in cells and increase the generation of reactive oxygen species further propelling the cells towards apoptosis. Furthermore, **IQ1** and **IQ2** cause significant decrease in relative telomere activity. In all biological investigations, **IQ2** was shown to induce a more significant biological response in cells as compared to **IQ1**. Thus, a multi-faceted indoloquinoxaline ligand capable of inducing a graded response in triggering DNA damage and inducing apoptosis may be a potential drug candidate to explore further in the future.

Chapter 3 delineates the **s**tabilization of *c-KIT 1* G-quadruplex using small molecules by exploring the optimal molecular architecture of aromatic chemical scaffold and side chains to achieve maximum binding interaction and cellular response in cancer cells. A library of

carbazole (**MC**) and dibenzofuran (**MD**) derivatives with identical side chain modifications were investigated for binding ability to G-quadruplexes. Biophysical evaluations revealed that the carbazole derivative **MC-4** efficiently stabilizes the *c-KIT 1* G-quadruplex. **MC-4** induced cytotoxicity and amplified the number of stabilized G4s in *c-KIT* expressing K562 cells. qRT-PCR assay and reporter luciferase assay further corroborated by the western blot assay establish that **MC-4** downregulates the expression of the *c-KIT* gene in K562 cells. **MC-4** was also found to arrest K562 cells in G0/G1 phase, triggers DNA damage and induces apoptosis. **MC-4** is a novel carbazole containing molecule that can provide an alternative approach to treat cancer cells with *c-KIT* upregulation.

Chapter 4 provides experimental evidences to support the presence of potential G-quadruplex in *mTOR* mRNA (rP2) which is present at the key location of the 5' UTR. Initial investigations to search for computationally predicted G-quadruplexes in the *mTOR* genome yielded the presence of a G-quadruplex, P2 in the non-coding strand of the *mTOR* gene near its TSS site. This sequence was found to interact with a well-explored porphyrin derivative, **TMPyP4**, known as one of the best G4 binders and capable of hampering the replicative machinery. Further analysis into the sequence let to the finding of a highly critical G-quadruplex, rP2 in the 5' UTR region of the *mTOR* mRNA which shares a similar structure with the P2 counterpart. rP2 was found to have an inhibitory role in the regulation of the translation of *mTOR* protein. The rP2 mutant was used to observe that without the G4 formation in the UTR region, the translation machinery can be upregulated. Investigations into the G-tracts responsible for formation of rP2 G-quadruplex showed that four non-consecutive G-tracts contribute to the G4 conformation. The effect of G4 binder, **TMPyP4** with rP2 using biophysical and biological assays showed that small molecules can downregulate *mTOR* gene expression *via* stabilization of G-quadruplex.

Hence, G-quadruplex structures present in both DNA and RNA exert regulatory influence over a spectrum of biological processes encompassing the central dogma within living cells. Perturbations in G4 dynamics are presumed to contribute to diseases, foremost among them being cancer, and the modulation of these structures through G4 ligands offers a promising avenue for cancer therapy. The work of this thesis has attempted to design and study pre-clinical drug candidates that hold potential in anti-cancer therapeutics. Another

aspect of the work has focused on identifying novel biomolecular targets such as *mTOR* which may be regulated using small molecules. In summation, the study of G4 structures, which represent a distinct three-dimensional conformation of nucleic acids, has ushered an era of biological exploration and innovative approaches in molecular medicine which needs further exploration and this thesis work has been an effort in this frontier.

References

(1) Jd, W. A structure for deoxyribose nucleic acid. *Nature* **1953**, *171*, 733-738.

(2) Zhao, J.; Bacolla, A.; Wang, G.; Vasquez, K. M. Non-B DNA structure-induced genetic instability and evolution. *Cellular and molecular life sciences* **2010**, *67*, 43-62.

(3) Hänsel-Hertsch, R.; Beraldi, D.; Lensing, S. V.; Marsico, G.; Zyner, K.; Parry, A.; Di Antonio, M.; Pike, J.; Kimura, H.; Narita, M. G-quadruplex structures mark human regulatory chromatin. *Nature genetics* **2016**, *48* (10), 1267-1272.

(4) Kouzine, F.; Wojtowicz, D.; Baranello, L.; Yamane, A.; Nelson, S.; Resch, W.; Kieffer-Kwon, K.-R.; Benham, C. J.; Casellas, R.; Przytycka, T. M. Permanganate/S1 nuclease footprinting reveals non-B DNA structures with regulatory potential across a mammalian genome. *Cell systems* **2017**, *4* (3), 344-356. e347.

(5) Watson, J.; Hays, F. A.; Ho, P. S. Definitions and analysis of DNA Holliday junction geometry. *Nucleic acids research* **2004**, *32* (10), 3017-3027.

(6) Georgakopoulos-Soares, I.; Chan, C. S.; Ahituv, N.; Hemberg, M. High-throughput techniques enable advances in the roles of DNA and RNA secondary structures in transcriptional and post-transcriptional gene regulation. *Genome biology* **2022**, *23* (1), 1-22.

(7) Chaudhuri, R.; Fatma, K.; Dash, J. Regulation of gene expression by targeting DNA secondary structures. *Journal of Chemical Sciences* **2021**, *133*, 1-17.

(8) Brooks, T. A.; Kendrick, S.; Hurley, L. Making sense of G-quadruplex and i-motif functions in oncogene promoters. *The FEBS journal* **2010**, *277* (17), 3459-3469.

(9) Bang, I. Untersuchungen über die Guanylsäure. *Biochem. Z* **1910**, *26*, 293-311.

(10) Gellert, M.; Lipsett, M. N.; Davies, D. R. Helix formation by guanylic acid. *Proceedings of the National Academy of Sciences* **1962**, *48* (12), 2013-2018.

(11) Davis, J. T. G-quartets 40 years later: from 5′-GMP to molecular biology and supramolecular chemistry. *Angewandte Chemie International Edition* **2004**, *43* (6), 668-698.

(12) Sen, D.; Gilbert, W. Formation of parallel four-stranded complexes by guanine-rich motifs in DNA and its implications for meiosis. *nature* **1988**, *334* (6180), 364-366.

(13) Todd, A. K.; Johnston, M.; Neidle, S. Highly prevalent putative quadruplex sequence motifs in human DNA. *Nucleic acids research* **2005**, *33* (9), 2901-2907.

(14) Huppert, J. L. Structure, location and interactions of G-quadruplexes. *The FEBS journal* **2010**, *277* (17), 3452-3458.

(15) Burge, S.; Parkinson, G. N.; Hazel, P.; Todd, A. K.; Neidle, S. Quadruplex DNA: sequence, topology and structure. *Nucleic acids research* **2006**, *34* (19), 5402-5415.

(16) Lane, A. N.; Chaires, J. B.; Gray, R. D.; Trent, J. O. Stability and kinetics of G-quadruplex structures. *Nucleic acids research* **2008**, *36* (17), 5482-5515.

(17) Sen, D.; Gilbert, W. A sodium-potassium switch in the formation of four-stranded G4-DNA. *Nature* **1990**, *344* (6265), 410-414.

(18) Palmer, B. F.; Clegg, D. J. Physiology and pathophysiology of potassium homeostasis. *Advances in physiology education* **2016**, *40* (4), 480-490.

(19) Huppert, J. L.; Balasubramanian, S. Prevalence of quadruplexes in the human genome. *Nucleic acids research* **2005**, *33* (9), 2908-2916.

(20) Spiegel, J.; Adhikari, S.; Balasubramanian, S. The structure and function of DNA G-quadruplexes. *Trends in Chemistry* **2020**, *2* (2), 123-136.

(21) Huppert, J. L.; Balasubramanian, S. G-quadruplexes in promoters throughout the human genome. *Nucleic acids research* **2007**, *35* (2), 406-413.

(22) Siddiqui-Jain, A.; Grand, C. L.; Bearss, D. J.; Hurley, L. H. Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription. *Proceedings of the National Academy of Sciences* **2002**, *99* (18), 11593-11598.

(23) Azzalin, C. M.; Reichenbach, P.; Khoriauli, L.; Giulotto, E.; Lingner, J. Telomeric repeat–containing RNA and RNA surveillance factors at mammalian chromosome ends. *Science* **2007**, *318* (5851), 798-801.

(24) Yangyuoru, P. M.; Di Antonio, M.; Ghimire, C.; Biffi, G.; Balasubramanian, S.; Mao, H. Dual Binding of an Antibody and a Small Molecule Increases the Stability of TERRA G-Quadruplex. *Angewandte Chemie International Edition* **2015**, *54* (3), 910-913.

(25) Xiao, C.-D.; Shibata, T.; Yamamoto, Y.; Xu, Y. An intramolecular antiparallel G-quadruplex formed by human telomere RNA. *Chemical Communications* **2018**, *54* (32), 3944-3946.

(26) Huppert, J. L.; Bugaut, A.; Kumari, S.; Balasubramanian, S. G-quadruplexes: the beginning and end of UTRs. *Nucleic acids research* **2008**, *36* (19), 6260-6268.

(27) Gibney, E.; Nolan, C. Epigenetics and gene expression. *Heredity* **2010**, *105* (1), 4-13.

(28) Kumari, S.; Bugaut, A.; Huppert, J. L.; Balasubramanian, S. An RNA G-quadruplex in the 5′ UTR of the NRAS proto-oncogene modulates translation. *Nature chemical biology* **2007**, *3* (4), 218-221.

(29) Morris, M. J.; Negishi, Y.; Pazsint, C.; Schonhoft, J. D.; Basu, S. An RNA G-quadruplex is essential for cap-independent translation initiation in human VEGF IRES. *Journal of the American Chemical Society* **2010**, *132* (50), 17831-17839.

(30) Gomez, D.; Lemarteleur, T.; Lacroix, L.; Mailliet, P.; Mergny, J. L.; Riou, J. F. Telomerase downregulation induced by the G-quadruplex ligand 12459 in A549 cells is mediated by hTERT RNA alternative splicing. *Nucleic acids research* **2004**, *32* (1), 371-379.

(31) Deng, Y.; Chan, S. S.; Chang, S. Telomere dysfunction and tumour suppression: the senescence connection. *Nature reviews cancer* **2008**, *8* (6), 450-458.

(32) Sauer, M.; Paeschke, K. G-quadruplex unwinding helicases and their function in vivo. *Biochemical Society Transactions* **2017**, *45* (5), 1173-1182.

(33) Mohaghegh, P.; Karow, J. K.; Brosh Jr, R. M.; Bohr, V. A.; Hickson, I. D. The Bloom's and Werner's syndrome proteins are DNA structure-specific helicases. *Nucleic acids research* **2001**, *29* (13), 2843-2849.

(34) Lattmann, S.; Stadler, M. B.; Vaughn, J. P.; Akman, S. A.; Nagamine, Y. The DEAH-box RNA helicase RHAU binds an intramolecular RNA G-quadruplex in TERC and associates with telomerase holoenzyme. *Nucleic acids research* **2011**, *39* (21), 9390-9404.

(35) González, V.; Hurley, L. H. The C-terminus of nucleolin promotes the formation of the c-MYC G-quadruplex and inhibits c-MYC promoter activity. *Biochemistry* **2010**, *49* (45), 9706-9714.

(36) Haeusler, A. R.; Donnelly, C. J.; Periz, G.; Simko, E. A.; Shaw, P. G.; Kim, M.-S.; Maragakis, N. J.; Troncoso, J. C.; Pandey, A.; Sattler, R. C9orf72 nucleotide repeat structures initiate molecular cascades of disease. *Nature* **2014**, *507* (7491), 195-200.

(37) Hanahan, D.; Weinberg, R. A. The hallmarks of cancer. cell, 100. *Transpl. Immunol* **2000**, *5*, 179-183.

(38) Brooks, T. A.; Hurley, L. H. The role of supercoiling in transcriptional control of MYC and its importance in molecular therapeutics. *Nature Reviews Cancer* **2009**, *9* (12), 849-861.

(39) Balasubramanian, S.; Hurley, L. H.; Neidle, S. Targeting G-quadruplexes in gene promoters: a novel anticancer strategy? *Nature reviews Drug discovery* **2011**, *10* (4), 261-275.

(40) Sun, D.; Hurley, L. H. The importance of negative superhelicity in inducing the formation of G-quadruplex and i-motif structures in the c-Myc promoter: implications for drug targeting and control of gene expression. *Journal of medicinal chemistry* **2009**, *52* (9), 2863-2874.

(41) Palm, W.; de Lange, T. How shelterin protects mammalian telomeres. *Annual review of genetics* **2008**, *42*, 301-334.

(42) Lim, C. J.; Cech, T. R. Shaping human telomeres: from shelterin and CST complexes to telomeric chromatin organization. *Nature Reviews Molecular Cell Biology* **2021**, *22* (4), 283-298.

(43) Hiyama, E.; Hiyama, K. Telomerase as tumor marker. *Cancer letters* **2003**, *194* (2), 221-233.

(44) Sun, D.; Thompson, B.; Cathers, B. E.; Salazar, M.; Kerwin, S. M.; Trent, J. O.; Jenkins, T. C.; Neidle, S.; Hurley, L. H. Inhibition of human telomerase by a G-quadruplex-interactive compound. *Journal of medicinal chemistry* **1997**, *40* (14), 2113-2116.

(45) Kim, M.-Y.; Vankayalapati, H.; Shin-Ya, K.; Wierzba, K.; Hurley, L. H. Telomestatin, a potent telomerase inhibitor that interacts quite specifically with the human telomeric intramolecular G-quadruplex. *Journal of the American Chemical Society* **2002**, *124* (10), 2098-2099.

(46) Maiti, S.; Saha, P.; Das, T.; Bessi, I.; Schwalbe, H.; Dash, J. Human telomeric G-quadruplex selective fluoro-isoquinolines induce apoptosis in cancer cells. *Bioconjugate Chemistry* **2018**, *29* (4), 1141-1154.

(47) Teng, F.-Y.; Jiang, Z.-Z.; Guo, M.; Tan, X.-Z.; Chen, F.; Xi, X.-G.; Xu, Y. G-quadruplex DNA: A novel target for drug design. *Cellular and Molecular Life Sciences* **2021**, *78* (19-20), 6557-6583.

(48) Dang, C. V. MYC on the path to cancer. *Cell* **2012**, *149* (1), 22-35.

(49) Dutta, D.; Debnath, M.; Müller, D.; Paul, R.; Das, T.; Bessi, I.; Schwalbe, H.; Dash, J. Cell penetrating thiazole peptides inhibit c-MYC expression via site-specific targeting of c-MYC G-quadruplex. *Nucleic acids research* **2018**, *46* (11), 5355-5365.

(50) Chaudhuri, R.; Prasanth, T.; Dash, J. Expanding the Toolbox of Target Directed Bio-Orthogonal Synthesis: In Situ Direct Macrocyclization by DNA Templates. *Angewandte Chemie International Edition* **2023**, *62* (7), e202215245.

(51) Brown, R. V.; Danford, F. L.; Gokhale, V.; Hurley, L. H.; Brooks, T. A. Demonstration that drug-targeted down-regulation of MYC in non-Hodgkins lymphoma is directly mediated through the promoter G-quadruplex. *Journal of Biological Chemistry* **2011**, *286* (47), 41018-41027.

(52) Adams, J. M.; Cory, S. The Bcl-2 protein family: arbiters of cell survival. *Science* **1998**, *281* (5381), 1322-1326.

(53) Amundson, S. A.; Myers, T. G.; Scudiero, D.; Kitada, S.; Reed, J. C.; Fornace Jr, A. J. An informatics approach identifying markers of chemosensitivity in human cancer cell lines. *Cancer research* **2000**, *60* (21), 6101-6110.

(54) Dexheimer, T. S.; Sun, D.; Hurley, L. H. Deconvoluting the structural and drug-recognition complexity of the G-quadruplex-forming region upstream of the bcl-2 P1 promoter. *Journal of the American Chemical Society* **2006**, *128* (16), 5404-5415.

(55) Wardelmann, E.; Merkelbach-Bruse, S.; Pauls, K.; Thomas, N.; Schildhaus, H.-U.; Heinicke, T.; Speidel, N.; Pietsch, T.; Buettner, R.; Pink, D. Polyclonal evolution of multiple secondary KIT mutations in gastrointestinal stromal tumors under treatment with imatinib mesylate. *Clinical Cancer Research* **2006**, *12* (6), 1743-1749.

(56) Bejugam, M.; Sewitz, S.; Shirude, P. S.; Rodriguez, R.; Shahid, R.; Balasubramanian, S. Trisubstituted isoalloxazines as a new class of G-quadruplex binding ligands: small molecule regulation of c-kit oncogene expression. *Journal of the American Chemical Society* **2007**, *129* (43), 12926-12927.

(57) Bejugam, M.; Gunaratnam, M.; Müller, S.; Sanders, D. A.; Sewitz, S.; Fletcher, J. A.; Neidle, S.; Balasubramanian, S. Targeting the c-Kit promoter G-quadruplexes with 6-substituted indenoisoquinolines. *ACS medicinal chemistry letters* **2010**, *1* (7), 306-310.

(58) Dash, J.; Shirude, P. S.; Hsu, S.-T. D.; Balasubramanian, S. Diarylethynyl amides that recognize the parallel conformation of genomic promoter DNA G-quadruplexes. *Journal of the American Chemical Society* **2008**, *130* (47), 15950-15956.

(59) Cox, A. D.; Fesik, S. W.; Kimmelman, A. C.; Luo, J.; Der, C. J. Drugging the undruggable RAS: Mission possible? *Nature reviews Drug discovery* **2014**, *13* (11), 828-851.

(60) Cogoi, S.; Paramasivam, M.; Spolaore, B.; Xodo, L. E. Structural polymorphism within a regulatory element of the human KRAS promoter: formation of G4-DNA recognized by nuclear proteins. *Nucleic acids research* **2008**, *36* (11), 3765-3780.

(61) Asamitsu, S.; Obata, S.; Yu, Z.; Bando, T.; Sugiyama, H. Recent progress of targeted G-quadruplex-preferred ligands toward cancer therapy. *Molecules* **2019**, *24* (3), 429.

(62) Sundquist, W. I.; Heaphy, S. Evidence for interstrand quadruplex formation in the dimerization of human immunodeficiency virus 1 genomic RNA. *Proceedings of the National Academy of Sciences* **1993**, *90* (8), 3393-3397.

(63) Tao, Y.; Zheng, Y.; Zhai, Q.; Wei, D. Recent advances in the development of small molecules targeting RNA G-quadruplexes for drug discovery. *Bioorganic Chemistry* **2021**, *110*, 104804.

(64) Monchaud, D.; Teulade-Fichou, M.-P. A hitchhiker's guide to G-quadruplex ligands. *Organic & biomolecular chemistry* **2008**, *6* (4), 627-636.

(65) Wang, Y.-H.; Yang, Q.-F.; Lin, X.; Chen, D.; Wang, Z.-Y.; Chen, B.; Han, H.-Y.; Chen, H.-D.; Cai, K.-C.; Li, Q. G4LDB 2.2: a database for discovering and studying G-quadruplex and i-Motif ligands. *Nucleic Acids Research* **2022**, *50* (D1), D150-D160.

(66) Burger, A. M.; Dai, F.; Schultes, C. M.; Reszka, A. P.; Moore, M. J.; Double, J. A.; Neidle, S. The G-quadruplex-interactive molecule BRACO-19 inhibits tumor growth, consistent with telomere targeting and interference with telomerase function. *Cancer research* **2005**, *65* (4), 1489-1496.

(67) Gunaratnam, M.; Greciano, O.; Martins, C.; Reszka, A. P.; Schultes, C. M.; Morjani, H.; Riou, J.-F.; Neidle, S. Mechanism of acridine-based telomerase inhibition and telomere shortening. *Biochemical pharmacology* **2007**, *74* (5), 679-689.

(68) Sun, D.; Guo, K.; Rusche, J. J.; Hurley, L. H. Facilitation of a structural transition in the polypurine/polypyrimidine tract within the proximal promoter region of the human VEGF gene by the presence of potassium and G-quadruplex-interactive agents. *Nucleic acids research* **2005**, *33* (18), 6070-6080.

(69) Guo, K.; Pourpak, A.; Beetz-Rogers, K.; Gokhale, V.; Sun, D.; Hurley, L. H. Formation of pseudosymmetrical G-quadruplex and i-motif structures in the proximal promoter region of the RET oncogene. *Journal of the American Chemical Society* **2007**, *129* (33), 10220-10228.

(70) Gonçalves, D. P.; Rodriguez, R.; Balasubramanian, S.; Sanders, J. K. Tetramethylpyridiniumporphyrazines—a new class of G-quadruplex inducing and stabilising ligands. *Chemical communications* **2006**, (45), 4685-4687.

(71) Fedoroff, O. Y.; Salazar, M.; Han, H.; Chemeris, V. V.; Kerwin, S. M.; Hurley, L. H. NMR-based model of a telomerase-inhibiting compound bound to G-quadruplex DNA. *Biochemistry* **1998**, *37* (36), 12367-12374.

(72) Seenisamy, J.; Bashyam, S.; Gokhale, V.; Vankayalapati, H.; Sun, D.; Siddiqui-Jain, A.; Streiner, N.; Shin-Ya, K.; White, E.; Wilson, W. D. Design and synthesis of an expanded porphyrin that has selectivity for the c-MYC G-quadruplex structure. *Journal of the American Chemical Society* **2005**, *127* (9), 2944-2959.

(73) Stanslas, J.; Hagan, D. J.; Ellis, M. J.; Turner, C.; Carmichael, J.; Ward, W.; Hammonds, T. R.; Stevens, M. F. Antitumor polycyclic acridines. 7. Synthesis and biological properties of DNA affinic tetra-and pentacyclic acridines. *Journal of medicinal chemistry* **2000**, *43* (8), 1563-1572.

(74) Pennarun, G.; Granotier, C.; Gauthier, L. R.; Gomez, D.; Hoffschir, F.; Mandine, E.; Riou, J.-F.; Mergny, J.-L.; Mailliet, P.; Boussin, F. D. Apoptosis related to telomere instability and cell cycle alterations in human glioma cells treated by new highly selective G-quadruplex ligands. *Oncogene* **2005**, *24* (18), 2917-2928.

(75) Piazza, A.; Boule, J.-B.; Lopes, J.; Mingo, K.; Largy, E.; Teulade-Fichou, M.-P.; Nicolas, A. Genetic instability triggered by G-quadruplex interacting Phen-DC compounds in Saccharomyces cerevisiae. *Nucleic acids research* **2010**, *38* (13), 4337-4348.

(76) Shin-ya, K.; Wierzba, K.; Matsuo, K.-i.; Ohtani, T.; Yamada, Y.; Furihata, K.; Hayakawa, Y.; Seto, H. Telomestatin, a novel telomerase inhibitor from Streptomyces anulatus. *Journal of the American Chemical Society* **2001**, *123* (6), 1262-1263.

(77) Drygin, D.; Siddiqui-Jain, A.; O'Brien, S.; Schwaebe, M.; Lin, A.; Bliesath, J.; Ho, C. B.; Proffitt, C.; Trent, K.; Whitten, J. P. Anticancer activity of CX-3543: a direct inhibitor of rRNA biogenesis. *Cancer research* **2009**, *69* (19), 7653-7661.

(78) Belmonte-Reche, E.; Morales, J. C. G4-iM Grinder: when size and frequency matter. G-Quadruplex, i-Motif and higher order structure search and analysis tool. *NAR genomics and bioinformatics* **2020**, *2* (1), lqz005.

(79) Garant, J.-M.; Perreault, J.-P.; Scott, M. S. Motif independent identification of potential RNA G-quadruplexes by G4RNA screener. *Bioinformatics* **2017**, *33* (22), 3532-3537.

(80) Lago, S.; Nadai, M.; Cernilogar, F. M.; Kazerani, M.; Domíniguez Moreno, H.; Schotta, G.; Richter, S. N. Promoter G-quadruplexes and transcription factors cooperate to shape the cell type-specific transcriptome. *Nature communications* **2021**, *12* (1), 1-13.

(81) Puig Lombardi, E.; Londoño-Vallejo, A. A guide to computational methods for G-quadruplex prediction. *Nucleic acids research* **2020**, *48* (1), 1-15.

(82) Kim, N. W.; Piatyszek, M. A.; Prowse, K. R.; Harley, C. B.; West, M. D.; Ho, P. L.; Coviello, G. M.; Wright, W. E.; Weinrich, S. L.; Shay, J. W. Specific association of human telomerase activity with immortal cells and cancer. *Science* **1994**, *266* (5193), 2011-2015.

(83) Shay, J. W.; Wright, W. E. Telomerase therapeutics for cancer: challenges and new directions. *Nature reviews Drug discovery* **2006**, *5* (7), 577-584.

(84) Tan, Z.; Tang, J.; Kan, Z.-Y.; Hao, Y.-H. Telomere G-quadruplex as a potential target to accelerate telomere shortening by expanding the incomplete end-replication of telomere DNA. *Current topics in medicinal chemistry* **2015**, *15* (19), 1940-1946.

(85) Gomez, D.; O'Donohue, M.-F. o.; Wenner, T.; Douarre, C.; Macadré, J. r.; Koebel, P.; Giraud-Panis, M.-J. p.; Kaplan, H.; Kolkes, A.; Shin-Ya, K. The G-quadruplex ligand telomestatin inhibits POT1 binding to telomeric sequences in vitro and induces GFP-POT1 dissociation from telomeres in human cells. *Cancer Research* **2006**, *66* (14), 6908-6912.

(86) Neidle, S. Quadruplex nucleic acids as targets for anticancer therapeutics. *Nature Reviews Chemistry* **2017**, *1* (5), 1-10.

(87) Salvati, E.; Leonetti, C.; Rizzo, A.; Scarsella, M.; Mottolese, M.; Galati, R.; Sperduti, I.; Stevens, M. F.; D'Incalci, M.; Blasco, M. Telomere damage induced by the G-quadruplex ligand RHPS4 has an antitumor effect. *The Journal of clinical investigation* **2007**, *117* (11), 3236-3247.

(88) Pal, S.; Fatma, K.; Ravichandiran, V.; Dash, J. Triazolyl Dibenzo [a, c] phenazines Stabilize Telomeric G-quadruplex and Inhibit Telomerase. *Asian Journal of Organic Chemistry* **2021**, *10* (11), 2921-2926.

(89) Goodwin, S.; Smith, A.; Horning, E. Alkaloids of Ochrosia elliptica Labill. *Journal of the American Chemical Society* **1959**, *81* (8), 1903-1908.

(90) Juret, P.; Tanguy, A.; Girard, A.; Le Talaer, J.; Abbatucci, J.; Dat-Xuong, N.; Le Pecq, J.; Paoletti, C. Preliminary trial of 9-hydroxy-2-methyl ellipticinium (NSC 264--137) in advanced human cancers. *European journal of cancer* **1978**, *14* (2), 205-206.

(91) Harmenberg, J.; Wahren, B.; Bergman, J.; Akerfeldt, S.; Lundblad, L. Antiherpesvirus activity and mechanism of action of indolo-(2, 3-b) quinoxaline and analogs. *Antimicrobial agents and chemotherapy* **1988**, *32* (11), 1720-1724.

(92) Opatowsky, Y.; Lax, I.; Tomé, F.; Bleichert, F.; Unger, V. M.; Schlessinger, J. Structure, domain organization, and different conformational states of stem cell factor-induced intact KIT dimers. *Proceedings of the National Academy of Sciences* **2014**, *111* (5), 1772-1777.

(93) Edling, C. E.; Hallberg, B. c-Kit—a hematopoietic cell essential receptor tyrosine kinase. *The international journal of biochemistry & cell biology* **2007**, *39* (11), 1995-1998.

(94) Lennartsson, J.; Rönnstrand, L. Stem cell factor receptor/c-Kit: from basic science to clinical implications. *Physiological reviews* **2012**, *92* (4), 1619-1649.

(95) Wang, W.; Shui, L.; Liu, Y.; Zheng, M. C-Kit, a double-edged sword in liver regeneration and diseases. *Frontiers in Genetics* **2021**, *12*, 598855.

(96) Ke, H.; Kazi, J. U.; Zhao, H.; Sun, J. Germline mutations of KIT in gastrointestinal stromal tumor (GIST) and mastocytosis. *Cell & Bioscience* **2016**, *6* (1), 1-10.

(97) Katagiri, S.; Chi, S.; Minami, Y.; Fukushima, K.; Shibayama, H.; Hosono, N.; Yamauchi, T.; Morishita, T.; Kondo, T.; Yanada, M. Mutated KIT Tyrosine Kinase as a Novel Molecular Target in Acute Myeloid Leukemia. *International journal of molecular sciences* **2022**, *23* (9), 4694.

(98) Zhong, L.; Li, Y.; Xiong, L.; Wang, W.; Wu, M.; Yuan, T.; Yang, W.; Tian, C.; Miao, Z.; Wang, T. Small molecules in targeted cancer therapy: Advances, challenges, and future perspectives. *Signal transduction and targeted therapy* **2021**, *6* (1), 1-48.

(99) Galanis, A.; Levis, M. Inhibition of c-Kit by tyrosine kinase inhibitors. *Haematologica* **2015**, *100* (3), e77.

(100) Heinrich, M. C.; Marino-Enriquez, A.; Presnell, A.; Donsky, R. S.; Griffith, D. J.; McKinley, A.; Patterson, J.; Taguchi, T.; Liang, C.-W.; Fletcher, J. A. Sorafenib Inhibits Many Kinase Mutations Associated with Drug-Resistant Gastrointestinal Stromal TumorsSorafenib Inhibits Kinase Mutations in Drug-Resistant GIST. *Molecular cancer therapeutics* **2012**, *11* (8), 1770-1780.

(101) López-Mejía, J. A.; Tallabs-Utrilla, L. F.; Salazar-Sojo, P.; Mantilla-Ollarves, J. C.; Sánchez-Carballido, M. A.; Rocha-Zavaleta, L. c-Kit Induces Migration of Triple-Negative Breast Cancer Cells and Is a Promising Target for Tyrosine Kinase Inhibitor Treatment. *International journal of molecular sciences* **2022**, *23* (15), 8702.

(102) Savage, D. G.; Antman, K. H. Imatinib mesylate—a new oral targeted therapy. *New England Journal of Medicine* **2002**, *346* (9), 683-693.

(103) Ilaria, F.; Valentina, P.; Sara, N. R.; Filippo, D. Multimeric G-quadruplexes: A review on their biological roles and targeting. *International Journal of Biological Macromolecules* **2022**.

(104) Shan, C.; Yan, J.-W.; Wang, Y.-Q.; Che, T.; Huang, Z.-L.; Chen, A.-C.; Yao, P.-F.; Tan, J.-H.; Li, D.; Ou, T.-M. Design, synthesis, and evaluation of isaindigotone derivatives to

downregulate c-myc transcription via disrupting the interaction of NM23-H2 with G-quadruplex. *Journal of Medicinal Chemistry* **2017**, *60* (4), 1292-1308.

(105) Ranjan, N.; Davis, E.; Xue, L.; Arya, D. P. Dual recognition of the human telomeric G-quadruplex by a neomycin–anthraquinone conjugate. *Chemical Communications* **2013**, *49* (51), 5796-5798.

(106) Tan, D. J. Y.; Das, P.; Winnerdy, F. R.; Lim, K. W.; Phan, A. T. Guanine anchoring: a strategy for specific targeting of a G-quadruplex using short PNA, LNA and DNA molecules. *Chemical Communications* **2020**, *56* (44), 5897-5900.

(107) Fernando, H.; Reszka, A. P.; Huppert, J.; Ladame, S.; Rankin, S.; Venkitaraman, A. R.; Neidle, S.; Balasubramanian, S. A conserved quadruplex motif located in a transcription activation site of the human c-kit oncogene. *Biochemistry* **2006**, *45* (25), 7854-7860.

(108) Hsu, S.-T. D.; Varnai, P.; Bugaut, A.; Reszka, A. P.; Neidle, S.; Balasubramanian, S. A G-rich sequence within the c-kit oncogene promoter forms a parallel G-quadruplex having asymmetric G-tetrad dynamics. *Journal of the American Chemical Society* **2009**, *131* (37), 13399-13409.

(109) Ducani, C.; Bernardinelli, G.; Högberg, B. r.; Keppler, B. K.; Terenzi, A. Interplay of three G-quadruplex units in the KIT promoter. *Journal of the American Chemical Society* **2019**, *141* (26), 10205-10213.

(110) Da Ros, S.; Nicoletto, G.; Rigo, R.; Ceschi, S.; Zorzan, E.; Dacasto, M.; Giantin, M.; Sissi, C. G-Quadruplex modulation of SP1 functional binding sites at the KIT proximal promoter. *International journal of molecular sciences* **2020**, *22* (1), 329.

(111) Monsen, R. C.; DeLeeuw, L. W.; Dean, W. L.; Gray, R. D.; Chakravarthy, S.; Hopkins, J. B.; Chaires, J. B.; Trent, J. O. Long promoter sequences form higher-order G-quadruplexes: an integrative structural biology study of c-Myc, k-Ras and c-Kit promoter sequences. *Nucleic acids research* **2022**, *50* (7), 4127-4147.

(112) Tassinari, M.; Richter, S. N.; Gandellini, P. Biological relevance and therapeutic potential of G-quadruplex structures in the human noncoding transcriptome. *Nucleic Acids Research* **2021**, *49* (7), 3617-3633.

(113) Wei, D.; Husby, J.; Neidle, S. Flexibility and structural conservation in a c-KIT G-quadruplex. *Nucleic acids research* **2015**, *43* (1), 629-644.

(114) Tian, T.; Chen, Y.-Q.; Wang, S.-R.; Zhou, X. G-Quadruplex: a regulator of gene expression and its chemical targeting. *Chem* **2018**, *4* (6), 1314-1344.

(115) Ou, T.-M.; Lu, Y.-J.; Zhang, C.; Huang, Z.-S.; Wang, X.-D.; Tan, J.-H.; Chen, Y.; Ma, D.-L.; Wong, K.-Y.; Tang, J. C.-O. Stabilization of G-quadruplex DNA and down-regulation of oncogene c-myc by quindoline derivatives. *Journal of medicinal chemistry* **2007**, *50* (7), 1465-1474.

(116) Li, M.-L.; Yuan, J.-M.; Yuan, H.; Wu, B.-H.; Huang, S.-L.; Li, Q.-J.; Ou, T.-M.; Wang, H.-G.; Tan, J.-H.; Li, D. Design, Synthesis, and Evaluation of New Sugar-Substituted Imidazole Derivatives as Selective c-MYC Transcription Repressors Targeting the Promoter G-Quadruplex. *Journal of Medicinal Chemistry* **2022**, *65* (19), 12675-12700.

(117) Kuang, G.; Zhang, M.; Kang, S.; Hu, D.; Li, X.; Wei, Z.; Gong, X.; An, L.-K.; Huang, Z.-S.; Shu, B. Syntheses and evaluation of new bisacridine derivatives for dual binding of G-quadruplex and i-motif in regulating oncogene c-myc expression. *Journal of Medicinal Chemistry* **2020**, *63* (17), 9136-9153.

(118) Doria, F.; Nadai, M.; Folini, M.; Di Antonio, M.; Germani, L.; Percivalle, C.; Sissi, C.; Zaffaroni, N.; Alcaro, S.; Artese, A. Hybrid ligand–alkylating agents targeting telomeric G-quadruplex structures. *Organic & biomolecular chemistry* **2012**, *10* (14), 2798-2806.

(119) Tera, M.; Ishizuka, H.; Takagi, M.; Suganuma, M.; Shin-ya, K.; Nagasawa, K. Macrocyclic hexaoxazoles as sequence-and mode-selective G-quadruplex binders. *Angewandte Chemie International Edition* **2008**, *47* (30), 5557-5560.

(120) Ghosh, A.; Trajkovski, M.; Teulade-Fichou, M. P.; Gabelica, V.; Plavec, J. Phen-DC3 Induces Refolding of Human Telomeric DNA into a Chair-Type Antiparallel G-Quadruplex through Ligand Intercalation. *Angewandte Chemie* **2022**, *134* (40), e202207384.

(121) Tera, M.; Iida, K.; Ishizuka, H.; Takagi, M.; Suganuma, M.; Doi, T.; Shin-ya, K.; Nagasawa, K. Synthesis of a potent G-quadruplex-binding macrocyclic heptaoxazole. *ChemBioChem* **2009**, *10* (3), 431-435.

(122) Ishikawa, R.; Yasuda, M.; Sasaki, S.; Ma, Y.; Nagasawa, K.; Tera, M. Stabilization of telomeric G-quadruplex by ligand binding increases susceptibility to S1 nuclease. *Chemical Communications* **2021**, *57* (59), 7236-7239.

(123) Ranjan, N.; Arya, D. P. Targeting C-myc G-quadruplex: dual recognition by aminosugar-bisbenzimidazoles with varying linker lengths. *Molecules* **2013**, *18* (11), 14228-14240.

(124) Xue, L.; Ranjan, N.; Arya, D. P. Synthesis and spectroscopic studies of the aminoglycoside (neomycin)− perylene conjugate binding to human telomeric DNA. *Biochemistry* **2011**, *50* (14), 2838-2849.

(125) Gomez, D.; Guedin, A.; Mergny, J.-L.; Salles, B.; Riou, J.-F.; Teulade-Fichou, M.-P.; Calsou, P. A G-quadruplex structure within the 5′-UTR of TRF2 mRNA represses translation in human cells. *Nucleic acids research* **2010**, *38* (20), 7187-7198.

(126) Chen, J.; Cheng, M.; Salgado, G. F.; Stadlbauer, P.; Zhang, X.; Amrane, S.; Guédin, A.; He, F.; Šponer, J.; Ju, H. The beginning and the end: flanking nucleotides induce a parallel G-quadruplex topology. *Nucleic Acids Research* **2021**, *49* (16), 9548-9559.

(127) McLuckie, K. I.; Waller, Z. A.; Sanders, D. A.; Alves, D.; Rodriguez, R.; Dash, J.; McKenzie, G. J.; Venkitaraman, A. R.; Balasubramanian, S. G-quadruplex-binding benzo [a] phenoxazines down-regulate c-KIT expression in human gastric carcinoma cells. *Journal of the American Chemical Society* **2011**, *133* (8), 2658-2663.

(128) Cui, X.; Lin, S.; Yuan, G. Spectroscopic probing of recognition of the G-quadruplex in c-kit promoter by small-molecule natural products. *International journal of biological macromolecules* **2012**, *50* (4), 996-1001.

(129) Shen, F.-H.; Jin, J.; Li, J.; Wang, Y.; Zhu, S.-H.; Lu, Y.-J.; Ou, T.-M.; Huang, Z.-S.; Huang, M.; Huang, Z.-Y. The G-quadruplex ligand, SYUIQ-FM05, targets proto-oncogene c-kit transcription and induces apoptosis in K562 cells. *Pharmaceutical biology* **2013**, *51* (4), 447-454.

(130) Zorzan, E.; Da Ros, S.; Musetti, C.; Shahidian, L. Z.; Coelho, N. F. R.; Bonsembiante, F.; Létard, S.; Gelain, M. E.; Palumbo, M.; Dubreuil, P. Screening of candidate G-quadruplex ligands for the human c-KIT promotorial region and their effects in multiple in-vitro models. *Oncotarget* **2016**, *7* (16), 21658.

(131) Mazzini, S.; Gargallo, R.; Musso, L.; De Santis, F.; Aviñó, A.; Scaglioni, L.; Eritja, R.; Di Nicola, M.; Zunino, F.; Amatulli, A. Stabilization of c-KIT G-Quadruplex DNA structures by the RNA polymerase I inhibitors BMH-21 and BA-41. *International journal of molecular sciences* **2019**, *20* (19), 4927.

(132) Guo, Q.-L.; Su, H.-F.; Wang, N.; Liao, S.-R.; Lu, Y.-T.; Ou, T.-M.; Tan, J.-H.; Li, D.; Huang, Z.-S. Synthesis and evaluation of 7-substituted-5, 6-dihydrobenzo [c] acridine derivatives as new c-KIT promoter G-quadruplex binding ligands. *European journal of medicinal chemistry* **2017**, *130*, 458-471.

(133) Wang, X.; Zhou, C.-X.; Yan, J.-W.; Hou, J.-Q.; Chen, S.-B.; Ou, T.-M.; Gu, L.-Q.; Huang, Z.-S.; Tan, J.-H. Synthesis and evaluation of quinazolone derivatives as a new class of c-KIT G-quadruplex binding ligands. *ACS medicinal chemistry letters* **2013**, *4* (10), 909-914.

(134) Gunaratnam, M.; Swank, S.; Haider, S. M.; Galesa, K.; Reszka, A. P.; Beltran, M.; Cuenca, F.; Fletcher, J. A.; Neidle, S. Targeting human gastrointestinal stromal tumor cells with a quadruplex-binding small molecule. *Journal of medicinal chemistry* **2009**, *52* (12), 3774-3783.

(135) Mitteaux, J.; Lejault, P.; Wojciechowski, F.; Joubert, A.; Boudon, J.; Desbois, N.; Gros, C. P.; Hudson, R. H.; Boulé, J.-B.; Granzhan, A. Identifying G-Quadruplex-DNA-disrupting small molecules. *Journal of the American Chemical Society* **2021**, *143* (32), 12567-12577.

(136) Cosconati, S.; Rizzo, A.; Trotta, R.; Pagano, B.; Iachettini, S.; De Tito, S.; Lauri, I.; Fotticchia, I.; Giustiniano, M.; Marinelli, L. Shooting for selective druglike G-quadruplex binders: evidence for telomeric DNA damage and tumor cell death. *Journal of medicinal chemistry* **2012**, *55* (22), 9785-9792.

(137) Monchaud, D.; Allain, C.; Bertrand, H.; Smargiasso, N.; Rosu, F.; Gabelica, V.; De Cian, A.; Mergny, J.-L.; Teulade-Fichou, M.-P. Ligands playing musical chairs with G-quadruplex DNA: a rapid and simple displacement assay for identifying selective G-quadruplex binders. *Biochimie* **2008**, *90* (8), 1207-1223.

(138) Merle, P.; Evrard, B.; Petitjean, A.; Lehn, J.-M.; Teulade-Fichou, M.-P.; Chautard, E.; De Cian, A.; Guittat, L.; Tran, P. L. T.; Mergny, J.-L. Telomere targeting with a new G4 ligand enhances radiation-induced killing of human glioblastoma cells. *Molecular cancer therapeutics* **2011**, *10* (10), 1784-1795.

(139) Carvalho, J.; Mergny, J.-L.; Salgado, G. F.; Queiroz, J. A.; Cruz, C. G-quadruplex, Friend or Foe: The Role of the G-quartet in Anticancer Strategies. *Trends in molecular medicine* **2020**, *26* (9), 848-861.

(140) Schmidt, A. W.; Reddy, K. R.; Knölker, H.-J. Occurrence, biogenesis, and synthesis of biologically active carbazole alkaloids. *Chemical reviews* **2012**, *112* (6), 3193-3328.

(141) Głuszyńska, A. Biological potential of carbazole derivatives. *European journal of medicinal chemistry* **2015**, *94*, 405-426.

(142) Panda, D.; Saha, P.; Das, T.; Dash, J. Target guided synthesis using DNA nano-templates for selectively assembling a G-quadruplex binding c-MYC inhibitor. *Nature communications* **2017**, *8* (1), 1-11.

(143) Chang, C.-C.; Chien, C.-W.; Lin, Y.-H.; Kang, C.-C.; Chang, T.-C. Investigation of spectral conversion of d (TTAGGG) 4 and d (TTAGGG) 13 upon potassium titration by a G-quadruplex recognizer BMVC molecule. *Nucleic acids research* **2007**, *35* (9), 2846-2860.

(144) Głuszyńska, A.; Juskowiak, B.; Kuta-Siejkowska, M.; Hoffmann, M.; Haider, S. Carbazole derivatives' binding to c-KIT G-quadruplex DNA. *Molecules* **2018**, *23* (5), 1134.

(145) Platella, C.; Ghirga, F.; Zizza, P.; Pompili, L.; Marzano, S.; Pagano, B.; Quaglio, D.; Vergine, V.; Cammarone, S.; Botta, B. Identification of Effective Anticancer G-Quadruplex-Targeting Chemotypes through the Exploration of a High Diversity Library of Natural Compounds. *Pharmaceutics* **2021**, *13* (10), 1611.

(146) Jun, J. V.; Chenoweth, D. M.; Petersson, E. J. Rational design of small molecule fluorescent probes for biological applications. *Organic & biomolecular chemistry* **2020**, *18* (30), 5747-5763.

(147) Kemp, J. A.; Kwon, Y. J. Cancer nanotechnology: current status and perspectives. *Nano convergence* **2021**, *8* (1), 1-38.

(148) Panda, D.; Debnath, M.; Mandal, S.; Bessi, I.; Schwalbe, H.; Dash, J. A nucleus-imaging probe that selectively stabilizes a minor conformation of c-MYC G-quadruplex and down-regulates c-MYC transcription in human cancer cells. *Scientific reports* **2015**, *5* (1), 1-16.

(149) Müller, D.; Saha, P.; Panda, D.; Dash, J.; Schwalbe, H. Insights from binding on quadruplex selective carbazole ligands. *Chemistry–A European Journal* **2021**, *27* (50), 12726-12736.

(150) Renčiuk, D.; Zhou, J.; Beaurepaire, L.; Guédin, A.; Bourdoncle, A.; Mergny, J.-L. A FRET-based screening assay for nucleic acid ligands. *Methods* **2012**, *57* (1), 122-128.

(151) De Cian, A.; Guittat, L.; Shin-ya, K.; Riou, J.-F.; Mergny, J.-L. Affinity and selectivity of G4 ligands measured by FRET. In *Nucleic Acids Symposium Series*, 2005; Oxford University Press: Vol. 49, pp 235-236.

(152) Figueiredo, J.; Peitinho, D.; Campello, M. P. C.; Oliveira, M. C.; Paulo, A.; Mergny, J.-L.; Cruz, C. Screening of Scaffolds for the Design of G-Quadruplex Ligands. *Applied Sciences* **2022**, *12* (4), 2170.

(153) Maleki, P.; Mustafa, G.; Gyawali, P.; Budhathoki, J. B.; Ma, Y.; Nagasawa, K.; Balci, H. Quantifying the impact of small molecule ligands on G-quadruplex stability against Bloom helicase. *Nucleic acids research* **2019**, *47* (20), 10744-10753.

(154) Jiang, Y.; Chen, A.-C.; Kuang, G.-T.; Wang, S.-K.; Ou, T.-M.; Tan, J.-H.; Li, D.; Huang, Z.-S. Design, synthesis and biological evaluation of 4-anilinoquinazoline derivatives as new c-myc G-quadruplex ligands. *European Journal of Medicinal Chemistry* **2016**, *122*, 264-279.

(155) Panda, D.; Saha, P.; Chaudhuri, R.; Prasanth, T.; Ravichandiran, V.; Dash, J. A competitive pull-down assay using G-quadruplex DNA linked magnetic nanoparticles to determine specificity of G-quadruplex ligands. *Analytical chemistry* **2019**, *91* (12), 7705-7711.

(156) Guertin, D. A.; Sabatini, D. M. Defining the role of mTOR in cancer. *Cancer cell* **2007**, *12* (1), 9-22.

(157) Saxton, R. A.; Sabatini, D. M. mTOR signaling in growth, metabolism, and disease. *Cell* **2017**, *168* (6), 960-976.

(158) Luo, L.; Sun, W.; Zhu, W.; Li, S.; Zhang, W.; Xu, X.; Fang, D.; Grahn, T. H. M.; Jiang, L.; Zheng, Y. BCAT1 decreases the sensitivity of cancer cells to cisplatin by regulating mTOR-mediated autophagy via branched-chain amino acid metabolism. *Cell death & disease* **2021**, *12* (2), 1-13.

(159) Wang, B.; Zhang, W.; Zhang, G.; Kwong, L.; Lu, H.; Tan, J.; Sadek, N.; Xiao, M.; Zhang, J.; Labrie, M. Targeting mTOR signaling overcomes acquired resistance to combined BRAF and MEK inhibition in BRAF-mutant melanoma. *Oncogene* **2021**, *40* (37), 5590-5599.

(160) Ben-Sahra, I.; Manning, B. D. mTORC1 signaling and the metabolic control of cell growth. *Current opinion in cell biology* **2017**, *45*, 72-82.

(161) Sabers, C. J.; Martin, M. M.; Brunn, G. J.; Williams, J. M.; Dumont, F. J.; Wiederrecht, G.; Abraham, R. T. Isolation of a Protein Target of the FKBP12-Rapamycin Complex in Mammalian Cells (∗). *Journal of Biological Chemistry* **1995**, *270* (2), 815-822.

(162) Kim, D.-H.; Sarbassov, D. D.; Ali, S. M.; King, J. E.; Latek, R. R.; Erdjument-Bromage, H.; Tempst, P.; Sabatini, D. M. mTOR interacts with raptor to form a nutrient-sensitive complex that signals to the cell growth machinery. *Cell* **2002**, *110* (2), 163-175.

(163) Ma, X. M.; Blenis, J. Molecular mechanisms of mTOR-mediated translational control. *Nature reviews Molecular cell biology* **2009**, *10* (5), 307-318.

(164) Shaw, R. J.; Cantley, L. C. Ras, PI (3) K and mTOR signalling controls tumour cell growth. *Nature* **2006**, *441* (7092), 424-430.

(165) Laplante, M.; Sabatini, D. M. mTOR signaling in growth control and disease. *Cell* **2012**, *149* (2), 274-293.

(166) Oh, W. J.; Jacinto, E. mTOR complex 2 signaling and functions. *Cell cycle* **2011**, *10* (14), 2305-2316.

(167) Meric-Bernstam, F.; Gonzalez-Angulo, A. M. Targeting the mTOR signaling network for cancer therapy. *Journal of Clinical Oncology* **2009**, *27* (13), 2278.

(168) Laplante, M.; Sabatini, D. M. mTOR signaling at a glance. *Journal of cell science* **2009**, *122* (20), 3589-3594.

(169) Mossmann, D.; Park, S.; Hall, M. N. mTOR signalling and cellular metabolism are mutual determinants in cancer. *Nature Reviews Cancer* **2018**, *18* (12), 744-757.

(170) Zhang, Y.; Ng, P. K.-S.; Kucherlapati, M.; Chen, F.; Liu, Y.; Tsang, Y. H.; de Velasco, G.; Jeong, K. J.; Akbani, R.; Hadjipanayis, A. A pan-cancer proteogenomic atlas of PI3K/AKT/mTOR pathway alterations. *Cancer cell* **2017**, *31* (6), 820-832. e823.

(171) Gremke, N.; Polo, P.; Dort, A.; Schneikert, J.; Elmshäuser, S.; Brehm, C.; Klingmüller, U.; Schmitt, A.; Reinhardt, H. C.; Timofeev, O. mTOR-mediated cancer drug resistance suppresses autophagy and generates a druggable metabolic vulnerability. *Nature communications* **2020**, *11* (1), 1-15.

(172) Guri, Y.; Hall, M. N. mTOR signaling confers resistance to targeted cancer drugs. *Trends in cancer* **2016**, *2* (11), 688-697.

(173) Lamming, D. W.; Ye, L.; Sabatini, D. M.; Baur, J. A. Rapalogs and mTOR inhibitors as anti-aging therapeutics. *The Journal of clinical investigation* **2013**, *123* (3), 980-989.

(174) Collins, G. P.; Popat, R.; Stathis, A.; Krasniqi, F.; Eyre, T. A.; Ng, C. H.; El-Sharkawi, D.; Schmidt, C.; Wicki, A.; Ivanova, E. A dose-escalation (DE) study with expansion evaluating safety, pharmacokinetics and efficacy of the novel, balanced PI3K/mTOR inhibitor PQR309 in patients with relapsed or refractory lymphoma. *Blood* **2016**, *128* (22), 5893.

(175) Abraham, R. T.; Gibbons, J. J. The mammalian target of rapamycin signaling pathway: twists and turns in the road to cancer therapy. *Clinical cancer research* **2007**, *13* (11), 3109-3114.

(176) Peralba, J. M.; DeGraffenried, L.; Friedrichs, W.; Fulcher, L.; Grünwald, V.; Weiss, G.; Hidalgo, M. Pharmacodynamic evaluation of CCI-779, an inhibitor of mTOR, in cancer patients. *Clinical Cancer Research* **2003**, *9* (8), 2887-2892.

(177) Boulay, A.; Zumstein-Mecker, S.; Stephan, C.; Beuvink, I.; Zilbermann, F.; Haller, R.; Tobler, S.; Heusser, C.; O'Reilly, T.; Stolz, B. Antitumor efficacy of intermittent treatment schedules with the rapamycin derivative RAD001 correlates with prolonged inactivation of

ribosomal protein S6 kinase 1 in peripheral blood mononuclear cells. *Cancer research* **2004**, *64* (1), 252-261.

(178) Liu, G. Y.; Sabatini, D. M. mTOR at the nexus of nutrition, growth, ageing and disease. *Nature reviews Molecular cell biology* **2020**, *21* (4), 183-203.

(179) Chen, Y.; Zhou, X. Research progress of mTOR inhibitors. *European journal of medicinal chemistry* **2020**, *208*, 112820.

(180) Neshat, M. S.; Mellinghoff, I. K.; Tran, C.; Stiles, B.; Thomas, G.; Petersen, R.; Frost, P.; Gibbons, J. J.; Wu, H.; Sawyers, C. L. Enhanced sensitivity of PTEN-deficient tumors to inhibition of FRAP/mTOR. *Proceedings of the National Academy of Sciences* **2001**, *98* (18), 10314-10319.

(181) Venkatesan, A. M.; Dehnhardt, C. M.; Delos Santos, E.; Chen, Z.; Dos Santos, O.; Ayral-Kaloustian, S.; Khafizova, G.; Brooijmans, N.; Mallon, R.; Hollander, I. Bis (morpholino-1, 3, 5-triazine) derivatives: potent adenosine 5′-triphosphate competitive phosphatidylinositol-3-kinase/mammalian target of rapamycin inhibitors: discovery of compound 26 (PKI-587), a highly efficacious dual inhibitor. *Journal of medicinal chemistry* **2010**, *53* (6), 2636-2645.

(182) Freitag, H.; Christen, F.; Lewens, F.; Grass, I.; Briest, F.; Iwaszkiewicz, S.; Siegmund, B.; Grabowski, P. Inhibition of mTOR's catalytic site by PKI-587 is a promising therapeutic option for gastroenteropancreatic neuroendocrine tumor disease. *Neuroendocrinology* **2017**, *105* (1), 90-104.

(183) Chen, Z.; Venkatesan, A. M.; Dos Santos, O.; Delos Santos, E.; Dehnhardt, C. M.; Ayral-Kaloustian, S.; Ashcroft, J.; McDonald, L. A.; Mansour, T. S. Stereoselective synthesis of an active metabolite of the potent PI3 kinase inhibitor PKI-179. *The Journal of organic chemistry* **2010**, *75* (5), 1643-1651.

(184) Beaufils, F.; Cmiljanovic, N.; Cmiljanovic, V.; Bohnacker, T.; Melone, A.; Marone, R.; Jackson, E.; Zhang, X.; Sele, A.; Borsari, C. 5-(4, 6-Dimorpholino-1, 3, 5-triazin-2-yl)-4-(trifluoromethyl) pyridin-2-amine (PQR309), a potent, brain-penetrant, orally bioavailable, pan-class I PI3K/mTOR inhibitor as clinical candidate in oncology. *Journal of medicinal chemistry* **2017**, *60* (17), 7524-7538.

(185) Borsari, C.; Rageot, D.; Beaufils, F.; Bohnacker, T.; Keles, E.; Buslov, I.; Melone, A.; Sele, A. M.; Hebeisen, P.; Fabbro, D. Preclinical development of PQR514, a highly potent PI3K inhibitor bearing a difluoromethyl–pyrimidine moiety. *ACS medicinal chemistry letters* **2019**, *10* (10), 1473-1479.

(186) Basu, B.; Dean, E.; Puglisi, M.; Greystoke, A.; Ong, M.; Burke, W.; Cavallin, M.; Bigley, G.; Womack, C.; Harrington, E. A. First-in-human pharmacokinetic and pharmacodynamic study of the dual m-TORC 1/2 inhibitor AZD2014. *Clinical cancer research* **2015**, *21* (15), 3412-3419.

(187) Varshney, D.; Spiegel, J.; Zyner, K.; Tannahill, D.; Balasubramanian, S. The regulation and functions of DNA and RNA G-quadruplexes. *Nature reviews Molecular cell biology* **2020**, *21* (8), 459-474.

(188) Rhodes, D.; Lipps, H. J. G-quadruplexes and their regulatory roles in biology. *Nucleic acids research* **2015**, *43* (18), 8627-8637.

(189) Long, W.; Zheng, B.-X.; Li, Y.; Huang, X.-H.; Lin, D.-M.; Chen, C.-C.; Hou, J.-Q.; Ou, T.-M.; Wong, W.-L.; Zhang, K. Rational design of small-molecules to recognize G-quadruplexes of c-MYC promoter and telomere and the evaluation of their in vivo antitumor activity against breast cancer. *Nucleic acids research* **2022**, *50* (4), 1829-1848.

(190) Monsen, R. C.; Chakravarthy, S.; Dean, W. L.; Chaires, J. B.; Trent, J. O. The solution structures of higher-order human telomere G-quadruplex multimers. *Nucleic acids research* **2021**, *49* (3), 1749-1768.

(191) Lee, D. S.; Ghanem, L. R.; Barash, Y. Integrative analysis reveals RNA G-quadruplexes in UTRs are selectively constrained and enriched for functional associations. *Nature communications* **2020**, *11* (1), 1-12.

(192) Weldon, C.; Dacanay, J. G.; Gokhale, V.; Boddupally, P. V. L.; Behm-Ansmant, I.; Burley, G. A.; Branlant, C.; Hurley, L. H.; Dominguez, C.; Eperon, I. C. Specific G-quadruplex ligands modulate the alternative splicing of Bcl-X. *Nucleic acids research* **2018**, *46* (2), 886-896.

(193) Eddy, J.; Maizels, N. Conserved elements with potential to form polymorphic G-quadruplex structures in the first intron of human genes. *Nucleic acids research* **2008**, *36* (4), 1321-1333.

(194) Robinson, J.; Raguseo, F.; Nuccio, S. P.; Liano, D.; Di Antonio, M. DNA G-quadruplex structures: more than simple roadblocks to transcription? *Nucleic Acids Research* **2021**, *49* (15), 8419-8431.

(195) Hänsel-Hertsch, R.; Di Antonio, M.; Balasubramanian, S. DNA G-quadruplexes in the human genome: detection, functions and therapeutic potential. *Nature reviews Molecular cell biology* **2017**, *18* (5), 279-284.

(196) Collie, G. W.; Promontorio, R.; Hampel, S. M.; Micco, M.; Neidle, S.; Parkinson, G. N. Structural basis for telomeric G-quadruplex targeting by naphthalene diimide ligands. *Journal of the American Chemical Society* **2012**, *134* (5), 2723-2731.

(197) Guilbaud, G.; Murat, P.; Recolin, B.; Campbell, B. C.; Maiter, A.; Sale, J. E.; Balasubramanian, S. Local epigenetic reprogramming induced by G-quadruplex ligands. *Nature chemistry* **2017**, *9* (11), 1110-1117.

(198) Balasubramanian, S.; Neidle, S. G-quadruplex nucleic acids as therapeutic targets. *Current opinion in chemical biology* **2009**, *13* (3), 345-353.

(199) Kikin, O.; D'Antonio, L.; Bagga, P. S. QGRS Mapper: a web-based server for predicting G-quadruplexes in nucleotide sequences. *Nucleic acids research* **2006**, *34* (suppl_2), W676-W682.

(200) Salvati, E.; Zizza, P.; Rizzo, A.; Iachettini, S.; Cingolani, C.; D'Angelo, C.; Porru, M.; Randazzo, A.; Pagano, B.; Novellino, E. Evidence for G-quadruplex in the promoter of vegfr-2 and its targeting to inhibit tumor angiogenesis. *Nucleic Acids Research* **2014**, *42* (5), 2945-2957.

(201) Biswas, B.; Kumari, P.; Vivekanandan, P. Pac1 signals of human herpesviruses contain a highly conserved G-quadruplex motif. *ACS infectious diseases* **2018**, *4* (5), 744-751.

(202) Zhang, Z.; Dai, J.; Veliath, E.; Jones, R. A.; Yang, D. Structure of a two-G-tetrad intramolecular G-quadruplex formed by a variant human telomeric sequence in K+ solution: insights into the interconversion of human telomeric G-quadruplex structures. *Nucleic acids research* **2010**, *38* (3), 1009-1021.

(203) Hatzakis, E.; Okamoto, K.; Yang, D. Thermodynamic stability and folding kinetics of the major G-quadruplex and its loop isomers formed in the nuclease hypersensitive element in the human c-Myc promoter: effect of loops and flanking segments on the stability of parallel-stranded intramolecular G-quadruplexes. *Biochemistry* **2010**, *49* (43), 9152-9160.

(204) Dai, J.; Chen, D.; Jones, R. A.; Hurley, L. H.; Yang, D. NMR solution structure of the major G-quadruplex structure formed in the human BCL2 promoter region. *Nucleic acids research* **2006**, *34* (18), 5133-5144.

(205) Dai, J.; Dexheimer, T. S.; Chen, D.; Carver, M.; Ambrus, A.; Jones, R. A.; Yang, D. An intramolecular G-quadruplex structure with mixed parallel/antiparallel G-strands formed in the human BCL-2 promoter region in solution. *Journal of the American Chemical Society* **2006**, *128* (4), 1096-1098.

(206) Yan, T.; Zhao, B.; Wu, Q.; Wang, W.; Shi, J.; Li, D.; Stovall, D. B.; Sui, G. Characterization of G-quadruplex formation in the ARID1A promoter. *International Journal of Biological Macromolecules* **2020**, *147*, 750-761.

(207) Neidle, S.; Balasubramanian, S. RSC biomolecular sciences. *RSC Biomolecular Sciences* **2006**, 1-252.

List of Publications

1) **Fatma, K**.; Thumpati, P.;Panda , D.;Ravichandiran, V.; Dash, J. Selective Recognition of *c-KIT* 1 G-quadruplex by Structural Tuning of Heteroaromatic Scaffolds and Side Chains. **(Manuscript under revision)**
2) **Fatma, K**.; Pal, S.; Burkhat, I.; Schwalbe, H; Dash, J. Hybrid G4 topology selection of h-TELO by click derived indoloquinolxaline small molecules trigger DNA damage and apoptosis. **(Manuscript under submission)**
3) **Fatma, K**.; Saha, P.; Dash, J. Evidence of Putative G-quadruplex formation in the 5' UTR of mTOR mRNA leads to high tunability of mTOR gene expression and downstream functions. **(Manuscript prepared and to be communicated shortly)**
4) Pal, S.;† **Fatma, K**.;† Ravichandiran, V.; Dash, J. Triazolyl Dibenzo [a, c] phenazines Stabilize Telomeric G-quadruplex and Inhibit Telomerase. *Asian Journal of Organic Chemistry* **2021**, *10* (11), 2921-2926. **(† Equal first author**).
5) Debnath, M.; **Fatma, K**.; Dash, J. Chemical regulation of DNA i-motifs for nanobiotechnology and therapeutics. *Angewandte Chemie* **2019**, *131* (10), 2968-2983.
6) Chaudhuri, R.; **Fatma, K**.; Dash, J. Regulation of gene expression by targeting DNA secondary structures. *Journal of Chemical Sciences* **2021**, *133*, 1-17.
7) Parvin, R.; **Fatma, K**.; Bera, D.; Dash, J.; Mukherjee, J.; Gachhui, R. The effect of temperature on the activity and stability of the thermostable enzyme caffeine dehydrogenase from Pichia manshurica CD1. *Biologia* **2023**, 1-10.
8) Dash, J.; Saha, P.; **Fatma, K**. Click and combinatorial approaches to quadruplex ligand discovery. In *Annual Reports in Medicinal Chemistry*, Vol. 54; Elsevier, 2020; pp 287-324.

ACES Asian Chemical Editorial Society

doi.org/10.1002/ajoc.202100468

ASIAN JOURNAL OF ORGANIC CHEMISTRY Communication

Special Collection

Triazolyl Dibenzo[*a,c*]phenazines Stabilize Telomeric G-quadruplex and Inhibit Telomerase

Sarmistha Pal[+],[a, b] Khushnood Fatma[+],[a] Velayutham Ravichandiran,[b] and Jyotirmayee Dash*[a]

Abstract: We herein report the synthesis and biophysical evaluation of triazolyl dibenzo[*a,c*]phenazine derivatives as a novel class of G-quadruplex ligands. The aromatic core facilitates π-π interaction and the flexible, protonable side chains interact with the phosphate backbone of DNA *via* electrostatic interactions. Förster resonance energy transfer (FRET) melting assay and isothermal titration calorimetry (ITC) studies suggest that these ligands show binding preference for the *hTELO* G-quadruplex over G-quadruplexes found in the promoter region of various oncogenes and duplex DNA. The *in vitro* telomeric repeat amplification protocol (Q-TRAP) assay reveals that these ligands reduce telomerase activity in cancer cells.

G-quadruplex (G4) structures are non-canonical higher-order structures formed by guanine-rich (G-rich) DNA sequences.[1,2] These structures consist of G-quartets, formed by the association of four guanine bases through the hydrogen bonding interactions of Hoogsteen type base-pairing.[1b] They are prevalent at crucial positions of the genome, including telomeric ends, promoter regions (e.g., *cMYC*,[3] *BCL 2*,[4] *c-KIT*,[5] *VEGF*,[6] *KRAS*[7]), introns,[8] untranslated regions (UTRs),[9,10] etc.[8,9,11,12] G4s present in the telomeric region in particular, were among the first G4 structures found to be biologically relevant.[13] In normal cells, single-stranded overhang region containing repetitive hexanucleotide sequence of $[TTAGGG]_n$ in telomeres plays crucial role in maintaining chromosomal stability and prevention of end-to-end fusion.[14] The length of the telomere is shortened at each cell division, leading to cellular aging, senescence and apoptosis.[15] Reverse transcriptase telomerase maintains telomere length and this is further protected from degradation by shelterin, a six-subunit protein complex.[16] In 85–90% malignant cells, telomerase activity is significantly upregulated, resulting in telomere lengthening and immortality of the cell.[17,18]

Telomerase plays a key role in tumorigenesis by lengthening the G-rich repeats at the telomeric ends and inducing cellular immortalization. The 3′ end of telomeric G-rich sequence can adopt monomeric or multimeric G4 structures. Small molecules can stabilize G4 and indirectly inhibit telomerase activity. The inhibition of telomerase is caused due to the lack of 3′ single-stranded telomeric DNA critical for hybridization with RNA subunit of telomerase enzyme for addition of TTAGGG repeats by the telomerase reverse transcriptase protein (Figure 1). Therefore, the telomeric hexanucleotide repeats represent a potential target in anticancer therapy. As a result, the stabilization of telomeric G-quadruplexes by small molecules has emerged as an attractive strategy for the development of anticancer drugs.[19–21]

Small molecule 2,6-diamidoanthraquinone was the first telomeric G4 interacting ligand to exhibit telomerase inhibition.[22] Since then, a number of studies explored the therapeutic potential of telomeric G-quadruplex. Small molecules like telomestatin,[23] BRACO-19,[24] 5,10,15,20-Tetrakis-(N-methyl-4-pyridyl)porphine (TMPyP4),[25] fluoro-isoquinolines,[26] 3,11-difluoro-6,8,13-trimethyl-8*H*-quino[4,3,2-*kl*]acridinium methosulfate (RHPS4),[27] pyridostatin,[28] berberine,[29] phenanthroline-dicarboxamide (PhenDC3),[30] bisquinolinium pyridodicarboxamide (PDC-360A),[31] Epiberberine (EPI),[32] naphthalene diimides (NDI),[33] oxazole telomestatin derivatives (OTD),[34] phthalocyanine,[35] BMPQ-1,[36] and a few more[37,38] are known for selective targeting of telomeric G4. We herein report triazole substituted dibenzo[*a,c*]phenazine derivatives (**DPa–c**) as *hTELO* quadruplex ligands with the ability to inhibit telomerase activity.

[a] *S. Pal,[+] K. Fatma,[+] Prof. Dr. J. Dash*
School of Chemical Sciences
Indian Association for the Cultivation of Science
Jadavpur, Kolkata-700032 (India)
Tel: +91-33-2473-4971, ext.1405
E-mail: ocjd@iacs.res.in

[b] *S. Pal,[+] Prof. V. Ravichandiran*
Department of Medicinal Chemistry
NIPER-KOLKATA
Chunilal Bhawan (Adjacent to BCPL),
168 Maniktala Main Road, P.O. Bengal Chemicals, P.S. Phoolbagan
Kolkata-700054, West Bengal (India)

[+] *These authors contributed equally.*

Supporting information for this article is available on the WWW under https://doi.org/10.1002/ajoc.202100468

This manuscript is part of a special collection celebrating the 10th Anniversary of the Asian Journal of Organic Chemistry.

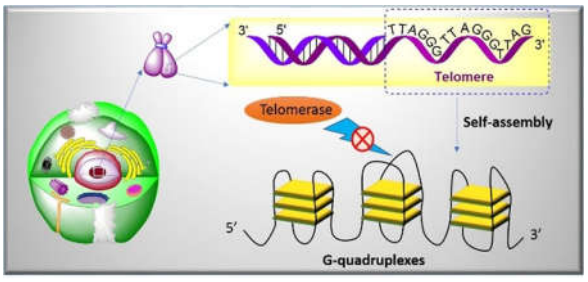


Figure 1. Inhibition of telomerase activity by G-quadruplex formation in telomeric region.

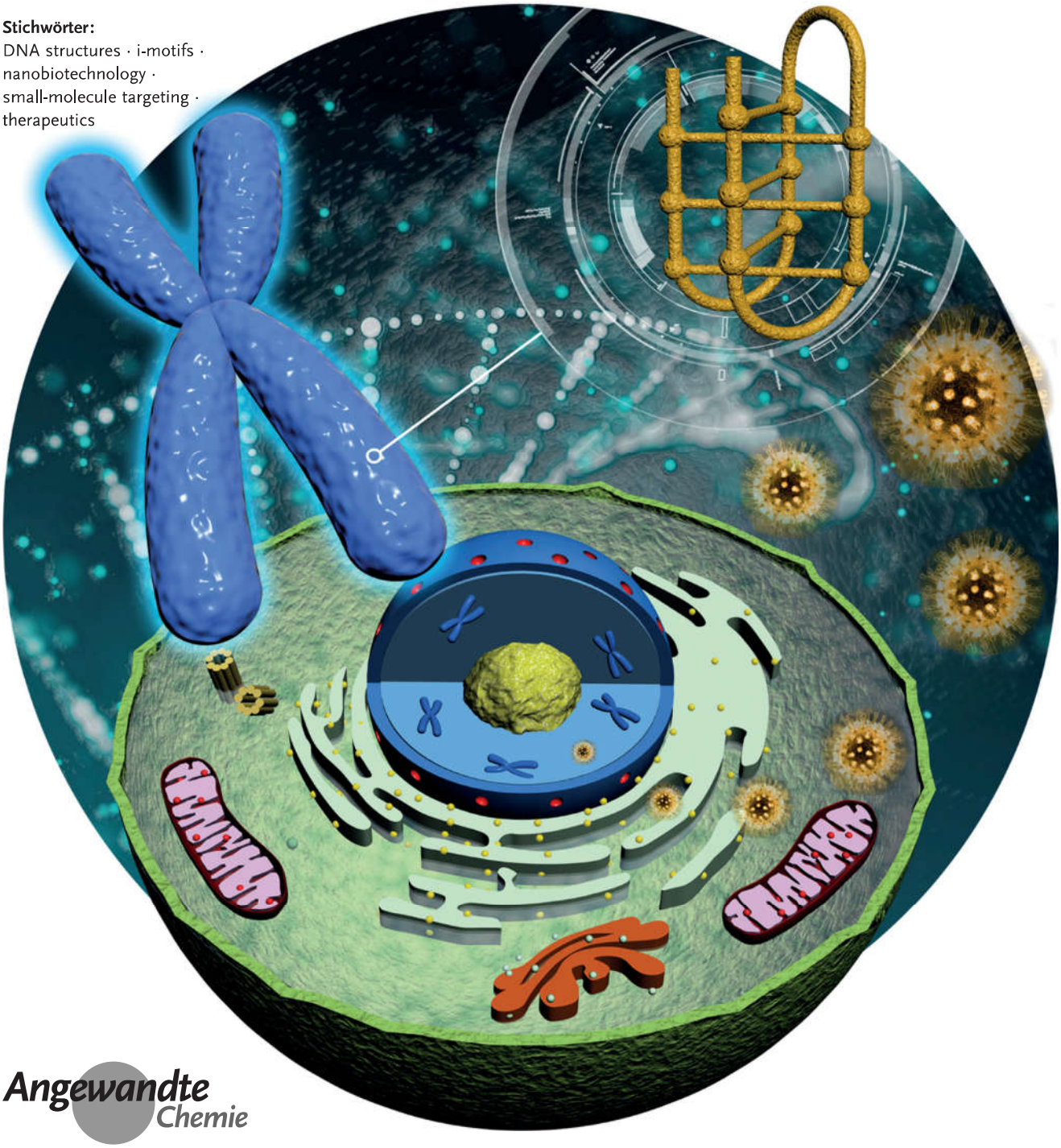
DNA Nanotechnology

International Edition: DOI: 10.1002/anie.201813288
German Edition: DOI: 10.1002/ange.201813288

Chemical Regulation of DNA i-Motifs for Nanobiotechnology and Therapeutics

*Manish Debnath, Khushnood Fatma und Jyotirmayee Dash**

Stichwörter:
DNA structures · i-motifs · nanobiotechnology · small-molecule targeting · therapeutics



J. Chem. Sci. (2021) 133:32
https://doi.org/10.1007/s12039-021-01898-2

PERSPECTIVE ARTICLE

Regulation of gene expression by targeting DNA secondary structures

RITAPA CHAUDHURI, KHUSHNOOD FATMA and JYOTIRMAYEE DASH*
School of Chemical Sciences, Indian Association for the Cultivation of Science, Jadavpur, Kolkata 700032, India
E-mail: ocjd@iacs.res.in

MS received 24 December 2020; revised 17 February 2021; accepted 25 February 2021

Abstract. G-quadruplexes (G4s) are guanine rich and i-motifs are cytosine rich non-canonical four-stranded DNA secondary structures present throughout the genome. Their presence in important regulatory regions of the genome marks their significance in biology. This perspective discusses the importance of both G-quadruplex and i-motif structures and highlights their discovery, structural uniqueness, prevalence, and role in the biological system. Herein, we discuss the ligands developed in our laboratory that target G-quadruplexes and i-motifs and modulate gene expression in cancer cells. We also shed light on some recent techniques developed by us to simplify the synthesis and screening of G4 and i-motif specific ligands and their importance in anti-cancer therapeutics.

Keywords. Anti-cancer therapeutics; Dynamic combinatorial chemistry; G-quadruplex; I-motif; *In situ* click chemistry.

1. Introduction

DNA is considered as the inheritable threads of life as it stores and transfers genetic information from one living organism to another. Usually, it adopts a right-handed double-helical structure stabilized by Watson Crick hydrogen bonding.[1] However, nucleic acids are highly polymorphic and under certain physiological conditions, DNA takes up various non-canonical structures like G-quadruplexes,[2,3] i-motifs,[4] hairpin loops,[5] triplexes,[6] etc. In this perspective, we have focused on G-quadruplexes and i-motifs and describe their structural features, function, and role in gene regulation. Both G-quadruplexes and i-motifs share a similarity in structure as both are non-canonical four-stranded nucleic acid secondary structures. G-quadruplexes and i-motifs are prevalent in the telomeric region as well as in the promoter region of various oncogenes like *c-MYC, c-KIT, BCL2, KRAS, VEGF*, etc. The nearly globular structure of G-quadruplexes formed in oncogene promoters stalls the RNA polymerase and prevents oncogene transcription. Thus, the stabilization of G-quadruplexes by small-molecule ligands is being considered an effective and alternative strategy to regulate oncogene expression. On the other hand, i-motifs have been implicated to regulate gene expression in a manner that could either upregulate or downregulate transcription (Figure 1). Apart from the crucial role of G-quadruplexes in gene regulation, there are exciting examples of utilizing G-quadruplexes and i-motifs in nanotechnology. Their structural oddities and tunable properties make them attractive candidates for the fabrication of synthetic ion channels, biosensors, biomolecular devices, etc.

2. Structural features of G-quadruplexes

G-quadruplexes are four-stranded DNA secondary structures, consisting of stacked planar G-tetrads, with either intra or intermolecular association. The four guanines of each tetrad are connected by cyclic Hoogsteen hydrogen bonding, stabilized by monovalent cations (K^+, Na^+, Figure 2). Sequences that can form G-quadruplexes are found throughout the human genome (376,000 sequences),[7] prevalently in the telomeres, and gene promoters (*c-MYC*[8], *VEGF*,[9] *BCL2*[10], *KRAS*,[11] and *c-KIT*[12]) and exhibit high kinetic and thermal stability. G-quadruplexes can be distinguished from each other by their differences in

*For correspondence

Published online: 02 April 2021

Biologia
https://doi.org/10.1007/s11756-023-01473-9

ORIGINAL ARTICLE

The effect of temperature on the activity and stability of the thermostable enzyme caffeine dehydrogenase from *Pichia manshurica* CD1

Rubia Parvin[1] · Khushnood Fatma[2] · Debbethi Bera[3] · Jyotirmayee Dash[2] · Joydeep Mukherjee[4] · Ratan Gachhui[1]

Received: 27 March 2023 / Accepted: 27 June 2023

Abstract

A novel thermostable caffeine dehydrogenase enzyme was isolated from *Pichia manshurica* strain CD1. At increased temperatures, the enzyme caffeine dehydrogenase exhibits increased activity. The enzyme also exhibits high stability after the heat treatment and remains up to 56% stable at 100 ºC compared to its untreated condition. A Comprehensive analysis was carried out to monitor the thermodynamic parameters (E_a, ΔG, ΔH, *and* ΔS) of the enzyme. In this study, the conformational changes of caffeine dehydrogenase in native and thermally treated states were investigated by CD spectroscopy and intrinsic tryptophan fluorescence spectroscopy. The far-UV CD spectra results showed that the protein was able to maintain its alpha-helical structure after the heat treatment and the near-UV CD spectra results showed that temperature had no such effect on the protein's tertiary structure. The results of tryptophan fluorescence spectroscopy of the enzyme showed a temperature-dependent dynamic quenching. The fluorescence intensity of the protein was decreased as the degree of freedom in the tryptophan- region of the protein was increased with temperature. Thus, the enzyme does not undergo denaturation at high temperatures and can maintain its conformation to remain catalytically active. Therefore, it can be concluded that the thermal stability of caffeine dehydrogenase is an essential key factor in increasing commercial interest due to its high catalytic efficiency and specificity.

Keywords *Pichia manshurica* CD1 · Caffeine dehydrogenase · Thermostability · Secondary structure · Fluorescence spectroscopy

Introduction

Caffeine dehydrogenase (E.C.1.17.5.2) is an intracellular enzyme that is also known as caffeine oxidase in the scientific literature. The enzyme is a member of the oxidoreductase family(Yu et al. 2008). The enzyme is most well-known for its ability to directly oxidize caffeine (Mohapatra et al. 2006). The enzyme is caffeine-specific, has less affinity for theobromine, and has no affinity for xanthine (Mohanty et al. 2012). Caffeine enzymatic catabolism in microbes has been reported to be performed by *N*-demethylation and oxidation pathways with demethylases and dehydrogenase or

✉ Ratan Gachhui
ratangachhui@yahoo.com

Rubia Parvin
parvinrubia@gmail.com

Khushnood Fatma
ockf@iacs.res.in

Debbethi Bera
beradebitthi@gmail.com

Jyotirmayee Dash
ocjd@iacs.res.in

Joydeep Mukherjee
joydeep.mukherjee@jadavpuruniversity.in

1 Department of Life Science & Biotechnology, Jadavpur University, Kolkata, West Bengal 700032, India

2 School of Chemical Sciences, Indian Association for the Cultivation of Science, Kolkata, West Bengal 700032, India

3 Department of Physics, Jadavpur University, Kolkata, West Bengal 700032, India

4 School of Environmental Studies, Jadavpur University, Kolkata, West Bengal 700032, India

Published online: 05 July 2023

Click and combinatorial approaches to quadruplex ligand discovery

Jyotirmayee Dash*, Puja Saha, Khushnood Fatma
School of Chemical Sciences, Indian Association for the Cultivation of Science, Jadavpur, Kolkata, India
*Corresponding author: e-mail address: ocjd@iacs.res.in

Contents

1. Introduction

"Finding a needle in the haystack" is the main aim of any ligand that is supposed to find and interact selectively with its sole target, especially in the cellular context where all biomolecules reside together and perform specific tasks to keep the system alive. The goal behind the design and synthesis of a ligand is its potential to bind with the target and to modulate its biological function for therapeutic purposes. In this regard, non-canonical four

Annual Reports in Medicinal Chemistry
ISSN 0065-7743
https://doi.org/10.1016/bs.armc.2020.04.003