SEQUENCE-SPECIFIC AND CONFORMATION-SPECIFIC TARGETING OF DUPLEX AND QUADRUPLEX DNA GROOVES WITH SMALL MOLECULES

by

RUPESH KUMAR NANJUNDA

Under the Direction of Dr. W. David Wilson

ABSTRACT

Small molecule mediated chemical intervention of biological processes using nucleic acid targets has proven extremely successful and is continually providing exciting new avenues for the development of anti-cancer agents and molecular probes. Correlation between structure and potential biological functions have been explored for DNA conformations that deviate from the standard B-form duplex–including triplexes, quadruplexes, i-motifs and so on. G-quadruplexes has certainly garnered much recognition due to an overwhelming rise in the evidences supporting their involvement in diverse biological process, and potential opportunities exist for regulation of biochemical processes through quadruplex-dependent mechanisms.

Small molecules that have been shown to induce and stabilize telomeric quadruplex conformations are found to have telomerase inhibition activity. The grooves of the quadruplexes offer an alternate recognition site for ligand interactions with potentially higher selectivity than the traditional terminal stacking sites. DB832, a bifuryl-phenyl diamidine, was recently reported to selectively recognize human telomeric G-quadruplex, as a stacked species,

with significant selectivity over duplex sequences. A series of biophysical studies were conducted to test the groove-binding mode of DB832, along with the selectivity for diverse quadruplex forming sequences. To gain better understanding of quadruplex groove-recognition by DB832, a series of structurally similar heterocyclic diamidines were also evaluated. The unique binding mode of DB832 as well as its selectivity for its DNA target may allow it to serve as a paradigm for the design of new class of highly selective quadruplex groove-binding molecules.

Beyond the alternative secondary structures, it is also becoming increasingly apparent that the structure and dynamics of the canonical Watson–Crick DNA double helix play pivotal roles in diverse biological functions. In recent years, compounds preferentially binding to duplex GC sequences have attracted the scientific community to further understand the DNA recognition rules. DB1878, a phenyl-furan-indole diamidine, was shown to recognize a mixed GC/AT motif as a stacked antiparallel dimer. Here, we have conducted detailed NMR structural studies of the complex of DB1878 with a GC/AT sequence. The structure reported here is completely different from the traditional DNA recognition exhibited by lexitropsins, and represents an entirely new motif for DNA minor groove recognition.

INDEX WORDS: DNA, Quadruplex, Duplex, Telomeres, Oncogenes, Telomerase, Diamidines, Groove binding, Selectivity, Specificity, Inhibition, Circular dichroism, Nuclear magnetic resonance, Surface plasmon resonance, Isothermal calorimetry, Thermal melting, Dimer, GC recognition

SEQUENCE-SPECIFIC AND CONFORMATION-SPECIFIC TARGETING OF DUPLEX AND QUADRUPLEX DNA GROOVES WITH SMALL MOLECULES

by

RUPESH KUMAR NANJUNDA

A Dissertation Submitted in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

in the College of Arts and Sciences

Georgia State University

2010

SEQUENCE-SPECIFIC AND CONFORMATION-SPECIFIC TARGETING OF DUPLEX AND QUADRUPLEX DNA GROOVES WITH SMALL MOLECULES

by

RUPESH KUMAR NANJUNDA

Committee Chair: Dr. W. David Wilson

Committee: Dr. Markus Germann

Dr. David W. Boykin

Electronic Version Approved:

Office of Graduate Studies

College of Arts and Sciences

Georgia State University

December 2010

न चोरहार्यं न च राजहार्यं न भ्रातृभाज्यं न च भारकारि।
व्यये कृते वर्धत एव नित्यं विद्याधनं सर्वधनप्रधानम्।।

It cannot be stolen by thieves, nor can it be taken away by kings. It cannot be divided among brothers; it does not cause a load on your shoulders.

If spent... it indeed always keeps growing. The wealth of Knowledge...
is the most superior wealth of all!

-Bhamini (from Sabhataranginii)

…. Dedicated to my dearest Maa and Paa

The greatest parents one can ever have!!!

ACKNOWLEDGEMENTS

It is absolutely impossible to list all the people who have influenced me throughout my doctoral career, nevertheless, first and foremost, I would like to express my deep sense of gratitude in acknowledging Dr. David Wilson for the excellent guidance, encouragement and affection he has showered on me during the course of my work. I have been highly inspired by his keen interest and ever-increasing enthusiasm in research. I consider myself lucky to have such a long time association with him both as a scientific advisor and as a caring friend, and, most importantly, for making me feel left out while other students complained about their advisors. Along with science, he has also taught me how to be a great individual; his passion for science is something every student should imbibe. He is the greatest mentor one can ever have!

I am deeply indebted to my committee members Dr. David Boykin and Dr. Markus Germann for their invaluable input in my research project, for their constructive criticism of this manuscript and, most importantly, for considering me worthy enough to receive a PhD degree! Special thanks to Dr. Germann for letting me use his well maintained NMR facility. My current knowledge of NMR is a result of his inspiring teaching techniques and due to the enlightening discussions I had with him. Special thanks to Dr. Boykin for his excellent compounds, without which none of this research would have ever been possible. Also, my deep gratitude for Dr. Sekar Chandrasekaran for helping me to get started on the Varian NMR instruments and for his constant advice on technical issues.

My thanks go out to all the members of Dr. Wilson's group, both past and present, especially to: Dr. Elizabeth White and Dr. Lei Wang, whose doctoral studies with small molecules to target quadruplexes and duplexes, respectively, formed the starting point of my

studies. Special thanks to Dr. Binh Nguyen (“*The Technical Guru”),* Farial Tanious, Dr. Manoj Munde and Dr. Yang Liu for their constant help with all the instruments in the lab and for useful discussions. I would like to thank Dr. Prashanth Athri for being such a great friend and his exceptional advice on all matters. Special thanks to Caterina Musetti – *the prettiest Italian I ever met!!* – for all her help with quadruplex research and showing me that there is life outside lab. *Ti voglio bene!*

The amount of work involved in a doctoral research is impossible without collaboration with other high-quality research groups. Many thanks for Dr. Nick Hud from Georgia Tech and Dr. Stephen Neidle from University of London for providing excellent compounds for quadruplex research.

Life outside the lab has been wonderful all throughout my graduate years and I thank each one of my friends: Ajay, Vijay, Stan, Sooka, Googs, Rishi, Ved, Vamsi, Sai, Ritesh, Jigar and many more for making it a memorable experience. Thank you Fools!

In order to complete a PhD thesis one needs plenty of endurance and patience. While these characteristics do come in handy for the one actually working on the thesis, they are even more important for the loved ones. On the one hand, this means the family, and I am deeply grateful to them for enduring my unwarranted snappy responses during seemingly endless times of stress. I thank them for their love and support as well as their confidence. I thank my brother who has been a great moral support and for taking care of all the affairs at home. My little nieces – *Aishu, Aditi, Divya* and *Teja* – the angels who always made me laugh and forget the rest of the world at the end of the day with their cute smiles, hugs and constant fights!

No word is big enough to thank my *Maa* and *Paa* for all their support and the hardship they faced to fulfill all my ambitions. They put my education as a first priority in their lives, and raised me to set high goals and standards for myself. They taught me to value honesty, courage, and humility above all other virtues. I dedicate this work to them, to honor their love, patience, and support.

On the other hand – even more critical – the partner, who has to bear all moods first-hand, ranging from just plain bad through to foul, all the way to hostile. If there were ever degrees for such achievements and for putting-up with me then *Hetal* certainly deserves a dozen of them. At the end of the day, it was she who had to handle all the piled-up bad moods and stress situations and, at the same time, keep her sanity. I thank her for her true love, her selfless support, and from-bottom-of-the-heart understanding! My deepest love, gratitude, and the biggest *Thank you* for her–the love of my life!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Chapter 6

LIST OF ABBREVIATIONS

1D	–	One-dimensional
2D	–	Two-dimensional
Å	–	Angstroms
A	–	Adenine
Aza	–	Azacyanines
C	–	Cytosine
COSY	–	Correlated Spectroscopy
CD	–	Circular Dichroism
d1	–	Relaxation delay
dA	–	Deoxyadenosine
dT	–	Deoxythymidine
DNA	–	Deoxyribonucleic Acid
DQF	–	Double Quantum Filtered
DSS	–	4,4-dimethyl-4-silapentane-1-sulfonic acid
EDTA	–	Ethylenediaminetetraethanoic acid
G	–	Guanine
G4-FID	–	Quadruplex-Fluorescence Intercalator Displacement
HEPES	–	4-(2-hydroxyethyl)-1-piperazineethanesulfonic Acid
HETCOR	–	Heteronuclear Correlation Spectroscopy
HSQC	–	Heteronuclear Single Quantum Correlation
HTel	–	Human Telomeric DNA
IC50	–	50% Inhibitory Concentration
ITC	–	Isothermal Calorimetry
mdeg	–	millidegrees

MES	–	2-(N-morpholino)ethanesulfonic Acid
ms	–	milliseconds
ND	–	Naphthalene Diimides
nm	–	nanometers
NMR	–	Nuclear Magnetic Resonance
NOE	–	Nuclear Overhauser Effect
NOESY	–	NOE Spectroscopy
ODN	–	Oligodeoxyribonucleotide
Oligo	–	Oligonucleotide
ppm	–	parts per million
RNA	–	Ribonucleic acid
RU	–	Response Units
SPR	–	Surface Plasmon Resonance
ss	–	single stranded
T	–	Thymine
TBA	–	Thrombin Binding Aptamer
TetTel	–	Tetrahymena Telomere
Tm	–	Thermal melting temperature
TO	–	Thiazole Orange
TOCSY	–	Total Correlation Spectroscopy
TRIS	–	Tris Hydroxymethylaminoethane
UV	–	Ultraviolet

1 INTRODUCTION

1.1 PERSPECTIVE

A great mystery of life is how a single-celled entity, beginning from the time of conception to birth to development, dividing and subdividing in an orderly manner, becomes in time a complex, individual organism. Science has, to a great extent, helped us in our never-ending quest for understanding the complex process of development by probing down to the level of atoms and molecules, to decode their fundamental properties how they are related to, contribute, and control diverse biochemical processes within the cells. The last fifty years have completely changed the way biological and medical researchers study and understand life, susceptibility to infectious and inherited diseases, and the molecular mechanisms of metabolic processes. One reason that brought about this understanding lies in the ability to access the information contained in biological macromolecules. Breakthroughs, combined with unanticipated discoveries, at various stages over decades have led to the creation of what is termed as modern-day medicine that has significantly altered the perception of human body. The reductionist's approach – the study of chemistry and physics of life – created an enormous wealth of biochemical and genetic data available for the rational design of drugs and the manipulation of the genome. The scientific advancements over the years have considerably increased our potential to understand and defend against lethal diseases. Science has gone to a great extent to solve the mystery of life, but it still is at an early stage. With the current knowledge, we can at least make an assertion that many of the underlying processes that guide the complex biochemical processes occurring in our body can be explained through science. Nevertheless, the fact remains that the more we decipher the secrets of the human body through science, the more complex it seems to be.

1.2 THE CENTRAL DOGMA

Within the countless number of cells in the human body, there are infinite number of biochemical processes undergoing, at any given time in a well-orchestrated manner, almost all of which involve the basic biomolecules: carbohydrates, lipids, proteins and nucleic acids. The five bases and twenty amino acids, which make the repertoire of nucleic acids and proteins respectively, are very well conserved throughout nature and, essentially, control and guide most of the biochemical processes occurring in the body. The central dogma of molecular biology, which formulates the sequential information from DNA to protein to be deterministic, was first articulated by Francis Crick in 1958[1]. This model summarizes the transfer of the most complex hereditary information in a very simplistic, but elegant way. Nucleic acids in all organisms carry the genetic information in the form of DNA, although some viruses use RNA as their genetic material. By the process of transcription, a stretch of DNA for a gene dictates the sequence for its correlative mRNA, and then the same mRNA serves as a template for the translation of amino acids into a sequence of polypeptides, which form a protein. Thus, the central dogma reveals a predictable chain of information from DNA to protein, whereupon sequential information for any one component offers determinative sequence information for the other components. The transfer of this information in unidirectional (i.e., DNA → RNA → PROTEINS), and each step is tightly regulated to maintain a balance between the individual components, a hallmark of normal cells. Cancer cells and viruses, however, exhibit deregulated signaling of all or any of the components of the central dogma, resulting in their unbridled proliferation. Also, the flow of information in cancer cells is not in accord with the simplest principles set forth by central dogma, reverse-transcriptase[2] and prion proteins[3] being prime examples.

Complex human diseases encompass a spectrum of genetic and environmental attributes that together affect the normal functioning of several molecular and cellular

pathways. They occur commonly in the population and are a major source of disability and death worldwide. The genes that contribute to complex diseases are notoriously difficult to identify, because they typically exert small effects on disease risk; in addition, the magnitude of their effects is likely to be modified by other unrelated genes as well as environmental factors[4]. One of the greatest challenges facing researchers today is to sort out how these contributing factors interact in a way that translates into effective strategies for disease diagnosis, prevention, and therapy. While major inroads have been made in identifying the genetic causes of disorders, little progress has been made in the discovery of common gene variations that predispose to complex diseases. On an exciting note, the genome sequencing projects have provided a step toward a broader and more complete understanding of challenges and opportunities in addressing some of the underlying questions for a successful therapeutic intervention. Further extensive investigation will help discover hitherto unknown biological functions and provide clues to the occurrence and control of many diseases.

1.3 BEYOND THE DOUBLE HELIX

Ever since the discovery of double stranded DNA, various higher order DNA structures have been discovered [5]. The remarkable ability of single strand oligonucleotides to form a multitude of structures is attributed to an excellent array of hydrogen bond donors and acceptors that nature has precisely incorporated among the nucleobases, and also various external factors such as salt, ion concentration, pH and the extent of hydration. The network of hydrogen bond donors and acceptors among the bases enable them to fold into structures that deviate from the classical Watson-Crick base pairing. Various triplexes[6], cruciforms[7], and i-motif[8] architectures have been reported in designed DNA sequences that have shown to exhibit interesting hydrogen-bonding patterns between the nucleotides (Figure 1.1). Multi-stranded DNA structures have assumed great importance in recent years with the realization

that they play important roles in various biological processes such as DNA recombination, replication, translation, gene expression and disease control, to name a few[9]. There is mounting evidence that DNA structural properties beyond the double helix significantly affect its interactions with proteins and play an important role in a number of biological processes[10].

1.4 G-QUADRUPLEXES

Guanosine molecules show a remarkable ability to self-assemble into highly complex patterns [11]. The self-association of guanine-rich DNA motifs under physiologically feasible ion concentrations led to the hypothesis that these structures might have significant biological importance. Guanine-rich DNA sequences, under certain physiological conditions, can form unique structures known as G-quadruplexes that have structural properties very different from the canonical DNA [5a, 12]. Stacking of G-tetrads leads to the formation of G-quadruplexes; complex and highly ordered helical structures. The G-tetrad motif is composed of four guanines arrayed in a square planar configuration and can self-assemble in the presence of cations (Figure 1.2, A). Adjacent guanine bases in a tetrad are hydrogen-bonded on their Watson-Crick and Hoogsteen edges, with their carbonyl groups oriented towards the centre of the quadruplex. Within a tetrad, the N1 imino proton and the C2 amino proton of a guanine residue are hydrogen bonded respectively to the C6 carbonyl oxygen and the N7 of the neighboring guanine residue, and a similar hydrogen bonding pattern continues until all the four guanines in a tetrad are cyclically connected by hydrogen bonds. Overall, eight hydrogen bonds are observed in a tetrad adding significant stability to the tetrads. The quadruplex is further stabilized through π-π stacking interactions of the stacked tetrads as well as by coordination with cations located between or within the tetrads. These structures exhibit extensive structural diversity and polymorphism relative to duplex DNA. In general, structural

polymorphism arises mostly from the nature of the loop, such as variations of strand stoichiometry, strand polarity, glycosidic torsion angle, and the location of the loops that link the guanine strands [5a, 12]. Meanwhile, the solution environment, such as proteins, ligands, or molecular crowding conditions, may also influence the topology of quadruplex [13]. G-quadruplexes can be folded from a single G-rich sequence intramolecularly or by the intermolecular association of dimeric or tetrameric strands. Other self-assembling motifs, like G-ribbons, were also identified in lipophilic guanosine derivatives [14]. Switching between G-ribbon and G-quadruplex structures was observed as a function of the solvent and cations in the solution [15].

1.4.1 Biological Role of G-Quadruplexes

Despite the fact that these unique structures were discovered and have been known for more than half a century, it is only during the past decade or so these G-quadruplex structures have come into the limelight, due to an overwhelming rise in the evidences supporting various hypotheses about the relevant role of G-quadruplexes in wide range of biological processes [16]. Potential quadruplex forming guanine rich sequences are found in biologically important regions of the genome including telomeres [12, 17], transcriptional regulatory regions of genes (such as insulin, VEGF, fragile X mental retardation gene) [8a, 18], immunoglobulin switch regions [19], promoter regions of oncogenes (such as *c-myc, c-kit, bcl-2, ret, k-ras*) [20] thereby raising exciting possibilities of moderating gene expression with the aid of G-quadruplexes. Because of their apparent role in controlling gene expression and their putative involvement in the inhibition of telomerase activity, G-quadruplexes are potential targets in cancer research [21]. Recent bioinformatics studies have shown that human genome contain as many as 376,000 potential quadruplex-forming sequences and an astounding 40% of these gene promoter regions have at least one quadruplex forming motif [22].

The discovery of a number of proteins in various organisms, such as yeast and oxytricha that can bind to telomeric DNA suggests that these quadruplex structures may also exist *in vivo* [23] and, interestingly, recent studies from different labs have provided evidence that these quadruplex structures do form *in vivo*. Autoradiography studies of metaphase chromosomes of normal and cancer cell lines have shown that a quadruplex binding ligand, ^{3}H-360A, accumulates in the nuclei of these cells and preferentially binds to the terminal regions of metaphase chromosomes of both normal and cancer cells[24]. Biochemical assays have shown that the cationic porphyrin, $TmPyP_4$, binds to and stabilizes a particular conformation of G-quadruplex structure in the NHE III_1 of *c-myc* promoter region thereby directly repressing *c-myc* transcriptional activation[25]. The strongest evidence to date for G-quadruplex existence *in vivo* synthetic single chain antibody fragments specifically targeted against a parallel intermolecular G-quadruplex assembled from the ciliated protozoan *Stylonichia lemnae* telomeric sequence[26]. The presence of a strong signal in the macronucleus by one of the antibody fragments confirmed the presence of quadruplex structures *in vivo*. These exciting results demonstrate that quadruplex structures can also form *in vivo* and ligands can be invariably employed to exploit these G-quadruplex structures as "druggable targets"[27].

1.4.2 Telomeres

Telomeres are specialized non-coding structural units capping the ends of chromosomes in eukaryotic cells of parasites and higher organisms [17d, 28]. They are found to play a vital role in maintaining the stability of chromosomes by preventing end-to-end fusion, recombination, nuclear degradation and cellular senescence [17d, 28a, 29]. Telomere length is regarded as a potential biomarker of aging[30]; there is a growing body of evidence indicating that shorter telomeres are associated with various diseases, including cancer, infectious

diseases, psychological stress, and cardiovascular disease[28a, 31]. Telomere dysfunction limits the proliferative capacity of human cells by activation of DNA damage responses, inducing senescence or apoptosis[30]. In humans, telomere shortening occurs in the vast majority of tissues during aging, and telomere shortening is accelerated in chronic diseases that increase the rate of cell turnover[30]. The leading hypothesis for telomere attrition is due to inflammation, exposure to infectious agents, and other types of oxidative stress, which damage telomeres and impair their repair mechanisms[32]. Several lines of evidence support this hypothesis, including observational findings that people exposed to infectious diseases have shorter telomeres.

1.4.3 End-Replication Problem and Telomerase

The semi-conservative nature of DNA replication prohibits DNA polymerase from replicating the 3'-end of the lagging strand of the DNA (termed as the "end replication problem") due to the absence of a RNA primer with a free 3'-OH group. Therefore, after each round of cell division and DNA replication, the extreme 3'-end of the telomeres shorten by about 50-150 base pairs and after about 60-70 cycles of cell division, the telomeres reach a critical length when the cell enters the senescence stage eventually leading to apoptosis and finally results in cell death[17d, 31d]. However, the telomere length is maintained at a constant level in tumor cell lines as well as in germ-line cells [29]. This is due to the activation of an enzyme called telomerase, which is absent for the normal function of most somatic cells that usually have longer telomeres, whereas widely expressed in immortal cells. In fact, telomeres are highly maintained in length in 80-85% of human tumor cells, which divide indefinitely by the action of telomerase[21d]. Telomerase is a ribonucleoprotein reverse-transcriptase enzyme consisting of an eleven bases long RNA template and a catalytic subunit (hTERT) with reverse transcriptase activity. With the aid of accessory proteins, the enzyme extends the 3' ends of the

DNA by consecutively adding the TTAGGG hexanucleotide repeats using the RNA template as a primer[21d]. However, in order to continue the elongation process, it is highly imperative that the primer maintain the single stranded conformation. Formation of any secondary structures, such as G-quadruplexes, impedes the hybridization of template RNA subunit of telomerase onto the primer during the elongation process and consequently resulting in the inhibition of telomerase activity. Therefore, inducing or stabilizing the telomeric quadruplex conformation might be one of the important methods of controlling and inhibiting the activity of telomerase. Moreover, the telomerase holoenzyme itself has potential recognition sites that can be exploited for the development of inhibitors. Antisense oligodeoxynucleotides[33], hammerhead ribozymes[34], peptide nucleic acids[35], chimeric RNA modules[36], reverse transcriptase inhibitors[37] and immunotherapy agents[38] are some of the agents that are shown to potentially recognize different structural and functional units of telomerase holoenzyme, and are actively studied as telomerase inhibitors. Since telomerase is necessary for the immortality of many cancer types, it is thought to be a potentially highly selective and attractive drug target for several anti-tumour strategies. Its action is detected in most primary human tumor specimens and tumour-derived cell lines, such as those of the prostate, breast, colon, lung and liver [21c-e, 31c, 39]. Hence, inactivation of telomerase may play an important role in cancer therapy. As aforementioned, telomerase activity can also be inhibited by stabilization of quadruplex conformation of telomeres and, therefore, small molecules that stabilize these quadruplex structures could act as telomerase inhibitors and can be employed as potential therapeutic agents.

Human telomeric sequences are one of the most extensively studied motifs due to their critical role in maintaining chromosomal integrity. Human telomeres consists of highly conserved tandem repeats of guanine rich hexanucleotide d[TTAGGG] ranging from 5-15 kb, while the extreme 3' terminus of this telomeric sequence is single stranded and composed of

only about 100-200 bases that folds into a "t-loop" structure[40]. The hexanucleotide repeats of d[TTAGGG] can fold into an array of quadruplex topologies *in vitro* under different physiological conditions[5a]. Structural information of telomeric quadruplex DNA under *in vivo* conditions is essential from drug design stand point of view. The first structure of the human telomeric DNA sequence, $AG_3(T_2AG_3)_3$ by NMR in Na^+ has shown that this sequence folds into an intramolecular quadruplex termed as an antiparallel basket structure with a mixture of diagonal and lateral TTA loops[41]. In the presence of K^+ in a crystalline state, an intramolecular parallel, propeller-type G-quadruplex conformation has been reported. Propeller-type G-quadruplexes have the loops running diagonally between the G-strands with the G-strands in a parallel arrangement[42]. Because the structure in K^+ solution is considered to be biologically more relevant, due to the high intracellular K^+ concentration, several attempts have been made to elucidate the folding topology of human telomeric quadruplex in K^+, and several structures were found that are inconsistent with the crystal structure. The equilibrium of G-quadruplex species in K^+ solution can be altered by several additional factors. For example, platinum-based cross-linking studies have shown that the basket-type structure coexists with other quadruplexes in both Na^+ and K^+ solutions[43]. A subsequent ^{125}I-radioprobing study has revealed that a chair-type conformation is the major species in K^+ solution[44]. Recently, sedimentation and fluorescence studies have revealed that the crystal structure of telomeric DNA is unlikely to be the major species in K^+ solution, and various forms are energetically similar [45]. A mixture of chair-type and parallel/antiparallel hybrid structures may coexist for telomeric DNA in K^+ solution. Circular dichroism studies of several modified human telomeric sequences with bromo-guanine substitutions revealed the formation of highly stable hybrid-type conformations, and were later further confirmed by high resolution NMR studies [46]. Recent structural studies by several groups also showed that the human telomeric quadruplex folds into a mixture of "hybrid-type" mixed parallel/antiparallel quadruplex conformations [47]. However, the NMR structures all have flanking sequences that confer additional

stabilization, so they cannot be directly compared with the crystal structure. Recently, it was reported that human telomeric DNA forms parallel-stranded intramolecular G-quadruplexes in K^+ solution under molecular crowding conditions [48]. Moreover, various labs have suggested a compact stacking structure for multimers of hybrid-type and parallel-type G-quadruplexes in human telomeric DNA. The "hybrid-type" structures of telomeric quadruplexes under physiological K^+ conditions maybe the predominant conformation *in vivo*.

1.4.4 Oncogenes

The *c-myc* gene belongs to the Myc gene family and functions as a gene-specific transcription factor through its protein product, c-Myc, for a wide range of human cancers. The c-Myc protein regulates almost 20% of all cellular genes and is also involved in cell cycle regulation, apoptosis, metabolism, cellular differentiation, and cell adhesion[49]. As a result, the aberrant over expression of *c-myc* is associated with a variety of malignant tumors including those of breast, colon, cervix, and myeloid leukemia[50]. In particular, c-Myc has been identified as one of the main activating factors for the human telomerase reverse transcriptase (hTERT) catalytic domain of the telomerase enzyme[51]. The nuclear hypersensitivity element III_1 (NHE III_1) upstream of the P1 promoter of *c-myc* is a G-rich strand containing a 27-base-pair sequence (*Pu27*, Table 1.1), which has the propensity to adopt a G-quadruplex structure. The presence of a quadruplex within this promoter region was initially proposed based on chemical probe studies, gel mobility measurements, and fluorescence resonance energy transfer (FRET) spectroscopy[52]. In later studies, the topological structures of several *c-myc* quadruplex sequences were determined by circular dichroism (CD), NMR, and mutational experiments [16a, 53]. The G-rich region of *c-myc* contains more than four consecutive G-strands, resulting in the formation of a dynamic mixture of four parallel G-quadruplex loop isomers in the native *Pu27* region. Furthermore, two different sequences

derived from the *Pu27* region have been analyzed by NMR[20d], revealed that both *myc-2345* (Table 1.1) and *myc-1245* (Table 1.1) fold into intramolecular propeller-type G-quadruplexes in K^+ solution. In this case, the core of three G-tetrads is formed by four G-stretches oriented in the same direction, with all the guanines in an *anti* glycosidic conformation and the three loops adopting double-chain-reversal structures, very similar to the crystalline state of telomeric G-quadruplex conformation in K^+. Similar structures have also been found in the *Pu24I* (Table 1.1) and *myc22*-G14T/G23T (Table 1.1) sequences [54]. A interesting structural feature of Pu24I revealed by NMR studies show that a guanine base (G24) at the 3′ end plugs back into the G-tetrad core by participating in G-tetrad formation and displacing another guanine (G10) of a continuous guanine tract in a loop. This configuration is maintained by a stable diagonal loop, which contains a G•(A-G) triad stacking on and capping the G-tetrad core. These new folding features result from the presence of five guanine tracts in the sequence that are different from the four guanine tracts in the *c-myc* sequences studied previously.

bcl-2 is a proto-oncogene, and its oncogenic property arises from decreasing the rate of cell death[55]. Its protein product, Bcl-2, is a mitochondrial membrane protein, which is present in delicate balance with related proteins and is involved in the control of programmed cell death, functioning as an apoptosis inhibitor [56]. Over-expression of *bcl-2* has been found in a wide range of human cancers, including B-cell and T-cell lymphomas, breast, cervical, prostate, and colorectal. In addition, it also functions in chemotherapy-induced apoptosis[55], which indicates its potential role in drug resistance. The human *bcl-2* gene contains a GC-rich region upstream of the P1 promoter, which is critical for the regulation of *bcl-2* gene expression. It can form a mixture of three distinct intramolecular G-quadruplexes (5′G4, MidG4, and 3′G4) resulting from the six runs of guanines, including one run of five guanines, two runs of four guanines each, and three runs of three guanines each (Table 1.1). With more than four consecutive G-tracts in the sequence, the G-quadruplex in *bcl-2* has the ability to

form either three or six different loop isomers [57]. The central G-quadruplex (MidG4, Table 1.1), which is the most stable of the major species formed in the *bcl-2* promoter region, is likely to form a mixed parallel/antiparallel structure consisting of three tetrads connected by loops and to give rise to three possible loop isomers. An NMR study of the shorter and mutated *bcl-2* quadruplex *bcl2*MidG4Pu23-G15T/G16T (Table 1.1) has shown that one of the topologies for this mixed parallel/antiparallel intramolecular quadruplex has two lateral loops and one propeller loop, similar to one of the Tel22 telomeric quadruplex topologies [20c, 58]. The G-rich strand located in the *bcl-2* P1 promoter plays a significant role in the regulation of *bcl-2* transcription [59]. Although the effects of G-quadruplex ligands on *bcl-2* expression remain to be deciphered, some studies have shown that some G-quadruplex ligands can induce apoptosis [60]. In particular, the ligand 12459 has been found to induce apoptosis characterized by dysfunction of Bcl-2 [61]. These findings suggest the possible role of G-quadruplex formation in the *bcl-2* promoter during apoptosis.

Recently, two G-rich sequences (*c-kit* native and *c-kit21,* Table 1.1) in the promoter region of the human *c-kit* gene have been identified, and biophysical studies have shown that these sequences can form G-quadruplexes [20e, 62]. In the *c-kit* native sequence, 87 base pairs upstream of the transcription start site of the human *c-kit* gene, a single G-quadruplex structure forms in K^+ solution. An NMR study has shown that the *c-kit87up* (Table 1.1) sequence forms a new intramolecular G-quadruplex[63]. Most strikingly, an isolated guanine (G10) is involved in G-tetrad core formation, despite the presence of four three-guanine tracts. There are four distinctive loops, including two single-residue and double-chain-reversal loops (A5, C9), a two-residue loop (C11, T12), and a five-residue stem-loop (A16, G17, G18, A19, G20). In view of the importance of predicting G-quadruplex topologies from sequence information, these new folds in which G residues in non-G-tract regions can participate in structural core formation are particularly worthy of attention. In the case of the *c-kit21*

sequence, a variety of quadruplex conformations have been identified. This sequence needs to be mutated in order to form a single quadruplex species, probably with a parallel fold.

1.4.5 Other G-Quadruplex Families

Guanine-rich sequences of human vascular endothelial growth factor (VEGF) and the *neuroblastoma (Rb)* oncogene promoter have also been shown to form G-quadruplexes by chemical footprinting and CD studies [16c, 64]. The G-rich sequences in these regions are known to exist in duplex form and ligands such as TMPyP4 and Telomestatin have been shown to induce parallel quadruplex conformations. Putative quadruplex forming sequences have also been discovered in hypoxia-inducible factor 1α (HIF-1α) promoter region, *k-ras* and *ret* oncogenes [65]. Other types of G-quadruplexes have also been found in the RNA region[66], peptide nucleic acids (PNA)[67] and locked nucleic acids (LNA)[68].

There are now reasons to believe that G-quadruplex structures are not merely an *in vitro* artifact. Both *in vitro* and *in vivo* data strongly support the physiological relevance of this nucleic acid structure at the telomere, one of the signature guanine-rich regions of the genome. While the precise role of this structure at chromosomal ends remains a matter of conjecture, a number of scenarios have been put forward, taking into account the diverse conformations of the G- quadruplex. The inability of telomerase to utilize the G-quadruplex fold of telomere has led to the emergence of a novel avenue for cancer therapy, through G-quadruplex stabilizing agents. These agents can trap DNA in a quadruplex conformation and, in this manner, may either inhibit telomere extension by telomerase or perturb telomere-capping mechanisms. This approach may be complicated by the structural heterogeneity of G-quadruplexes that arise from human telomeric DNA. In addition, questions remain regarding the *in vivo* mechanism of action of existing G-quadruplex stabilizers. Nevertheless, given the

clinical significance of the G-quadruplex, research activities on telomeric and genomic quadruplexes will continue to grow. With a better understanding of the biological functions and structural properties of G-quadruplexes, it is expected that a wealth of new drugs that are less cytotoxic and that have higher selectivity will emerge in the near future.

1.5 TARGETING G-QUADRUPLEXES WITH SMALL MOLECULES

Studying small molecule interactions with quadruplex forming structures is a rapidly developing area in the field of drug design and synthesis driven by the need for understanding at the molecular level for the development of potential candidates into effective chemotherapeutic agents. The basic principles that aid in the design and synthesis of small molecules for duplex DNA recognition can also be employed for quadruplex recognition. Compounds that intercalate between the base pairs of duplex DNA can be used as lead candidates to develop small molecules that can interact with quadruplexes by either intercalation or end-stacking. Several compounds to-date have been shown to bind and stabilize the human telomeric quadruplex conformations. The quadruplex-small molecule interactions, in most cases, occur via the external stacking of the ligand on the G-quartets of either one or both ends of the quadruplex. These compounds are traditionally planar and have aromatic moieties that can efficiently stack on the large aromatic surface area of the G-quartet via π-π interactions. Some of the compounds designed by exploiting the structural properties of G-quartets and have shown to inhibit telomerase activity are di- and tri-substituted acridines such as BSU6039[69] and BRACO-19[60d, 70], cationic porphyrins such as TmPyP4[53, 71] and Se2SAP[72], macrocycles such as PIPER[73], BOQ1[74] and MMQ3[75], carbazoles[76], metal mediated compounds such as Pt-MPQ[77], Ni(II)-salphen[78] and Cu-ttpy[79], and neutral molecules such as telomestatin[16d, 60b, 71b] and HXDV[80]. These compounds have been studied in great detail, and in all cases end stacking is the observed primary mode of binding with

telomeric quadruplex structure, whereas, some compounds such as $TmPyP_4$ exhibit non-specific binding such as external binding to the loops or the backbone of the quadruplex structure.

Quadruplex grooves offer an attractive recognition site for interaction of small molecules, but there has been very little success in designing molecules that target quadruplex grooves. Small molecules that have been discovered to sequence-specifically interact with the grooves of duplex DNA can be used as paradigms to develop compounds to target the grooves of the quadruplex. Spectroscopic studies have shown that DODC, a carbocyanine dye, interacts with the grooves of a dimeric hairpin quadruplex DNA[81]. Recent NMR studies have shown that distamycin, a classic duplex-DNA minor groove binder, also binds to an intermolecular quadruplex conformation as an antiparallel dimer in the opposite grooves making specific hydrogen bonding with the donor and acceptors along the grooves of the quadruplex[82]. ^{1}H-NMR studies have shown that diaryl amides can potentially recognize grooves of a *parallel* quadruplex conformation formed in the promoter region of *c-kit2* oncogene with high degree of selectivity over duplex DNA[83]. However, small molecules that can target grooves of biologically relevant telomeric quadruplex conformation have not been reported so far. Nonetheless, these results point out that quadruplex groove recognition is indeed a viable strategy for drug design and needs much further attention.

1.6 RESEARCH GOALS

DB832, a bifuryl-phenyl diamidine, was recently reported to selectively recognizing human telomeric G-quadruplex conformation with significant selectivity over duplex sequences. Circular dichroism analysis revealed that DB832 can potentially recognize multiple grooves of the telomeric quadruplex as a stacked agent. The first goal of this research described

in Chapter 2 involved a multifaceted approach providing complementary lines of evidence to comprehensively characterize the binding mode of DB832. An array of established spectroscopic techniques, circular dichroism, NMR, thermal melting, fluorescence intercalator displacement (G4-FID) assay, surface plasmon resonance, and isothermal calorimetry techniques have been used to test this. DB832 serves as a paradigm to show that the grooves of the human telomere can indeed be selectively targeted and can serve as a starting point for the design of new molecules that may have therapeutic use as anti-cancer or anti-trypanosomal agents.

With the abundance of various quadruplex forming structures in the human genome, selectivity within different quadruplex conformations becomes an important task. Binding to non-targeted quadruplex sequences results in compound loss and may have unintentional effects on regulation of non-targeted genes. The second goal of this research was to evaluate the selectivity of DB832 to various quadruplex forming sequences in the genome and is presented in Chapter 3. CD studies show that DB832 is highly selective for different human telomeric quadruplex conformations and potentially interacts as a stacked species in the grooves. DB832 binding to non-telomeric sequences such as *c-myc*, *bcl-2*, TBA and intermolecular quadruplex sequences are non-specific in nature, and most likely interact as monomers by stacking at their terminal tetrads. NMR studies with the wild-type and modified telomeric sequences further suggest that DB832 can bind to a selective quadruplex conformation from a mixture of structures.

With the goal of gaining a better understanding of all the structural elements of DB832 contributing towards quadruplex recognition, a series of structurally similar aromatic diamidines, designed and developed by the laboratory of Dr. David Boykin, with DB832 as the prototype, were evaluated. Particular emphasis was directed towards compounds that

maintained the core structure of DB832: the 5-5-6 ring systems. Systematic atom-wise and group-wise modifications were undertaken on the 5-5-6 scaffold of DB832 that were hypothesized to be important for the compound binding. Chapter 4 describes the characterization of these ligands with telomeric quadruplex-DNA using an array of biophysical techniques. These ligands form a novel class of quadruplex-interactive agents with significant selectivity over duplex-DNA, and can be used as lead candidates to design compounds with improved selectivity and affinity.

The goal of the work described in Chapter 5 was the characterization of different classes of quadruplex-interactive agents with telomeric and oncogenic quadruplex forming sequences using surface plasmon resonance. All the ligands exhibited significant selectivity for quadruplex structures over duplex sequences. The synthetic ease of these ligands can facilitate in carrying out more systematic substitutions, coupled with the opportunities to modulate the side-chains, for better quadruplex recognition.

The goal of understanding groove-binding was also probed with a new duplex target compound, DB1878, with a unique GC-containing duplex sequence, and is described in Chapter 6. Detailed NMR studies show that DB1878 interacts as a cooperative antiparallel dimer in the minor groove of –ATGA– containing sequences. However, the stacking of the ligand in the minor groove is unique from all the ligands that have been reported as dimers so far. The presence of guanine is of utmost importance for the dimer formation. Structural details of the dimer complex could potentially aid in design and development of small molecules that can better interact with the grooves of the guanine-rich quadruplex-DNA.

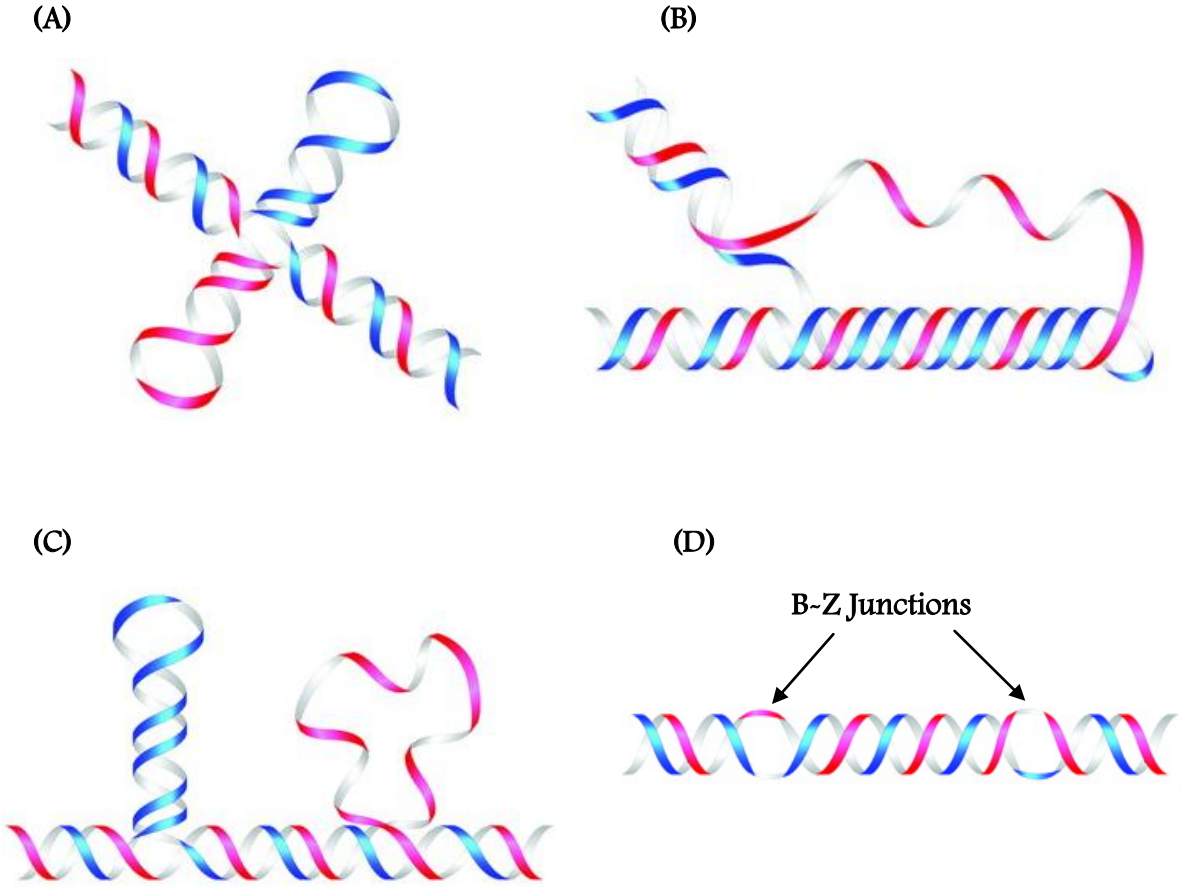
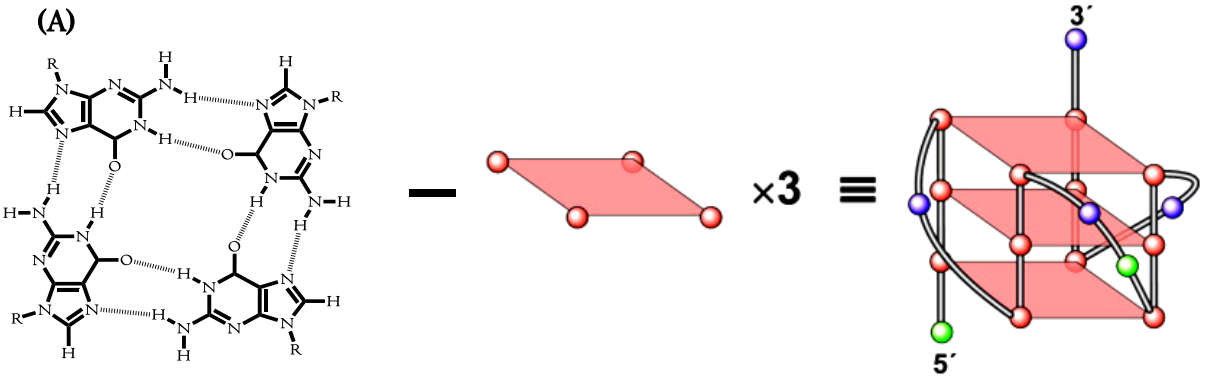


Figure 1.1: Some non-B DNA conformations commonly adopted by DNA. (A) cruciforms, (B) triplexes, (C) slipped structures, and (D) Left-handed Z-DNA.

Cartoon representations directly obtained from Bacolla et al[84], without further permission.



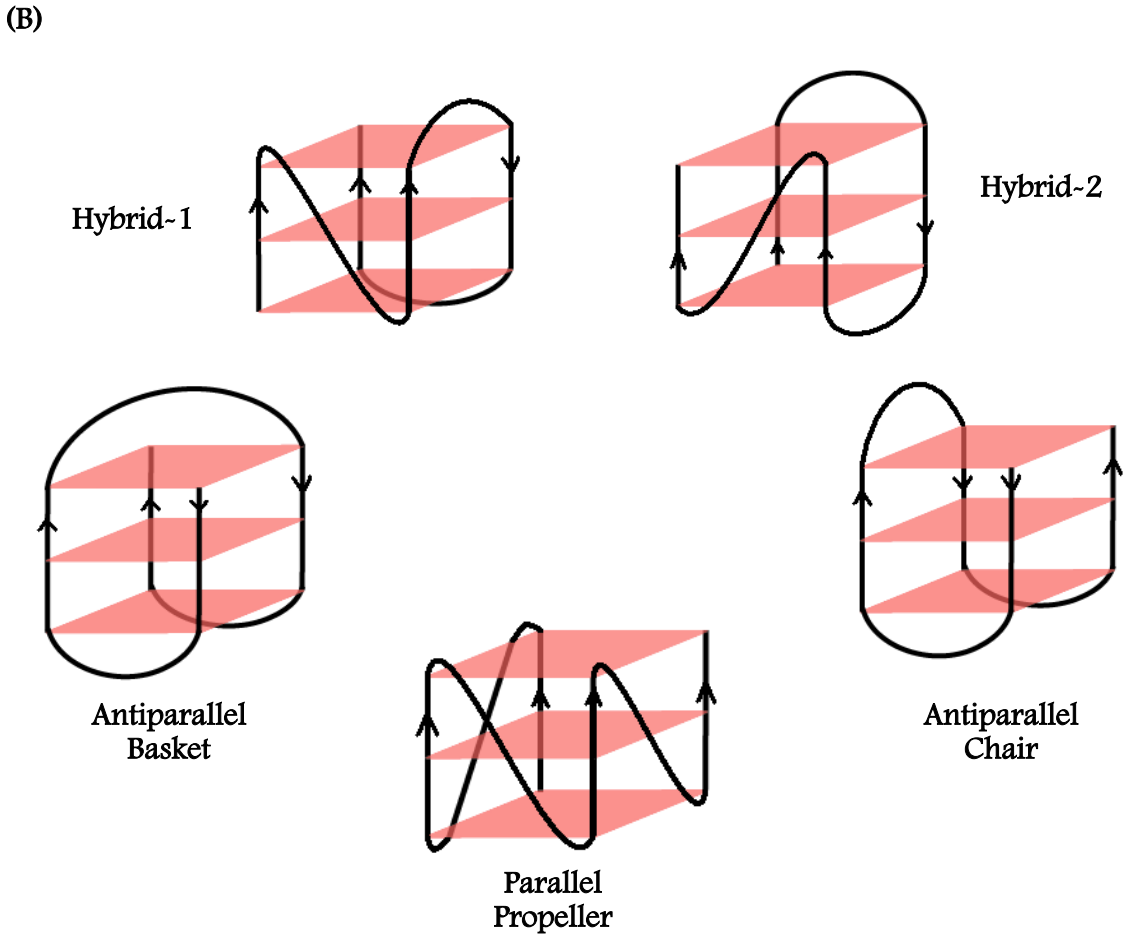


Figure 1.2: Organization of a G-Quadruplex (A), and some commonly observed intramolecular quadruplex folds (B).

Table 1.1: Quadruplex forming sequences most commonly found in the genome.

Region	Variant	Sequence
Human Telomere		TTAGGGTTAGGGTTAGGGTTAGGGTTA
c-myc	*Pu27*	T_1GGGGAGGGT_{10}GGGGAGGGTG$_{20}$GGGAAGG
	myc-2345	T_1GAGGGTGGG$_{10}$GAGGGTGGGG$_{20}$AA
	myc-1245	T_1GGGGAGGGT_{10}TTTTAGGGTG$_{20}$GGGA
	Pu24I	TG$_5$AGGGT_{10}GGIGAGGGTG$_{20}$GGGAAGG
	myc22-G14T/G23T	TGAGGGTGGGT_{14}AGGGTGGGT_{23}AA
bcl-2	*Pu39*	A_1GGGGCGGGC_{10}GCGGGAGGAA_{20}GGGGGCGGGA_{30}GCGGGGCTG
	5'-G4	A_1GGGGCGGGC_{10}GCGGGAGGAA_{20}GGGGGC
	3'-G4	C_{10}GCGGGAGGAA_{20}GGGGGCGGGA_{30}GCGGGGCTG
	MidG4	CGGGC_{10}GCGGGAGGAA_{20}GGGGGCGGGA_{30}GC
	MidG4-G15T/G16T	GGGCGCGGGAGGAA$T_{15}T_{16}$GGGCGGG
c-kit	*c-kit native*	A_1GGGAGGGCG_{10}CTGGGAGGAG_{20}GGGCTG
	c-kit87up	A_1GGGAGGGCG_{10}CTGGGAGGAG_{20}GG
	c-kit21	C_1GGGCGGGCG_{10}CGAGGGGAGG_{20}GGAGGC

1.7 REFERENCES

[1] F. Crick, *Nature* **1970**, *227*, 561.

[2] J. Hurwitz, J. P. Leis, *J. Virol.* **1972**, *9*, 116.

[3] S. B. Prusiner, *Proceedings of the National Academy of Sciences of the United States of America* **1998**, *95*, 13363.

[4] P. Kiberstis, L. Roberts, *Science* **2002**, *296*, 685.

[5] a)A. T. Phan, V. Kuryavyi, D. J. Patel, *Curr Opin Struct Biol* **2006**, *16*, 288; b)V. Bloomfield, D. Crothers, I. Tinoco, *University Science Books* **2000**.

[6] a)M. D. Frank-Kamenetskii, S. M. Mirkin, *Annual Review of Biochemistry* **1995**, *64*, 65; b)R. W. Roberts, D. M. Crothers, *Proceedings of the National Academy of Sciences of the United States of America* **1991**, *88*, 9397.

[7] D. R. F. Leach, *Bioessays* **1994**, *16*, 893.

[8] a)K. Guo, V. Gokhale, L. H. Hurley, D. Sun, *Nucl. Acids Res.* **2008**, *36*, 4598; b)K. Gehring, J. L. Leroy, M. Gueron, *Nature* **1993**, *363*, 561.

[9] a)R. Sundararajan, L. Gellon, R. M. Zunder, C. H. Freudenreich, *Genetics* **2010**, *184*, 65; b)I. Voineagu, C. H. Freudenreich, S. M. Mirkin, *Molecular Carcinogenesis* **2009**, *48*, 309; c)J. L. Callahan, K. J. Andrews, V. A. Zakian, C. H. Freudenreich, *Mol. Cell. Biol.* **2003**, *23*, 7849; d)G. Zheng, T. Kochel, R. W. Hoepfner, S. E. Timmons, R. R. Sinden, *Journal of Molecular Biology* **1991**, *221*, 107; e)A. Dayn, S. Malkhosyan, S. M. Mirkin, *Nucl. Acids Res.* **1992**, *20*, 5991.

[10] a)A. Bacolla, M. Wojciechowska, B. Kosmider, J. E. Larson, R. D. Wells, *DNA Repair* **2006**, *5*, 1161; b)A. Bacolla, R. D. Wells, *Journal of Biological Chemistry* **2004**, *279*, 47411; c)S. C. Raghavan, P. C. Swanson, X. Wu, C.-L. Hsieh, M. R. Lieber, *Nature* **2004**, *428*, 88.

[11] a)J. T. Davis, G. P. Spada, *Chemical Society Reviews* **2007**, *36*; b)S. Lena, P. Neviani, S. Masiero, S. Pieraccini, Gian P. Spada, *Angewandte Chemie International Edition* **2010**, *49*, 3657.

[12] S. Burge, G. N. Parkinson, P. Hazel, A. K. Todd, S. Neidle, *Nucleic Acids Res* **2006**, *34*, 5402.

[13] a)B. I. Kankia, G. Barany, K. Musier-Forsyth, *Nucl. Acids Res.* **2005**, *33*, 4395; b)F. M. Chen, *Biochemistry* **1992**, *31*, 3769; c)B. I. Kankia, L. A. Marky, *Journal of the American Chemical Society* **2001**, *123*, 10799.

[14] E. Mezzina, P. Mariani, R. Itri, S. Masiero, S. Pieraccini, G. P. Spada, F. Spinozzi, J. T. Davis, G. Gottarelli, *Chemistry* **2001**, *7*, 388.

[15] a)A. Calzolari, R. Di Felice, E. Molinari, A. Garbesi, *Physica E: Low-dimensional Systems and Nanostructures* **2002**, *13*, 1236; b)X. Liu, I. C. M. Kwan, S. Wang, G. Wu, *Organic Letters* **2006**, *8*, 3685; c)I. C. M. Kwan, X. Mo, G. Wu, *Journal of the American Chemical Society* **2007**, *129*, 2398.

[16] a)A. Siddiqui-Jain, C. L. Grand, D. J. Bearss, L. H. Hurley, *Proc Natl Acad Sci U S A* **2002**, *99*, 11593; b)K. Shin-ya, K. Wierzba, K.-i. Matsuo, T. Ohtani, Y. Yamada, K. Furihata, Y. Hayakawa, H. Seto, *Journal of the American Chemical Society* **2001**, *123*, 1262; c)D. Sun, K. Guo, J. J. Rusche, L. H. Hurley, *Nucleic Acids Res* **2005**, *33*, 6070; d)M.-Y. Kim, H. Vankayalapati, K. Shin-ya, K. Wierzba, L. H. Hurley, *Journal of the American Chemical Society* **2002**, *124*, 2098.

[17] a)N. Maizels, *Nat Struct Mol Biol* **2006**, *13*, 1055; b)J. D. Griffith, L. Comeau, S. Rosenfield, R. M. Stansel, A. Bianchi, H. Moss, T. de Lange, *Cell* **1999**, *97*, 503; c)A. Meeker, A. De Marzo, *Curr Opin Oncol* **2004**, *16*, 32 ; d)D. Rhodes, L. Fairall, T. Simonsson, R. Court, L. Chapman, *EMBO Rep* **2002**, *3*, 1139

[18] a)P. Catasti, X. Chen, R. K. Moyzis, E. M. Bradbury, G. Gupta, *Journal of Molecular Biology* **1996**, *264*, 534; b)A. Lew, W. J. Rutter, G. C. Kennedy, *Proc Natl Acad Sci U S A* **2000**, *97*, 12508; c)J. C. Darnell, K. B. Jensen, P. Jin, V. Brown, S. T. Warren, R. B. Darnell, *Cell* **2001**, *107*, 489; d)M. Fry, L. A. Loeb, *Proceedings of the National Academy of Sciences of the United States of America* **1994**, *91*, 4950.

[19] C. Schaffitzel, J. Postberg, K. Paeschke, H. J. Lipps, **2010**, pp. 159.

[20] a)S. Cogoi, L. E. Xodo, *Nucl. Acids Res.* **2006**, *34*, 2536; b)H. Fernando, A. P. Reszka, J. Huppert, S. Ladame, S. Rankin, A. R. Venkitaraman, S. Neidle, S. Balasubramanian, *Biochemistry* **2006**, *45*, 7854; c)J. Dai, T. S. Dexheimer, D. Chen, M. Carver, A. Ambrus, R. A. Jones, D. Yang, *J Am Chem Soc* **2006**, *128*, 1096; d)A. T. Phan, Y. S. Modi, D. J. Patel, *J Am Chem Soc* **2004**, *126*, 8710; e)A. K. Todd, S. M. Haider, G. N. Parkinson, S. Neidle, *Nucleic Acids Res* **2007**, *35*, 5799; f)K. Guo, A. Pourpak, K. Beetz-Rogers, V. Gokhale, D. Sun, L. H. Hurley, *Journal of the American Chemical Society* **2007**, *129*, 10220.

[21] a)J. L. Mergny, C. Helene, *Nat Med* **1998**, *4*, 1366; b)Y. Xu, Y. Suzuki, K. Kaminaga, M. Komiyama, *NUCLEIC ACIDS SYMP SER (OXF)* **2009**, *53*, 63; c)D. Sun, L. H. Hurley, *Methods Enzymol* **2001**, *340*, 573; d)J. W. Shay, Y. Zou, E. Hiyama, W. E. Wright, *Hum. Mol. Genet.* **2001**, *10*, 677; e)G. Saretzki, *Cancer Lett* **2003**, *194*, 209

[22] a)J. L. Huppert, A. Bugaut, S. Kumari, S. Balasubramanian, *Nucl. Acids Res.* **2008**, *36*, 6260; b)J. L. Huppert, *Biochimie* **2008**, *90*, 1140.

[23] H. J. Lipps, D. Rhodes, *Trends in Cell Biology* **2009**, *19*, 414.

[24] C. Granotier, G. Pennarun, L. Riou, F. Hoffschir, L. R. Gauthier, A. De Cian, D. Gomez, E. Mandine, J.-F. Riou, J.-L. Mergny, P. Mailliet, B. Dutrillaux, F. D. Boussin, *Nucl. Acids Res.* **2005**, *33*, 4182.

[25] A. Siddiqui-Jain, C. Grand, D. Bearss, L. Hurley, *Proc Natl Acad Sci U S A* **2002**, *99*, 11593

[26] C. Schaffitzel, I. Berger, J. Postberg, J. Hanes, H. J. Lipps, A. Plückthun, *Proceedings of the National Academy of Sciences of the United States of America* **2001**, *98*, 8572.

[27] S. Neidle, G. N. Parkinson, *Biochimie* **2008**, *90*, 1184.

[28] a)M. Blasco, *Curr Opin Genet Dev* **2003**, *13*, 70 ; b)E. M. Rezler, D. J. Bearss, L. H. Hurley, *Curr Opin Pharmacol* **2002**, *2*, 415.

[29] a)R. McElligott, R. J. Wellinger, *Embo J* **1997**, *16*, 3705; b)E. H. Blackburn, *Nature* **2000**, *408*, 53.

[30] H. Jiang, E. Schiffer, Z. Song, J. Wang, P. Zürbig, K. Thedieck, S. Moes, H. Bantel, N. Saal, J. Jantos, M. Brecht, P. Jenö, M. N. Hall, K. Hager, M. P. Manns, H. Hecker, A. Ganser, K. Döhner, A. Bartke, C. Meissner, H. Mischak, Z. Ju, K. L. Rudolph, *Proceedings of the National Academy of Sciences* **2008**, *105*, 11299.

[31] a)J. M. J. Houben, H. J. J. Moonen, F. J. van Schooten, G. J. Hageman, *Free Radical Biology and Medicine* **2008**, *44*, 235; b)T. Minamino, H. Miyauchi, T. Yoshida, Y. Ishida, H. Yoshida, I. Komuro, *Circulation* **2002**, *105*, 1541; c)S. Zimmermann, U. Martens, *Cellular and Molecular Life Sciences* **2007**, *64*, 906; d)H. Jiang, Z. Ju, K. Rudolph, *Zeitschrift für Gerontologie und Geriatrie* **2007**, *40*, 314.

[32] P. Ilmonen, A. Kotrschal, D. J. Penn, *PLoS ONE* **2008**, *3*, e2143.

[33] a)X.-H. Fu, J.-S. Zhang, N. Zhang, Y.-D. Zhang, *Combination of telomerase antisense oligonucleotides simultaneously targeting hTR and hTERT produces synergism of inhibition of telomerase activity and growth in human colon cancer cell line, Vol. 11*, **2005**; b)S. Wang, L. Lin, Z. Chen, R. Lin, S. Chen, W. Guan, X. Wang, *Chinese Science Bulletin* **2002**, *47*, 993.

[34] Y. Yokoyama, Y. Takahashi, A. Shinohara, Z. Lian, X. Wan, K. Niwa, T. Tamaya, *Cancer Research* **1998**, *58*, 5406.

[35] C. G. S. Masood A Shammas, David R Corey, Robert J Shmookler Reis, *Oncogene* **1999**, *18*, 10.

[36] A. N. Elayadi, A. Demieville, E. V. Wancewicz, B. P. Monia, D. R. Corey, *Nucl. Acids Res.* **2001**, *29*, 1683.

[37] a)C. Strahl, E. Blackburn, *Mol. Cell. Biol.* **1996**, *16*, 53; b)E. Pascolo, C. Wenz, J. Lingner, N. Hauel, H. Priepke, I. Kauffmann, P. Garin-Chesa, W. J. Rettig, K. Damm, A. Schnapp, *Journal of Biological Chemistry* **2002**, *277*, 15566.

[38] H. Chen, Y. Li, T. Tollefsbol, *Molecular Biotechnology* **2009**, *41*, 194.

[39] a)E. Rezler, D. Bearss, L. Hurley, *Curr Opin Pharmacol* **2002**, *2*, 415 ; b)M. Shammas, R. Shmookler Reis, M. Akiyama, H. Koley, D. Chauhan, T. Hideshima, R. Goyal, L. Hurley, K. Anderson, N. Munshi, *Mol Cancer Ther* **2003**, *2*, 825 ; c)G. Saretzki, T. von Zglinicki, *Drugs Today (Barc)* **2003**, *39*, 265

[40] Y. Xu, H. Sato, K.-i. Shinohara, M. Komiyama, H. Sugiyama, *NUCLEIC ACIDS SYMP SER (OXF)* **2007**, *51*, 243.

[41] Y. Wang, D. J. Patel, *Structure* **1993**, *1*, 263.

[42] G. N. Parkinson, M. P. H. Lee, S. Neidle, *Nature* **2002**, *417*, 876.

[43] S. Redon, S. Bombard, M.-A. Elizondo-Riojas, J.-C. Chottard, *Nucl. Acids Res.* **2003**, *31*, 1605.

[44] Y. He, R. D. Neumann, I. G. Panyutin, *Nucl. Acids Res.* **2004**, *32*, 5359.

[45] J. Li, J. J. Correia, L. Wang, J. O. Trent, J. B. Chaires, *Nucl. Acids Res.* **2005**, *33*, 4649.

[46] a)Y. Xu, Y. Noguchi, H. Sugiyama, *Bioorg Med Chem* **2006**, *14*, 5584; b)A. Matsugami, Y. Xu, Y. Noguchi, H. Sugiyama, M. Katahira, *Febs J* **2007**, *274*, 3545.

[47] a)A. T. Phan, K. N. Luu, D. J. Patel, *Nucleic Acids Res* **2006**, *34*, 5715; b)A. Ambrus, D. Chen, J. Dai, T. Bialis, R. A. Jones, D. Yang, *Nucl. Acids Res.* **2006**, *34*, 2723; c)J. Dai, M. Carver, D. Yang, *Biochimie* **2008**, *90*, 1172.

[48] Y. Xue, Z.-y. Kan, Q. Wang, Y. Yao, J. Liu, Y.-h. Hao, Z. Tan, *Journal of the American Chemical Society* **2007**, *129*, 11185.

[49] C. V. Dang, K. A. O'Donnell, K. I. Zeller, T. Nguyen, R. C. Osthus, F. Li, *Seminars in Cancer Biology* **2006**, *16*, 253.

[50] C. V. Dang, L. M. S. Resar, E. Emison, S. Kim, Q. Li, J. E. Prescott, D. Wonsey, K. Zeller, *Exp Cell Res* **1999**, *253*, 63.

[51] K. Padmanabhan, K. P. Padmanabhan, J. D. Ferrara, J. E. Sadler, A. Tulinsky, *Journal of Biological Chemistry* **1993**, *268*, 17651.

[52] a)T. Simonsson, M. Kubista, P. Pecinka, *Nucl. Acids Res.* **1998**, *26*, 1167; b)T. Simonsson, R. Sjöback, *Journal of Biological Chemistry* **1999**, *274*, 17379.

[53] J. Seenisamy, E. M. Rezler, T. J. Powell, D. Tye, V. Gokhale, C. S. Joshi, A. Siddiqui-Jain, L. H. Hurley, *Journal of the American Chemical Society* **2004**, *126*, 8702.

[54] a)A. Ambrus, D. Chen, J. Dai, R. A. Jones, D. Yang, *Biochemistry* **2005**, *44*, 2048; b)A. T. Phan, V. Kuryavyi, H. Y. Gaw, D. J. Patel, *Nat Chem Biol* **2005**, *1*, 167.

[55] J. M. Adams, S. Cory, *Oncogene* **0000**, *26*, 1324.

[56] D. T. Chao, S. J. Korsmeyer, *Annual Review of Immunology* **1998**, *16*, 395.

[57] T. S. Dexheimer, D. Sun, L. H. Hurley, *J Am Chem Soc* **2006**, *128*, 5404.

[58] J. Dai, D. Chen, R. A. Jones, L. H. Hurley, D. Yang, *Nucleic Acids Res* **2006**, *34*, 5133.

[59] J. U. Seto M, Hockett RD, Graninger W, Bennett S, Goldman P, Korsmeyer SJ, *Embo J* **1988**, *7*.

[60] a)G. Pennarun, C. Granotier, L. R. Gauthier, D. Gomez, F. Hoffschir, E. Mandine, J.-F. Riou, J.-L. Mergny, P. Mailliet, F. D. Boussin, **2005**, *24*, 2917; b)T. Tauchi, K. Shin-ya, G. Sashida, M. Sumi, A. Nakajima, T. Shimamoto, J. H. Ohyashiki, K. Ohyashiki, *Oncogene* **2003**, *22*, 5338; c)C. Leonetti, S. Amodei, C. D'Angelo, A. Rizzo, B. Benassi, A. Antonelli, R. Elli, M. F. G. Stevens, M. D'Incalci, G. Zupi, A. Biroccio, *Molecular Pharmacology* **2004**, *66*, 1138; d)A. M. Burger, F. Dai, C. M. Schultes, A. P. Reszka, M. J. Moore, J. A. Double, S. Neidle, *Cancer Research* **2005**, *65*, 1489.

[61] C. Douarre, D. Gomez, H. Morjani, J.-M. Zahm, M.-F. O'Donohue, L. Eddabra, P. Mailliet, J.-F. Riou, C. Trentesaux, *Nucl. Acids Res.* **2005**, *33*, 2192.

[62] a)S. Rankin, A. P. Reszka, J. Huppert, M. Zloh, G. N. Parkinson, A. K. Todd, S. Ladame, S. Balasubramanian, S. Neidle, *J Am Chem Soc* **2005**, *127*, 10584; b)H. Fernando, A. P. Reszka, J. Huppert, S. Ladame, S. Rankin, A. R. Venkitaraman, S. Neidle, S. Balasubramanian, *Biochemistry* **2006**, *45*, 7854.

[63] A. T. Phan, V. Kuryavyi, S. Burge, S. Neidle, D. J. Patel, *J Am Chem Soc* **2007**, *129*, 4386.

[64] Y. Xu, H. Sugiyama, *Nucl. Acids Res.* **2006**, *34*, 949.

[65] a)K. Guo, A. Pourpak, K. Beetz-Rogers, V. Gokhale, D. Sun, L. H. Hurley, *J Am Chem Soc* **2007**, *129*, 10220; b)S. Cogoi, L. E. Xodo, *Nucleic Acids Res* **2006**, *34*, 2536; c)R. De Armond, S. Wood, D. Sun, L. H. Hurley, S. W. Ebbinghaus, *Biochemistry* **2005**, *44*, 16341.

[66] M. Wieland, J. S. Hartig, *Chemistry & Biology* **2007**, *14*, 757.

[67] E. A. Englund, Q. Xu, M. A. Witschi, D. H. Appella, *Journal of the American Chemical Society* **2006**, *128*, 16456.

[68] J. T. Nielsen, K. Arar, M. Petersen, *Nucl. Acids Res.* **2006**, *34*, 2006.

[69] N. H. Campbell, M. Patel, A. B. Tofa, R. Ghosh, G. N. Parkinson, S. Neidle, *Biochemistry* **2009**, *48*, 1675.

[70] J. Debray, W. Zeghida, M. Jourdan, D. Monchaud, M.-L. Dheu-Andries, P. Dumy, M.-P. Teulade-Fichou, M. Demeunynck, *Organic & biomolecular chemistry* **2009**, *7*.

[71] a)N. Nagesh, R. Buscaglia, J. M. Dettler, E. A. Lewis, *Biophysical Journal* **2010**, *98*, 2628; b)M.-Y. Kim, M. Gleason-Guzman, E. Izbicka, D. Nishioka, L. H. Hurley, *Cancer Research* **2003**, *63*, 3247.

[72] J. Seenisamy, S. Bashyam, V. Gokhale, H. Vankayalapati, D. Sun, A. Siddiqui-Jain, N. Streiner, K. Shin-ya, E. White, W. D. Wilson, L. H. Hurley, *Journal of the American Chemical Society* **2005**, *127*, 2944.

[73] a)C. Sissi, L. Lucatello, A. Paul Krapcho, D. J. Maloney, M. B. Boxer, M. V. Camarasa, G. Pezzoni, E. Menta, M. Palumbo, *Bioorganic & Medicinal Chemistry* **2007**, *15*, 555; b)W. M. David, J. Brodbelt, S. M. Kerwin, P. W. Thomas, *Anal Chem* **2002**, *74*, 2029.

[74] C. Allain, D. Monchaud, M.-P. Teulade-Fichou, *Journal of the American Chemical Society* **2006**, *128*, 11890.

[75] a)J.-L. Mergny, L. Lacroix, M.-P. Teulade-Fichou, C. Hounsou, L. Guittat, M. Hoarau, P. B. Arimondo, J.-P. Vigneron, J.-M. Lehn, J.-F. Riou, T. Garestier, C. Hélène, *Proceedings of the National Academy of Sciences of the United States of America* **2001**, *98*, 3062; b)M.-P. Teulade-Fichou, C. Carrasco, L. Guittat, C. Bailly, P. Alberti, J.-L. Mergny, A. David, J.-M. Lehn, W. D. Wilson, *Journal of the American Chemical Society* **2003**, *125*, 4732.

[76] a)C.-C. Chang, J.-Y. Wu, C.-W. Chien, W.-S. Wu, H. Liu, C.-C. Kang, L.-J. Yu, T.-C. Chang, *Analytical Chemistry* **2003**, *75*, 6177; b)F.-C. Huang, C.-C. Chang, P.-J. Lou, I.-C. Kuo, C.-W. Chien, C.-T. Chen, F.-Y. Shieh, T.-C. Chang, J.-J. Lin, *Molecular Cancer Research* **2008**, *6*, 955.

[77] a)R. Kieltyka, P. Englebienne, N. Moitessier, H. Sleiman, **2010**, pp. 223; b)H. Bertrand, S. Bombard, D. Monchaud, M.-P. Teulade-Fichou, *NUCLEIC ACIDS SYMP SER (OXF)* **2008**, *52*, 163; c)I. Ourliac-Garnier, M.-A. Elizondo-Riojas, S. Redon, N. P. Farrell, S. Bombard, *Biochemistry* **2005**, *44*, 10620.

[78] J. E. Reed, A. A. Arnal, S. Neidle, R. Vilar, *Journal of the American Chemical Society* **2006**, *128*, 5992.

[79] Y. L. Jiang, Z. P. Liu, *Mini Reviews in Medicinal Chemistry* **2010**, *10*, 726.

[80] C. M. Barbieri, A. R. Srinivasan, S. G. Rzuczek, J. E. Rice, E. J. LaVoie, D. S. Pilch, *Nucl. Acids Res.* **2007**, *35*, 3272.

[81] Q. Chen, I. D. Kuntz, R. H. Shafer, *Proceedings of the National Academy of Sciences of the United States of America* **1996**, *93*, 2635.

[82] L. Martino, A. Virno, B. Pagano, A. Virgilio, S. Di Micco, A. Galeone, C. Giancola, G. Bifulco, L. Mayol, A. Randazzo, *Journal of the American Chemical Society* **2007**, *129*, 16048.

[83] J. Dash, P. S. Shirude, S.-T. D. Hsu, S. Balasubramanian, *Journal of the American Chemical Society* **2008**, *130*, 15950.

[84] A. Bacolla, R.D. Wells, *Journal of Biological Chemistry* **2004**, *279*, 47411.

2 EVALUATION OF DB832 AS A POTENTIAL GROOVE BINDER FOR HUMAN TELOMERIC QUADRUPLEX DNA AS A STACKED SPECIES

2.1 INTRODUCTION

Contrary to popular belief that nucleic acids, particularly DNA, are relatively static rigid units compared to other biomolecules [1], nucleic acids exhibit remarkable structural plasticity with conformations ranging from simple linear strands to complex scaffolds such as t-RNA, Holliday junctions and compaction in nucleosomes [2]. Such conformational transitions, highly dictated by local sequence composition, can be substantially modulated with salt, pH, degree of hydration and, most importantly, proteins and small-molecules, to influence diverse cellular functions. Guanine-rich sequences have an intrinsic ability to form such conformationally dynamic structures known as G-quadruplexes, comprised of planar arrays of G-quartets connected through a network of Hoogsteen hydrogen bonds [2c]. These G-quartets, coupled with coordination with various metal ions and by combinatorial possibilities of loop residues, can stack into a plethora of highly dynamic G-quadruplex architectures [3].

G-quadruplex motifs have been shown to form in several biologically active regions including telomeres and oncogene promoter regions where there is a significant enrichment of guanine-rich clusters [4]. Furthermore, recent genome-wide studies suggest that as many as 376,000 potential quadruplexes could exist within the human genome and more than 40% of human genes contain a putative G-quadruplex in their promoter regions [5]. Despite the fact these structures were discovered more than half a century ago [6], it is only during the past decade or so they have garnered much recognition due to an overwhelming rise in the evidences supporting their involvement in diverse biological process [7]. G-quadruplex formation and stabilization in the telomeric G-rich single-strand component has been shown to significantly reduce telomerase activity [7b, 7c, 8]; also ligand-mediated G-quadruplex

stabilization in oncogene promoter regions can significantly down-regulate gene transcription levels [9]. Fundamental design strategies that have been successfully formulated to design ligands for duplex recognition can also be extrapolated for quadruplex recognition. In fact, CX3543, a promising candidate currently in phase II clinical trials, was designed from fluoroquinolones that target duplex DNA [10] and exerts its antiproliferative activity through a quadruplex-mediated chemical intervention [11]. It has now become more evident that these G-quadruplex structures have very unique chemical and biological properties, and potential opportunities exist for regulation of biochemical processes through quadruplex-dependent mechanisms.

Human telomeric DNA has certainly generated the greatest interest in the G-quadruplex field because of the natural existence of a potential quadruplex-forming guanine-rich TTAGGG repeats at their 3' ends, and the mixture of G-quadruplex conformations it can fold into [12]. The *single-strand* overhang is an essential substrate for the normal functioning of telomerase enzyme; the formation of a G-quadruplex structure by the overhang essentially disrupts the telomere elongation mechanism of the enzyme [4i, 7b, 8c, 13]. Enhanced expression of telomerase is a key marker for and is essential for the survival of many types of cancer cells [14]. Small molecules that have been shown to induce and stabilize telomeric quadruplex conformations are found to have telomerase inhibition activity [15]. Therefore, stabilizing these unique structures with small molecules is rapidly developing as an important anticancer strategy to successfully design and develop potential lead candidates into effective chemotherapeutic agents. Several compounds to-date have been shown to bind and stabilize G-quadruplex conformations, and the most common recurring theme of quadruplex recognition is via stacking of the ligands at one or both ends of the quadruplex [3c, 16]. Most of these ligands that bind by this mode are highly planar aromatic, macrocyclic scaffolds with very little conformational flexibility [16] and, in some cases, these scaffolds are tethered with

side chains to preferably interact with quadruplex grooves or the loops and enhance their affinity with the target [15a, 15e, 17]. Some of the strongest binding telomeric quadruplex end stackers with highest telomerase inhibition and promising anticancer activity include telomestatin, BRACO-19, and RHPS4 [18]. Essentially all of the known quadruplex-end stackers are designed using duplex DNA intercalators as paradigms and, even though, some of the ligands have significant selectivity for quadruplexes, the majority of them exhibit very little selectivity among different quadruplex DNAs. From the target selectivity perspective, it is important that the drug bind to relatively very few-to none of the non-target sequences to circumvent any deleterious effects and also loss of valuable compound. Therefore, successfully designing ligands that exhibit *significant selectivity* for quadruplex structures and effectively inhibit telomerase is absolutely imperative.

The grooves of the quadruplexes offer an alternate recognition site for ligand interactions with potentially higher selectivity than the traditional terminal stacking sites. Moreover, the structural polymorphism exhibited by G-quadruplexes in the genome has in itself posed significant challenges in designing ligands to distinguish between quadruplexes or to induce interconversion between different conformers [12a, 12d-f, 19]. Nevertheless, by successfully exploiting the distinct structural differences in groove geometries among quadruplex, coupled with diverse hydrogen-bonding possibilities, ligands can be successfully engineered to discriminate quadruplex grooves from the grooves of other quadruplex and of duplex DNA. However, there has been very little success so far in developing small molecules that strongly and selectively target quadruplex grooves. Considering the wealth of quadruplex structural information now available, groove-targeting is clearly an area that requires much more consideration. Some of the ligands that have been reported so far to interact with quadruplex grooves exhibit very little selectivity over duplex DNA and, moreover, the reported groove-binding ligands have only been characterized with fairly simple tetrameric or dimeric

G-quadruplex systems [20]. For instance, distamycin, a well-known DNA minor-groove binder, was reported to bind as an antiparallel stacked dimer in the opposite grooves of truncated telomeric sequence of *Oxytricha*, d(TGGGGT) [20a]. This sequence is reported to fold into a simple *interstrand parallel* conformation with four readily accessible equivalent grooves. Additionally, distamycin molecules have been reported to stack on the terminal G-tetrads of very similar G-quadruplex conformations [20b], and that the distamycin groove-recognition is favored only at substantial ligand concentrations [20a]. Similarly, virtual-screening and ^{1}H-NMR results from the same group have shown a series of compounds from a library interacting with grooves of the aforementioned quadruplex conformation [21]. In a very important finding, cyanine dyes such as DODC were predicted to bind in the grooves of human telomeric quadruplex conformation [22]. Cyanine molecules are, however, well-known for forming stacked aggregates in the grooves of DNA duplexes [23] and it maybe difficult to develop as selective G-quadruplex agents. A recent exciting study has shown that diarylethynyl amides can potentially recognize grooves of a *parallel* quadruplex conformation formed in the promoter region of *c-kit2* oncogene with high degree of selectivity over duplex DNA [24]. The terminal amide and amino groups of these ligands were hypothesized to form favorable interactions with the G-quadruplex grooves and loops.

Small molecules that can strongly and selectively recognize quadruplex grooves in biologically significant regions such as human telomeres have not been discovered or characterized so far. We recently reported on a bifuryl-phenyl diamidine, DB832 (Figure 2.1-C), capable of selectively recognizing human telomeric G-quadruplex conformation with significant selectivity over duplex sequences [25]. Circular dichroism analysis revealed that DB832 can potentially recognize *multiple grooves* of the telomeric quadruplex as a stacked agent. This highly encouraged us to pursue a high-resolution solution structure of DB832 complexed in the grooves of human telomeric quadruplex using multidimensional NMR.

However, the unique quadruplex recognition mode of DB832 coupled with the conformational plasticity of human telomeric quadruplex, and some of the inherent limitations associated with NMR methods, presented major experimental hurdles that essentially constrained our structure elucidation process. As a result, in this study, we formulated a multi-method approach that would provide complementary lines of evidence to comprehensively characterize the groove-binding trait exhibited by DB832. An array of established spectroscopic techniques, circular dichroism, NMR, thermal melting, fluorescence intercalator displacement (G4-FID) assay, surface plasmon resonance, and isothermal calorimetry techniques have been used to test this. In the process, we have also evaluated distamycin and F1190, two small molecules that were recently reported as quadruplex groove-binders for simple quadruplex systems, for their selectivity for human telomeric quadruplex sequences, and have found rather contradicting evidence that proves otherwise. DB832 serves as a paradigm to show that the grooves of the human telomere can indeed be targeted and it can also serve as the starting point for the design of new molecules that may have therapeutic use as anti-cancer or anti-trypanosomal agents. Nevertheless, the dearth of compounds representing the quadruplex groove-binding class distinctly highlights the experimental challenges involved in successfully designing small molecules with significant target selectivity and, therefore, makes it a much more challenging and attractive field to pursue.

2.2 MATERIALS AND METHODS

2.2.1 Oligonucleotides, compounds and buffer

The unlabeled and 5'-biotin labeled oligonucleotides Tel22, d[AGGG(TTAGGG)$_3$]; Tel24, d[TTGGG(TTAGGG)$_3$A]; Tel26, d[AAAGGG(TTAGGG)$_3$AA]; *c-myc*, d[T(AGGGTGGGG)$_2$AA]; U6U7, d[TTAGGGUUAGGG]; d[TGGGGT]; ODN9, d[AG_{Br}GGTTAG_{Br}GGTTAGGGTTAG_{Br}GG]; ODN4, d[AGGGTTAGGGTTA$G_{Br}G_{Br}G_{Br}$TTAGGG];

Dickerson, d[CGAATTCGTTTTCGAATTCG] were purchased from The Midland Certified Reagent Company or Integrated DNA Technologies with HPLC purification and mass spectrometry characterization. NMR analysis performed on these sequences further confirmed the purity of these sequences. The concentration of oligonucleotides was determined from absorbance at 260 nm using molar extinction coefficient obtained from nearest-neighbor principle. F1190 was purchased from Life Chemicals. Thiazole Orange and TMPyP4 were purchased from Mid-Century Chemicals. The synthesis of DB832 will be described elsewhere. A 1 mM stock solution of each compound was prepared in double deionized water and in deuterated water for NMR experiments. This stock solution was diluted to required concentrations with appropriate buffer right before their usage.

2.2.2 Circular Dichroism Studies

All CD experiments were performed at 25 °C in 10 mM K_2HPO_4 buffer containing 80 mM KCl and 1 mM EDTA. CD spectra were recorded using a Jasco J-810 spectrophotometer with a 1-cm pathlength quartz cell at a scan speed of 50 nm/min and response time of 1 second. Appropriate amount of compounds were sequentially titrated from the stock solution into the DNA solution in the cuvette until the desired mole ratios of compound to quadruplex were obtained. The spectra were averaged over four scans. A buffer baseline scan was collected in the same cuvette and subtracted from the average scan of each ratio. Data were processed and plotted using Kaleidagraph 4.0 software.

2.2.3 Thermal Melting Studies

Thermal denaturation studies were conducted on a Cary 300 BIO UV-visible spectrophotometer in quartz cuvettes of 1 cm pathlength. Compound-DNA solutions were prepared in low salt buffer containing 10 mM TRIS buffer (pH 7.4), 10 mM KCl and 1 mM EDTA. The absorbance of the oligonucleotides was monitored at the recommended wavelength of 295 nm for quadruplex sequences and 260 nm for duplex sequences as a function of

temperature. Several duplex and hairpin DNA oligonucleotides without compound were used as a control. Samples of compound to DNA ratios from 0:1 to 4:1 was prepared. Cuvettes were mounted in a thermal block, and the solution temperatures were monitored by a thermistor in a reference cuvette with a computer-controlled heating rate of 0.5 °C/min. Experiments were generally conducted at quadruplex concentrations in the range of 2-3 μM in TRIS buffer containing 10 mM KCl. Data were analyzed and plotted using Kaleidagraph 4.0 software.

2.2.4 Nuclear Magnetic Resonance Studies

Quadruplex DNA samples were prepared in degassed phosphate buffer containing 80 mM KCl, 10 mM K_2HPO_4 and 0.1 mM EDTA and reconstituted in 90% H_2O:10% D_2O. DSS was employed as an internal reference. Intramolecular quadruplex concentrations were in the range of 0.1 mM to 0.3 mM unless otherwise mentioned. The final DNA samples were adjusted to pH 7.0 using 1M HCl or 1M KOH solutions. Finally, the NMR samples were heated past their transition temperature and annealed to room temperature several times before collecting the spectra. Experiments were performed on a Varian Unity 600 spectrometer. DB compounds were titrated into the quadruplex DNA with compound to DNA ratios varying from 1 to 4. Temperature-dependent ^{1}H-spectra were recorded from 15°C to 45°C using jump-return and WATERGATE methods for solvent suppression. All NMR data were processed and analyzed with a combination of VNMR (Varian Inc.) and MNova (Mestrelab Research) software.

2.2.5 Quadruplex-Fluorescence Intercalator Displacement Assay (G4-FID)

All the experiments were performed using a Varian CARY Eclipse Fluorescence Spectrophotometer in a 1 ml cuvette starting from 0.5 μM solution of pre-folded DNA mixed with 1 μM of the fluorophore thiazole orange (an excess is required to be sure that all the binding sites are occupied by the fluorophore) in 10 mM TRIS/50 mM KCl buffer at pH 7.5. The addiction of ligand was followed after 3 minutes of equilibration time and after this the fluorescence spectrum was recorded with excitation wavelength at 501 nm (thiazole orange

maximum absorbance wavelength) and scanned from 520 nm to 700 nm. The 1/30% fluorophore displacement value was then plotted as a function of the concentration of added ligand.

2.2.6 Surface Plasmon Resonance Studies

Biosensor SPR experiments were performed with a four-channel BIAcore 2000 optical biosensor system (BIAcore, Inc.) and streptavidin-coated sensor chips. All DNA samples, for either duplex- or quadruplex-binding experiments, were used as single strands to prevent dissociation in the SPR flow system. The chips were prepared for use by conditioning with a series of 1 min injections of 1 M NaCl in 50 mM NaOH followed by extensive washing with buffer. 5'-Biotinylated DNA samples (25-50 nM) in HBS buffer were immobilized on the flow cell surface by non-covalent capture as previously described [26]. Three flow cells were used to immobilize DNA samples, and any one of the flow cell was left blank as a control. Interaction analysis was performed by using steady-state methods with multiple injections of increasing compound concentrations over the immobilized DNA surface at 25 °C. Biosensor experiments were conducted in filtered, degassed HEPES buffer (10 mM HEPES, 100 mM KCl, 3 mM EDTA, 0.005 v/v of 10% P20 BIACORE surfactant, pH 7.3) at 25 °C. Flow cell 1 was left blank as a reference, while flow cells 2-4 were immobilized with DNA on a streptavidin-derivatized gold chip (SA chip from BIAcore) by manual injection of DNA stock solutions (flow rate of 1 μL/min) until the desired value of DNA response was obtained (350-400 RU). Compound solutions were prepared in with the running buffer by serial dilutions from stock solution. Typically, a series of different ligand concentrations (100 nM to 50 μM) were injected onto the chip (flow rate of 50 μL/min, 5-10 min) until a constant steady-state response was obtained followed by a dissociation period (buffer, 10 min). After every cycle, the chip surface was regenerated (20 s injection of 10 mM glycine solution, pH 2.0) followed by multiple buffer injections.

The instrument response (RU) in the steady-state region is proportional to the amount of bound drug and was typically determined by linear averaging over a 10-20 s or longer time span, depending on the length of the steady-state plateau. The predicted maximum response per bound compound in the steady-state region (RUmax) was determined from the DNA molecular weight, the amount of DNA on the flow cell, the compound molecular weight, and the refractive index gradient ratio of the compound and DNA, as previously described [27]. To obtain the binding constants, the data were evaluated with different interaction models to obtain an optimal fit using BIAevaluation (BIAcore Inc.) and Kaleidagraph (Synergy Software) software for nonlinear least-squares optimization of the binding parameters:

One site: $r = (K_1C_{free})/(1 + K_1C_{free})$

Two site: $r = (K_1C_{free} + 2K_1K_2C_{free}^2)/(1 + K_1C_{free} + K_1K_2C_{free}^2)$

Three site: $r = (K_1C_{free} + 2K_1K_2C_{free}^2 + 3K_1K_2K_3C_{free}^3)/(1 + K_1C_{free} + K_1K_2C_{free}^2 + K_1K_2K_3C_{free}^3)$

where K_1, K_2 and K_3 are equilibrium constants for three types of binding sites and C_{free} is the concentration of the compound in equilibrium with the complex and is fixed by the concentration in the flow solution.

2.2.7 Isothermal Calorimetry

ITC experiments were performed with a MicroCal VP-ITC (MicroCal Inc., Northampton, MA, USA). Phosphate buffer containing 10 mM K_2HPO_4, 3 mM EDTA, and 80 mM KCl and pH adjusted to 7.4 was used for the ITC experiments. The compound was injected into the DNA in the sample cell in 5 μL increments. The observed heat for each injection was determined by integration of the injection peak areas with respect with time. Blank titrations were conducted by injecting the compound into the sample cell containing only buffer under the same conditions. The corrected interaction heat was determined by subtracting the blank heat from that for the compound/DNA titration.

2.3 RESULTS AND DISCUSSION

2.3.1 DB832 Binds to the Human Telomeric DNA as a Stacked Species in Multiple Grooves

Circular dichroism has emerged as an important non-invasive technique for determination of the conformation of biomolecules and also provides insights about the binding modes of small molecules with DNA using pattern recognition [28]. Binding of ligands to DNA can be easily identified by the changes exhibited in the CD pattern by monitoring the compound and DNA wavelengths. When an achiral ligand, which has no CD by itself in solution, binds to a chiral macromolecule, an induced CD (ICD) signal is observed in the wavelength region corresponding to the bound achiral ligand [29]. This ICD signal can be a weak positive or negative, as in the case of duplex DNA intercalators, or can be a large positive signal, as in the case of duplex DNA groove binders [26b, 30]. Compounds that are known to interact with quadruplex structures have demonstrated preferential stacking at the terminal G-tetrads. Quadruplex end stackers, similar to duplex DNA intercalators, also exhibit a weak positive or negative ICD signal upon complex formation. In some cases, compounds are shown to effectively stack at the ends even without exhibiting any significant ICD signals [31]. Groove binding with quadruplexes has been a rarely observed phenomenon. As a result, a clear CD pattern that can used to identify possible groove binding agents has not been established until now. Shafer et al., using docking method, identified a carbocyanine dye, DODC (Figure 2.1-B), as a potential groove binder, and studied its interaction with an intermolecular dimeric hairpin quadruplex using CD [22]. This was the first ever reported binding study of a quadruplex interacting agent using CD. DODC binding to the quadruplex resulted in an induced CD signal in the wavelength region corresponding to the absorbance of the bound DODC. The large positive and negative ICD signal was hypothesized to be a consequence of the dye molecules stacking in the grooves of the quadruplex (Figure 2.3-A, adopted from the above reference).

CD experiments were carried out with DODC with other very similar quadruplex systems to confirm that a similar ICD pattern is exhibited for groove binding. Figure 2.3-B shows the CD spectra of DODC also with a dimeric hairpin quadruplex (Materials and Methods, Section 2.2.1), recently reported to exist as a mixture of parallel and antiparallel conformers. DODC exhibits an exciton-type splitting upon complex formation suggesting that compounds that can form stacked species possibly in the grooves of the quadruplex have an exciton-type splitting in the region corresponding to the absorbance of the bound ligand. DODC also exhibited a very similar ICD pattern upon complex formation with the wild-type human telomeric sequence, Tel22. Therefore, the large ICD pattern exhibited by small molecules upon complex formation is characteristic of the ligands stacking in the grooves of quadruplex systems. RHPS4 (Figure 2.1-A), a well-known small molecule shown to stack at the ends of human telomeric quadruplex DNA by NMR, was evaluated by CD to see if a ICD pattern exists for this class of compounds. As expected, RHPS4 did not exhibit any ICD upon complex formation (Figure 2.3-C) further suggesting that end-stacking compounds have very weak to almost no ICD with quadruplexes.

DB832 was recently reported to bind to the human telomeric quadruplex DNA, Tel22 d[$AG_3(T_2AG_3)_3$], possibly in the grooves of the quadruplex conformation as exemplified by a large positive and negative induced CD signals (Figure 2.5-A) [25]. The large magnitude of the induced CD was hypothesized due to the stacking of DB832 in different grooves of the telomeric quadruplex structure. Using the induced CD pattern exhibited by DB832 as a model for quadruplex groove recognition, CD studies of DB832 were carried out with Tel24, d[$TTG_3(T_2AG_3)_3A$], to investigate if a pattern similar exists (Figure 2.5-B). NMR studies have shown that Tel24 sequence forms a mixed parallel/antiparallel hybrid type structure in K^+ (Figure 2.2-A) [32]. When DB832 is titrated into Tel24, a large exciton-type ICD signal, very similar to Tel22, is observed in the wavelength region corresponding to the absorbance of

bound DB832. This suggests that the DB832 is also binding as a stacked species possibly in the grooves of this quadruplex system. Moreover, the DNA wavelength region does not exhibit much change upon compound binding suggesting that DB832 is preferentially binding to the pre-formed hybrid conformations.

Based on the CD studies of Tel22 and Tel24, however, it is rather difficult to definitively establish the groove binding aspect of DB832. Therefore, if such a system exists, where the grooves of the quadruplex systems are rendered sterically inaccessible, small molecules that selectively bind in the grooves can be completely prevented from interacting in the grooves. Structural features of G-quartets reveal that the aromatic proton, H8, of the guanine bases is directly pointed into the grooves of the quadruplex. If those protons could be replaced by a bulky substituent, small molecules can be directly inhibited from interacting in the grooves. Bromo-guanines turn out to be the best possible candidate to test such a hypothesis. Sugiyama and coworkers have successfully employed these modified guanine bases and shown that stable quadruplex architectures could be maintained with this modification [12a, 33]. The bromine substitution on the C8 position of guanine further locks the *syn* glycosidic conformation of guanine nucleotides further adding to the stability of the fold [34]. Sugiyama and coworkers recently showed that ODN9, d(AG_2GGTTAG_8GGTTAGGGTTAG_{20}GG), folds in a mixed parallel/antiparallel hybrid conformation, which is likely the structure the unmodified human telomere adopts in solution in the presence of K^+ [34]. The three bromoguanine substitutions correspond to the guanines in the terminal tetrad at the 5'-end of the hybrid fold. These bulky bromine substituents selectively inhibit binding in all the three accessible grooves of the quadruplex (Figure 2.2-C), thereby, effectively hindering the interaction of any groove binding molecules. The fourth groove in the hybrid quadruplex conformation is inaccessible due to a diagonally running loop connecting two adjacent strands.

CD studies of DB832 were conducted with the ODN9 sequence to further test the groove binding mode. CD spectra of DB832 titrated into ODN9 shown in figure 2.5-C. Most interestingly, DB832 does not exhibit the characteristic exciton-type splitting, that is observed with unmodified telomeric sequences, with the ODN9 sequence. A weak ICD signal at significantly high ratios of DB832 is observed, suggesting that the interaction of DB832 with ODN9 is primarily at the ends. Since DB832 does not exhibit any exciton-type splitting with this hybrid quadruplex structure when the grooves are rendered inaccessible by bromo-guanine substitution, it provides evidence that the mode of binding of this compound to the human telomere is groove binding as a stacked species.

Another modified human telomeric sequence with bromine atoms at different positions was also tested to confirm the groove binding of DB832. ODN4, d(AGGGTTAGGGTTA$G_{14}G_{15}G_{16}$TTAGGG), was also shown to fold into a mixed parallel/antiparallel hybrid conformation, similar to the unmodified human telomeric sequence [34]. However, in this case, the three bromine substitutions were located in a single quadruplex groove (Figure 2.2-D). Therefore, small molecules that interact in the grooves should still be able to bind in the remainder of the two accessible grooves. CD studies of DB832 were conducted with ODN4 sequence (Figure 2.5-D). Interestingly, exciton-type ICD signal is still observed with this sequence upon titration of DB832. However, the magnitude of the ICD is comparatively smaller than what is observed with the unmodified telomeric sequences. The ICD signal generated upon complex formation is due to the formation of stacked species of DB832 in the remaining two grooves of the hybrid fold, resulting in decreased intensity.

The 27-nt G-rich sequence in the NHE-III promoter element of the *c-myc* oncogene has been shown to fold into a parallel quadruplex architecture [35]. The connecting loop bases between the guanine stretches can form different loop isomers while maintaining the same

parallel scaffold in different variants of *c-myc*. Moreover, the connecting bases form a double-chain reversal loop and completely block all the grooves of the quadruplex. This parallel quadruplex fold is also an excellent system to test for small molecules that interact primarily in the grooves of the quadruplex. The loop bases should act as a natural steric blockage for groove interacting ligands. To test this idea, CD titrations of DB832 were carried out with a dual mutant *c-myc* variant, d[T(AGGGTGGGG)$_2$AA], that was shown by NMR to form a stable parallel conformation [35b]. Figure 2.6-C shows the CD spectra of DB832 with *c-myc* sequence. The absence of any exciton-type ICD signal indicates that the compound is not able to form stacked species due to the steric blockage of the grooves. This also suggests the preferential affinity of DB832 for the hybrid-type quadruplex scaffolds and the exciton-type splitting observed with the telomeric sequences is as a result of stacking molecules in the grooves.

The combined CD results of Tel22, Tel24, ODN9, ODN4 and *c-myc* strongly suggest the groove binding mode of DB832 with the human telomeric quadruplex sequence.

2.3.2 DB832 is the First Reported Quadruplex Groove Binder for Biologically Relevant Telomeric Quadruplex Sequences

Distamycin-A (Figure 2.1-D) is a well known small molecule that has high affinity for AT-rich duplex sequences [36]. Small molecules that are excellent duplex binders are often tested for their affinity for quadruplex systems. Distamycin-A was shown to interact with simple quadruplex systems, such as d[TG$_4$T]. The interaction of distamycin with the aforementioned sequence and very similar quadruplex systems has been shown primarily by stacking at the terminal tetrads by NMR [20b]. However, a recent high resolution NMR study has shown that distamycin can also stack in the grooves of the same quadruplex system, but at significantly higher concentrations [20a]. The ligand binds as an antiparallel dimer in the two

diametrically opposite grooves of the quadruplex with a 4:1 stoichiometry. However, there have not been any studies related to distamycin binding in the grooves of biologically relevant quadruplex systems. Considerable resources are still being devoted to test the quadruplex interacting potential of distamycin to non-targeted simple model systems.

To test the groove binding potential of distamycin with biologically relevant quadruplex conformations, such as the human telomeric system, CD studies of distamycin were conducted with Tel22 sequence (Figure 2.6-A). Interestingly, the CD spectrum of distamycin does not show any exciton-type ICD signals proposed for ligands forming stacked species in the grooves of the quadruplex. The interaction of distamycin at low CD concentrations might be by stacking at the ends of this quadruplex system, supporting the NMR studies with simple quadruplex systems at low concentrations. Moreover, thermal melting studies (data not shown) of distamycin with various telomeric systems showed a negligible change in the stability of the quadruplex. Therefore, the validity of distamycin, as a groove binder, for biologically relevant quadruplex systems is a matter of question.

Virtual screening and docking studies of a library of small molecules have shown that a subset of ligands also selectively target the grooves of the simple quadruplex systems [21]. However, their affinities for duplex systems have not been reported so far. These compounds were proposed by NMR to bind in the grooves of the d[TG_4T] system with optimum hydrophobic and van der Waals interactions along the grooves of the quadruplex. To further test the selectivity of these ligands, F1190 (Figure 2.1-E), the most potent compound from the series, was chosen to evaluate its interaction potential with a relevant quadruplex system, such as Tel22. CD spectra of F1190 with Tel22 (Figure 2.6-B) exhibited virtually no ICD signal even at high ratios of the ligand. Absence of the ICD suggests that this compound might not be interacting in the grooves of the quadruplex system.

Apart from these two reported molecules as groove binders for simple quadruplex models, no other small molecule has been reported to selectively target the grooves of human telomeric quadruplex DNA. Therefore, DB832 is the first heterocyclic diamidine that is shown to selectively bind as stacked species in the grooves of the telomeric DNA with preferential binding for the hybrid-type conformation.

2.3.3 DB832 Induces the Formation of Hybrid Conformation in Telomeric Quadruplex DNA

To further confirm the groove binding mode of DB832, Fluorescence Intercalator Displacement Assay (FID) and competition CD experiments were performed (by Caterina Musetti from Dr. Wilson's group). FID assay is based on the displacement by small molecules of a fluorescent probe from DNA [37]. It allows the ranking of a set of putative ligands based on their ability to displace the probe from the specific DNA region where the probe interacts. Also based on the affinity of a ligand, a general idea about the binding site of the ligand can be obtained. The choice of what molecule to use as probe was based on recent studies with Thiazole Orange (TO) [38]. This dye was used as a displacement probe because its excitation wavelength (501 nm) is higher than most compounds and therefore does not interfere with the assay. Furthermore, TO, shown to bind as an end-stacker, has a binding affinity in the order of *ca.* 3 x 10^6 M^{-1} for human telomeric DNA [39], making it an ideal probe to study. Figure 2.7-A shows the graph of the displacement assay of DB832 and a well known quadruplex end stacker, TMPyP4. Distamycin was also tested as a control molecule based on CD results. DB832 has a binding affinity of *ca.* 1.4 x 10^6 M^{-1} with Tel22 quadruplex sequence (discussed later in the SPR results). From the graph, it is evident that DB832 is not able to displace TO even at the highest concentrations. This is most likely not due to binding constant differences but due to the fact that DB832 preferentially interacts with the grooves of the human telomeric sequence. In fact, the same assay, performed with *c-myc,* indicates that the compound is able to displace

30% of the TO at low concentrations (4.5 μM). Since *c-myc* is known to fold in a parallel conformation with no accessible grooves and based also on results obtained by CD studies, this assay reinforces the hypothesis that DB832 interacts with the human telomeric grooves. TMPyP4, which has a strong affinity for the ends of the human telomeric quadruplex, displaces TO completely even at modest concentrations. Distamycin, which was earlier shown by CD not to interact with telomeric quadruplex DNA, also does not displace TO even at high concentrations.

To test the results of FID assay, a competition CD experiment was performed with Tel22 and DB832 with TO as a probe (Figure 2.7-B). TO has been shown to preferentially bind to the antiparallel conformation of human telomeric DNA [37, 39]. CD spectra (Figure 2.7-B) of Tel22 reveal the presence of 260 nm and 290 nm peaks in the DNA absorbance region, characteristic of a hybrid-type quadruplex conformation. Upon titration of TO, there is a decrease in the intensity of 260 nm peak and a subsequent increase of the 290 nm peak, indicative of an increased antiparallel characteristic of the sequence upon complex formation. TO, being a quadruplex end stacker, do not exhibit any ICD signals upon complex formation either. When DB832 is titrated upon saturation with TO, significant changes are observed in the DNA wavelength region (Figure 2.7-C). The peak at 290 nm decreases in the magnitude slowly whereas the 260 nm peak remains the same. However, when the compound exhibits the induced CD signal formation, a drastic reduction in the 290 nm peak is observed with a concomitant change in the 260 nm peak. Upon complete saturation, the DNA region exhibits peaks at both 260 nm and 290 nm, very similar to the starting conformation - the hybrid conformation - of the DNA but of lesser magnitude. Therefore, DB832 was able to induce a conformational switch from an antiparallel fold, induced by TO, back to the hybrid fold.

2.3.4 DB832 Exhibits High Selectivity for Human Telomeric Quadruplex DNA over Duplex DNA

With the abundance of duplex DNA in genome, targeting relatively very few non-duplex structures such as quadruplexes becomes a complex task. Non-selective binding to secondary sites might result in deleterious effects and cytotoxicity, therefore, structure and sequence selectivity is of utmost importance to achieve desired results. In order to determine the selectivity of DB compounds for human telomeric quadruplex sequence over duplex DNA, thermal melting studies were performed with telomeric quadruplex DNA. Figure 2.4 shows the UV melting profile of DB832 with a duplex sequence and two human telomeric quadruplex forming sequences, Tel24 and Tel26 (Materials and Methods). DB832, showed a ΔTm of ~12 °C and an impressive 20 °C for Tel24 and Tel26 quadruplex sequences respectively at 6:1 ratios indicating signification stabilization of the quadruplex conformation. From the UV profile it is highly apparent that DB832 has a very low stabilizing potential for duplex (ΔTm ~ 3 °C). The quadruplex stabilization potential observed for DB832 rank among the strongest for quadruplex-interacting small molecules reported so far. It is highly apparent that the DB832 is significantly stabilizing the quadruplex conformation, while maintaining a low degree of selectivity for duplex sequences. Therefore, DB832 makes an ideal candidate for further study as a highly selective quadruplex interacting agent.

2.3.5 Multiple Binding Modes Observed in NMR

Telomeric quadruplex sequences in K^+ exists as a mixture of conformations [3a, 12f], with the hybrid structure predominant. The NMR structure of Tel24 reveals that this sequence folds into a mixed parallel/antiparallel hybrid-1 type structure [32]. To test the groove binding mode exhibited by DB832 from CD, ^{1}H-NMR titration studies were conducted with Tel24. Imino protons of guanines are excellent probes to determine if a G-rich sequence folds into a

quadruplex conformation [40]. These imino protons resonate between 10.0 to 12.5 ppm in a quadruplex structure and are readily observable in a water sample. Moreover, the presence of multiple quadruplex conformations can also be readily addressed by monitoring the imino protons. Binding of a small molecule to a quadruplex structure affects the chemical environment of imino protons which can also be identified from simple titration experiments. Figure 2.8 shows the guanine imino proton NMR spectra of DB832 with the Tel24 sequence up to ratios 2:1 at 25 °C. At higher compound to quadruplex ratios, especially at NMR concentrations, solubilities of the compound and the complex become an issue. Considerable aggregation is observed resulting in significant line broadening and rendering it difficult to obtain spectra for analysis. In the absence of any compound (Fig 2.8, 0:1 ratio), the spectra show distinct number of imino protons corresponding to the total number of guanines in the Tel24 sequence. This shows that the sequence forms primarily a single structure in K^+, and in this case it is the hybrid-1 type structure. Titration of DB832 into the DNA solution resulted in significant changes in the imino proton spectra and spectral broadening and shifts revealed that the ligands were in intermediate exchange on the NMR timescale. A number of imino resonances broaden in a site-specific fashion, consistent with the local chemical environment of various residues being perturbed by ligand binding. Distinct chemical shift changes can be readily observed for most of the imino protons. Imino protons of G3, G9, G17 and G21, which constitute the 5' terminal tetrad of the quadruplex, are the most perturbed even at the lowest compound to quadruplex ratio. This suggests the inherent sensitivity of this technique even for the smallest changes produced, at relatively low ratios, upon complex formation. The 5' G-tetrad imino protons broadened considerably and disappeared completely at high ratios.

The high perturbation of the imino protons at the top end is consistent with the stacking of the ligand on the 5' G-quartet. NMR studies of distamycin, Hoechst-33258, and ethidium with several quadruplex forming sequences has revealed similar perturbations of the imino protons of the top tetrad of the quadruplex indicating stacking on the top end of the

quadruplex [40]. Small molecules that have excellent duplex-DNA groove binding properties, such as Hoechst-33258 and distamycin, can also preferentially bind as end-stackers with quadruplex systems. Therefore, it is equally conceivable for the currently studied ligands to exhibit a similar end-stacking as the initial mode of binding. Interestingly, broadening occurs for the aromatic residues of the 5' G-tetrads (G3, G9, G17 and G21 of 5'-tetrad) and A20·T1 base pair (spectra not shown), providing support for a mixed intercalation-exterior stacking mode of binding (i.e., between the 5'-tetrad and the capping A20·T1 base pair of the loops). Therefore NMR results show the stacking of DB832 on the top quartet at low ratios. In most of the cases, imino protons of G11and G23, that constitutes the bottom tetrad, exhibited very little change. This is probably due to the presence of a stable capping A24·T13 base pair at the 3' terminal, offering significant protection from solvent exchange. Also, relatively unchanged imino proton chemical shifts of the bottom tetrad indicates that 3' end of the Tel24 quadruplex structure is least likely the binding site for DB832. Imino protons of G4 and G22 that correspond to the adjacent guanine of the middle tetrad exhibit small downfield shifts accompanied by slow broadening; however, imino protons of G10 and G16 of the same tetrad exhibit very little changes. G4 and G22 are located in a groove that has optimal dimensions for compounds to form stacked species. Therefore, after the initial end-stacking mode at the 5' terminal, there is a very likely possibility that DB832 is targeting the "medium" groove encompassed by G3, G5, G23, G21 plane. The groove adjacent to the medium groove, comprised by the plane of G15, G17, G21, and G23 residue is classified as the narrowest groove. Therefore, it is highly unlikely for DB832 to efficiently form stacked species in the narrow groove. Absence of any major changes in imino protons of guanines (G16 and G15) further confirms this.

Figure 2.9 shows the aromatic proton spectra of Tel24 titrated with DB832. Similar to the imino protons, aromatic protons of the bases can also be used as probes to verify the interaction of ligands [41]. The aromatic protons of the guanine residues, involved in the tetrad formation, point directly into the grooves of the quadruplex. Ligands that can bind in the grooves of the quadruplex can perturb the local environment of these protons and, therefore, these protons can be used to monitor the drug-quadruplex interaction. Titration of DB832 into the DNA solution resulted in significant changes in the aromatic proton spectra and spectral broadening and shifts revealed that the ligands were in intermediate exchange on the NMR timescale. A number of aromatic resonances broaden in a site-specific fashion, again consistent with the local chemical environment of various residues being perturbed by ligand binding. Interestingly, broadening occurs for the aromatic residues of the 5'-end G-tetrad: G3, G9 and G21, providing support for a mixed intercalation-exterior stacking mode of binding (i.e., between the exterior tetrads and the capping A20·T1 base pairs of the loops). Surprisingly, G17 undergoes significant downfield shifts followed by broadening. Aromatic protons of A20·T1 base-pair undergo significant perturbations (broad, downfield shifts) upon ligand binding indicating a preferential stacking at the 5'-terminal tetrad. The guanine aromatic protons of the 3'-end tetrad undergo very small changes, except for G5, suggesting the 3'-end might be an optimal binding site for stacking of DB832. This was also observed with the imino protons of the same residues. However, the aromatic protons of A24·T13, that forms a stable base-pair at the 3'-end undergoes broad downfield shifts. Therefore, the ligand might be exhibiting a weak stacking at the 3'-terminal tetrad. The stacking of the ligand might be occurring by a conformational rearrangement of the two base-pairs to accommodate the ligands in an efficient manner. Aromatic protons of G4 and G22 that constitute the middle tetrad, and G5 of the bottom tetrad, undergo small upfield broadening. Whether this change is due to the binding of the ligands in the grooves is not clear. SPR and ITC studies (discussed

later in Section 2.3.6) show a weak secondary binding by the ligand which is hypothesized to be groove binding. Therefore, it is a possibility that such weak interactions might not be easily detectable using traditional NMR methods.

2.3.6 DB832 Binds to the Human Telomere Cooperatively with High Stoichiometry

Well over half a century ago, George Scatchard, a renowned physical chemist, noted that the following questions *'How many? How tightly? Where? Why?* must be answered in order to understand the fundamentals of ligand binding to DNA [42]. There is not one single experimental technique that can provide the answers to all these questions and, thus, a multifaceted approach has to be employed to elucidate drug-DNA interactions. Effective drug design requires sufficiently clear, thermodynamic, kinetic and structural details of the drug-DNA interactions, a picture that can only be unraveled by combining the information from different methodologies. With regards to DB832, sufficient information is available to answer *Where? How tightly?* and, to some extent, *Why?* the ligand is selective for quadruplex DNA. However, the question of *How many?* has always been the most challenging to answer.

To decipher the number of DB832 molecules that are bound to a quadruplex unit, a combination of CD, ITC and SPR techniques were used. Titration curves from CD experiments were plotted for the CD signal at the maximum absorbance wavelength of bound DB832 as a function of molar ratio for different telomeric quadruplex sequences (Figure 2.10). From the plot, it is clear that DB832 binds to the quadruplex system with a high stoichiometry. Due to the high stoichiometry of the system, it is difficult to fit the data in order to obtain the exact number of bound molecules. However, a straightforward analysis of the plot reveals about 5-6 DB832 molecules are bound per quadruplex unit. Closer inspection of the titration curves for different telomeric sequences also show a clear breakpoint at different added ratios of DB832.

The first break occurs at an added ratio of around 3:1 for all the telomeric sequences. However, the second break varies for different quadruplex sequences, but is generally occurring around 6:1 compounds /quadruplex ratio. The two breakpoints exhibited by the compound upon complex formation suggest multiple binding events exhibited by the compound. This was also highly evident from NMR studies (Section 2.3.5). The initial binding observed until 3:1 ratios, characterized by weak ICD signal, is due to the stacking of the ligand at the terminal tetrads of the quadruplex, as also suggested by NMR results. The latter binding observed at higher ratios until around 6:1, characterized by strong excition-type splitting, is due to the recognition of quadruplex grooves as stacked species. The sigmoidal shape of the curve also indicates that the binding of DB832 to the quadruplex sequences occurs with positive cooperativity. Small molecules, such as porphyrins, can aggregate on the surface of quadruplex-DNA, resulting in false high stoichiometries, due to the non-specific binding. However, in the case of DB832, a clear saturation of the ICD signal is observed at higher molar ratios (Figure 2.10, ratios above 10 are not shown) suggesting that the high stoichiometry is not due to the aggregation or non-specific interactions, and the binding of DB832 to the telomeric quadruplex DNA is indeed distinct.

To further test the binding stoichiometry of DB832, ITC experiments were conducted with telomeric quadruplex DNA, Tel24, with different DNA concentrations (Figures 2.12-A, B). Plots of the observed net binding heat/mol versus molar ratio were obtained by subtracting the integrated peak areas for the blank titration (DB832 into buffer) from the areas in the DB832-DNA interaction titration. Figure 2.12-A shows the ITC plot of DB832 titrated into 10 μM quadruplex DNA. From the plot it is clear that DB832 is indeed binding with high stoichiometry and with multiple binding modes as characterized by distinct breaks in the titration curve at different ratios. Surprisingly, the compound exhibits some non-specific binding at ratios higher ratios (> 6:1), but the heat observed at those ratios are very small. ITC

titrations were also conducted with a higher DNA concentration (20 μM Figure 2.12-B) to understand the thermodynamics at lower drug-DNA ratios. From the titration curve it is again clear that the ligand is binding with distinct binding modes with the telomeric DNA. An initial strong binding (up to ratio 2:1) is followed by a weaker secondary binding (ratios above 2:1). The initial strong binding, as suggested by CD and NMR, is probably due to the stacking of DB832 at either one or both ends of the quadruplex. The latter weaker binding is due to the recognition of multiple grooves of the quadruplex as stacked species. Due to the high stoichiometry, it was difficult to obtain complete thermodynamic profile for DB832 binding to the DNA. Nevertheless, based on the CD and ITC data analysis, results suggest that approximately 6 molecules of DB832 are bound per quadruplex unit.

SPR was used to quantitatively evaluate the interactions between DB832 and a series of telomeric quadruplex DNA in order to gain insight into their affinities (Figure 2.11: A-F). From the sensorgrams, the large response obtained suggests that DB832 is binding to all the telomeric sequences, and exhibit fast association and dissociation kinetics. Sensorgrams of Tel24 and Tel26, Figures 2.11-B and C respectively, are shown for high concentrations of DB832 (up to 50 μM) to show the behavior exhibited by the ligand, characterized by high response values. Responses obtained at such high ligand concentrations are generally due to non-specific binding, and are not included during evaluation of binding affinities. From the steady-state binding plots, fit to a two-site model for low ligand concentrations (up to 1 μM), we can readily notice that the primary binding constant (K_1) for different sequences fall in the range of 1 x 10^6 to 4 x 10^6 M^{-1}, and is followed by a weak binding constant (K_2) that is 10-fold weaker in all cases. Although the primary binding constant obtained for DB832 is not high, it does fall in the range of values that is generally observed for compounds that interact at the ends of the quadruplex [25]. The secondary, 10-fold weaker, binding observed is most likely the binding event occurring due to the interaction of DB832 in the grooves of the

quadruplex as stacked species. Thermodynamic studies of distamycin-A interacting with a parallel quadruplex conformation were also characterized with binding affinities in the same range (~4 x 10^5 M^{-1}) [20a]. Therefore, this might be a general interaction affinity that is observed for compounds that bind in the grooves of the quadruplexes. Therefore, as aforementioned, small molecules that are excellent duplex-DNA groove binders, such as Hoechst-33258 and distamycin, initially interact at the ends of the quadruplex conformations due to their weak affinity for quadruplex grooves. This again highlights the difficulty in designing compounds that are very specific to the quadruplex grooves; however, ligands that are shown to target grooves of quadruplexes can be used as paradigms to develop more potent ligands with enhanced affinities.

2.4 CONCLUSION

Quadruplex grooves offer an attractive target to design new classes of compounds that can bind strongly and selectively. Since groove dimensions vary significantly according to the type of quadruplex, groove-binding also offers the opportunity for obtaining increased selectivity for a particular quadruplex structure. Selectivity of a compound for its target quadruplex structure is important in order to reduce cytotoxicity from duplex-binding as well as prevent drug loss from binding to non-targeted quadruplex sites.

DB832 is a heterocyclic diamidine that recognizes multiple sites of the human telomeric DNA, showing virtually no binding to duplex DNA. This is the first reported small molecule that is shown to recognize grooves of the wild-type telomeric quadruplex conformation. Although, the initial binding mode observed with DB832 is end-stacking, the latter binding which is the recognition of the grooves as a stacked species is a much more interesting phenomenon. The unique binding mode of DB832 as well as its selectivity for its

DNA target may allow it to serve as the starting point for the design of a new class of highly selective groove-binding molecules. Potentially, the individual monomer units could be covalently linked, dramatically increasing the affinity and selectivity of the compound for human telomeric DNA, leading to enhanced telomerase inhibition and anti-protozoan activity with decreased cytotoxic side effects.

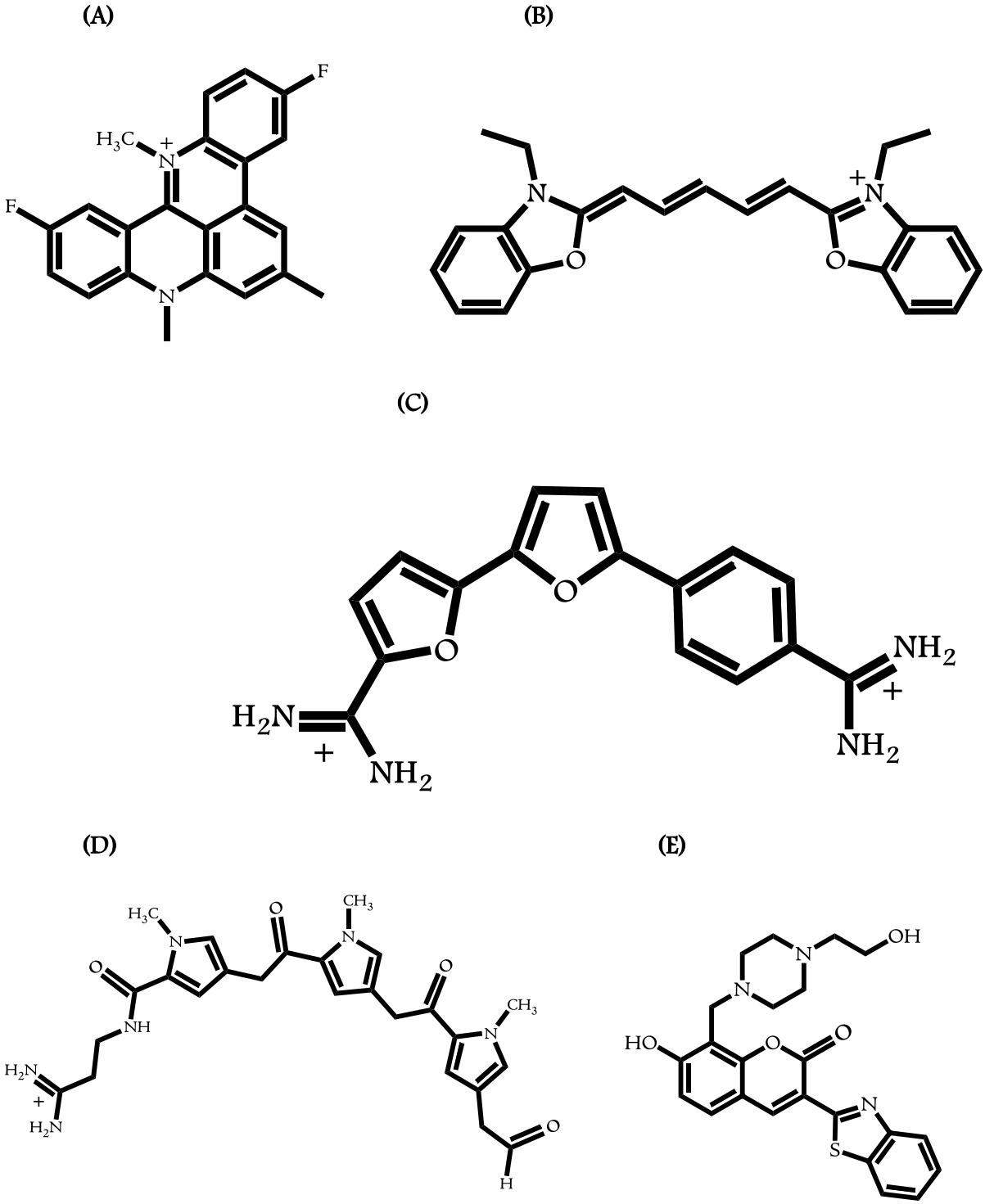


Figure 2.1: Chemical structures of RHPS4 (A), DODC (B), DB832 (C), Distamycin (D) and F1190 (E).

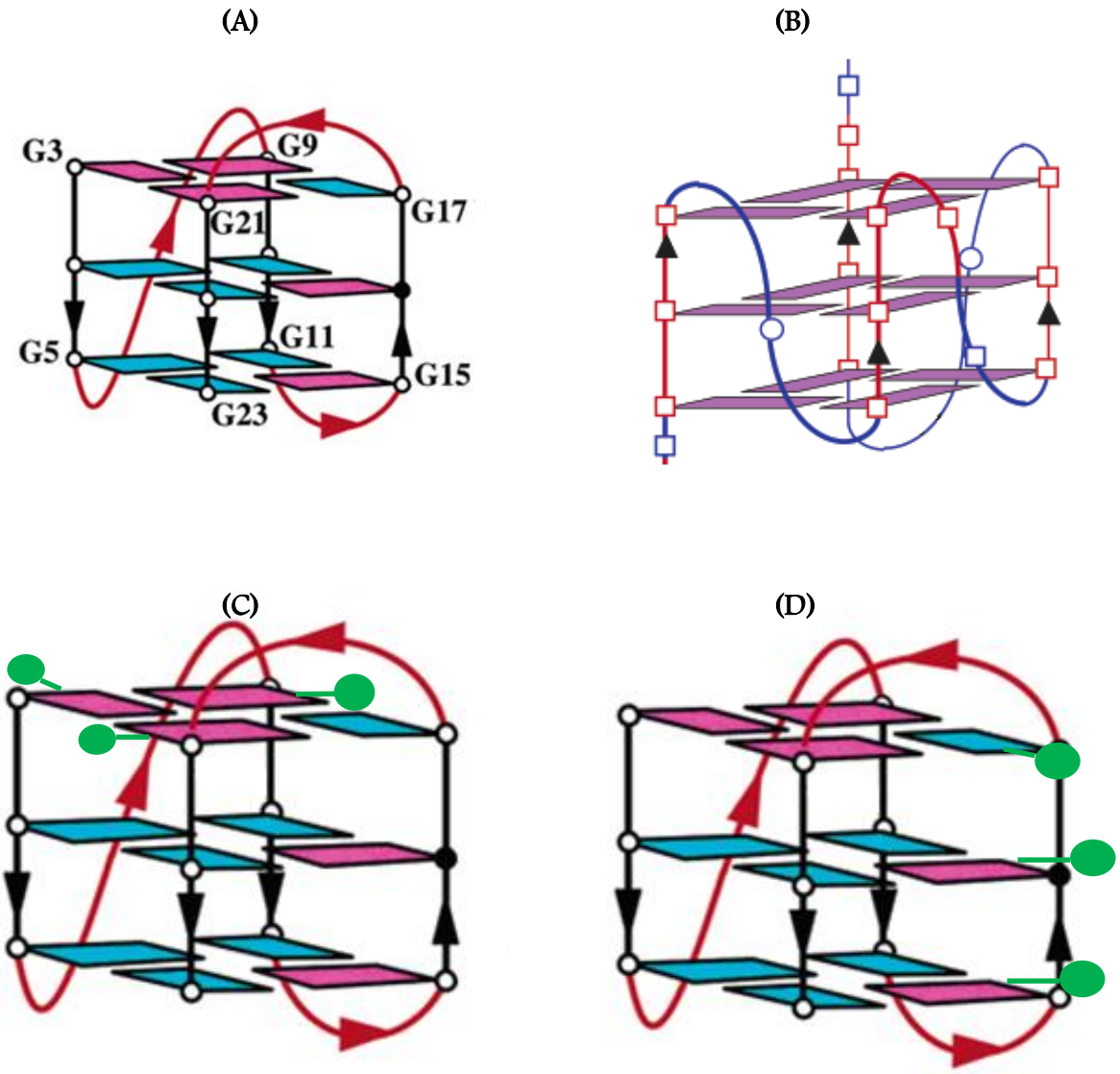


Figure 2.2: Folding patterns of different quadruplex-forming motifs: Tel24 [32] (A), *c-myc* [40] (B), ODN9 [34] (C), and ODN4 [12a] (D) used in this study. The bulky bromine atoms are represented as green ball and stick.

All the quadruplex cartoons are directly obtained from their published journals without further permission.

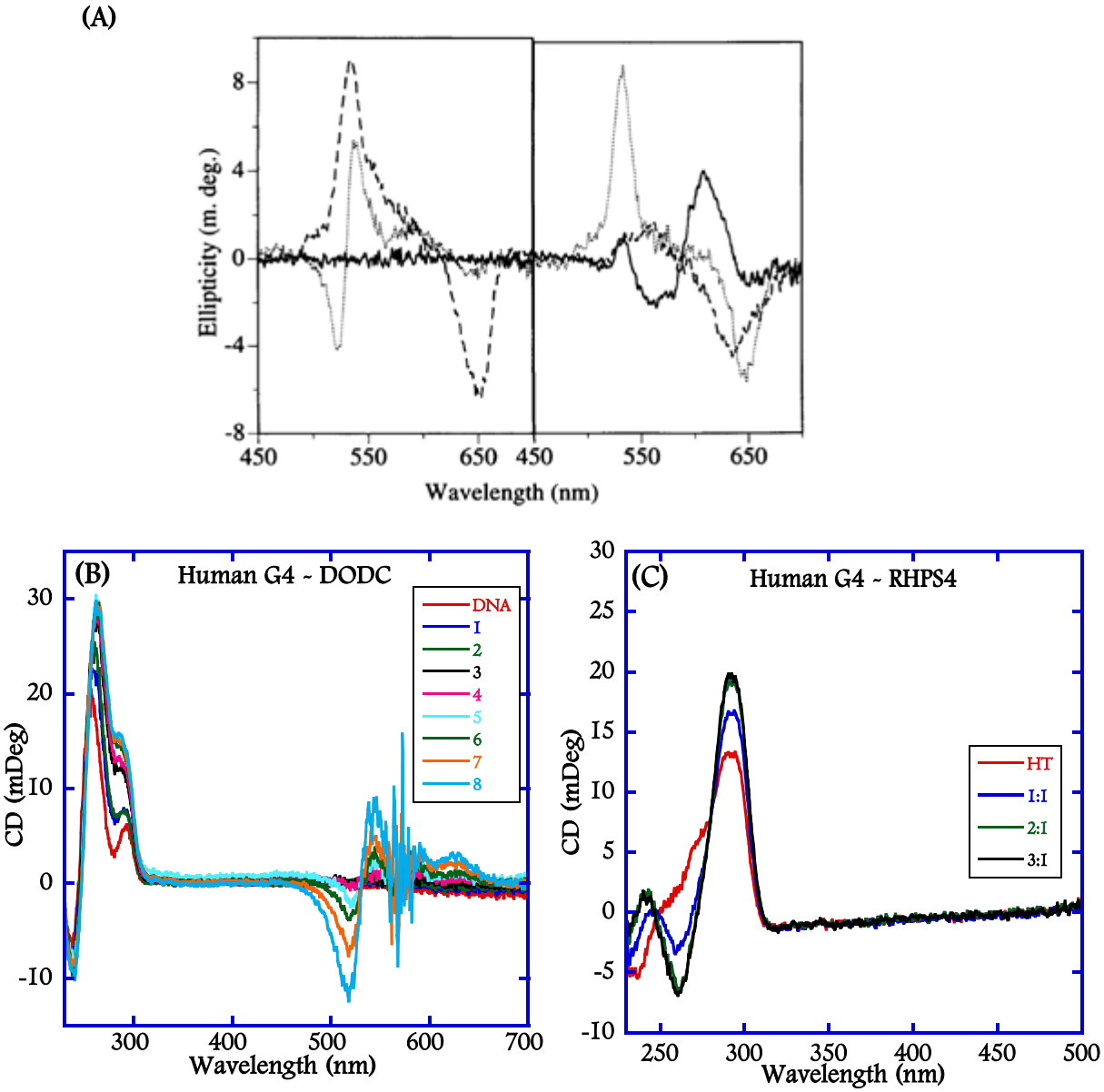


Figure 2.3: (A) Induced CD spectra of DODC with intermolecular quadruplex forming sequences and a duplex sequence as reported by Shafer et al. (B) CD spectra of DODC with a dimeric quadruplex sequence, U6U7 in K+, and (C) CD spectra of Tel22 in K+ with RHPS4, a well-known human telomeric quadruplex end-stacking agent.

The exciton-type splitting observed in (A-dashed lines, dotted lines) is proposed for the formation of stacked species by the DODC dye possibly in the grooves of the quadruplex. A similar excition-type is observed for DODC with a dimer quadruplex. A weak or an absence of exciton-type splitting is characteristic of compounds that are known to stack at the ends of the quadruplex, as shown by RHPS4 (C). Buffer conditions: Phosphate buffer containing 100 mM K+ at 25 °C.

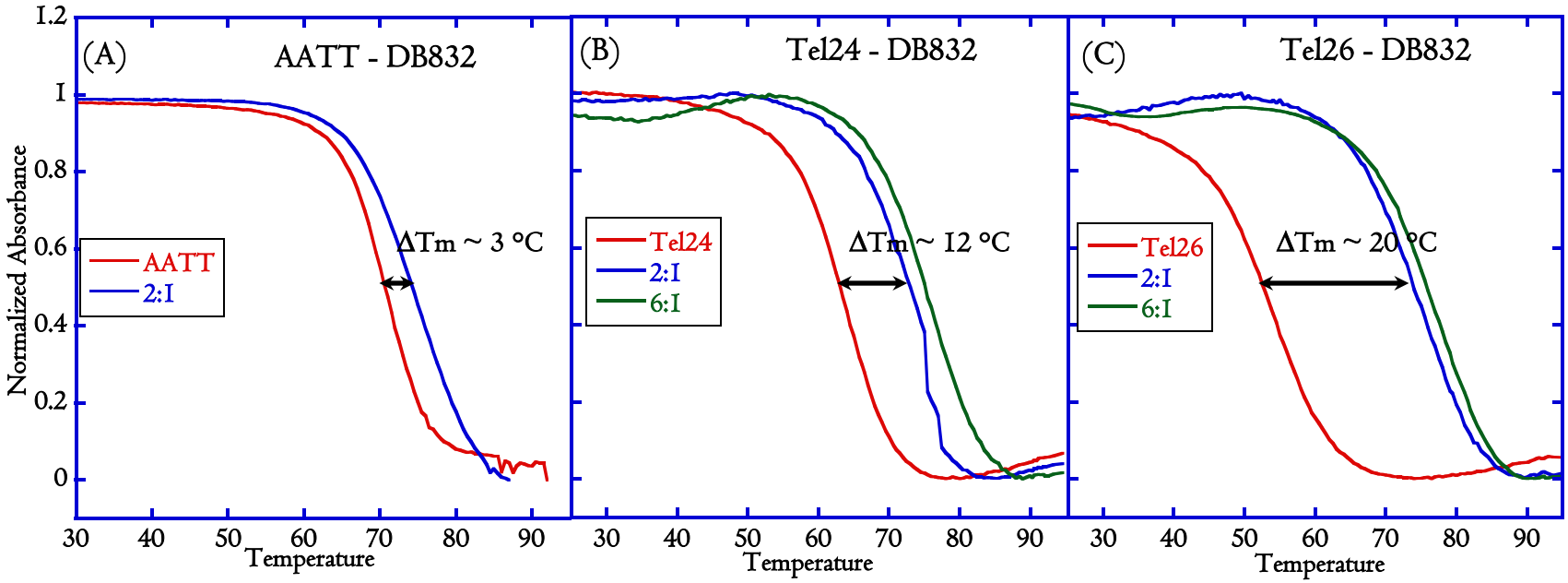


Figure 2.4: UV melting profiles of the hairpin duplex, d(CGCGAATTCGTCTCCGAATTCGCG), monitored at 260 nm (A) modified human telomere sequences, Tel24 (B), and Tel26 (C) monitored at 295 nm in the absence and presence of DB832 in phosphate buffer containing 100 mM K^+.

DB832:DNA ratio for duplex sequence was 2:1, while, DB832:DNA ratio for quadruplex sequences was 2:1 and 6:1. DB832 increases the melting temperature of the Tel24 by approximately 12 °C, and that of Tel26 by an impressive 20 °C at 6:1 ratios, while it has a negligible effect on the melting temperature of the duplex sequence. DNA/quadruplex concentration for Tm and CD experiments is ~3-5 μM unless otherwise mentioned.

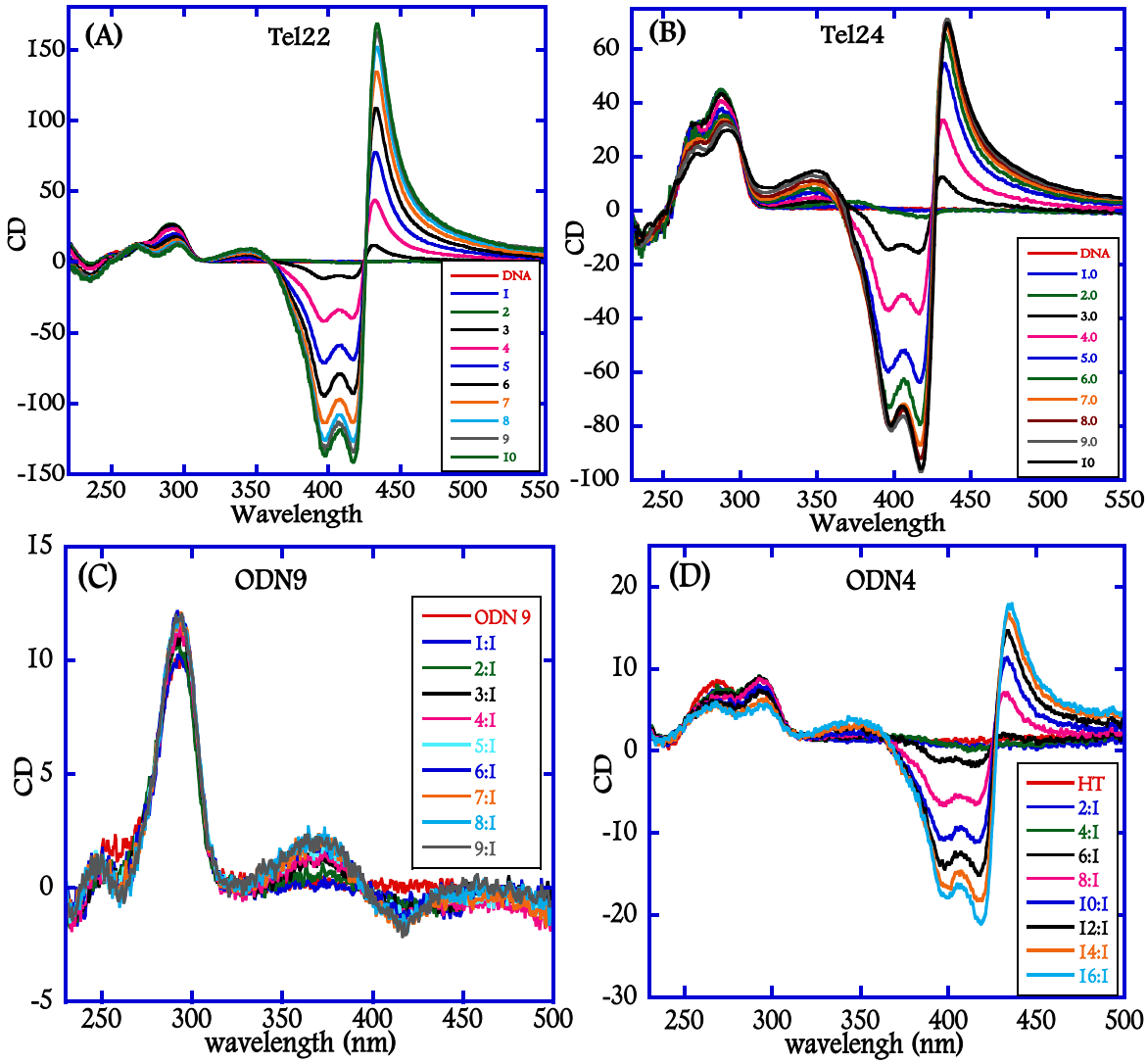


Figure 2.5: CD spectra of DB832 titrated into (A) Tel22 (B) Tel24, (C) ODN9, and (D) ODN4 quadruplex forming sequences.

DB832 was titrated into DNA solutions until saturation point is reached in the ICD region. Tel22 and Tel24, which has multiple accessible grooves, can facilitate the formation of stacked DB832 molecules in their grooves. ODN9, with all the grooves blocked by the bulky bromine atoms exhibit no exciton-type splitting, confirming the groove binding mode of DB832 as stacked species. The small induced CD observed with ODN9 is possibly due to the stacking of DB832 at the terminal tetrads. ODN4, with one the grooves completely blocked by bromine atoms, exhibit an exciton-type splitting but of lesser magnitude. The decrease in the magnitude of ICD signal might be due to the blockage of one of the grooves preventing the formation of stacked species in that groove. Buffer conditions: Phosphate buffer containing 100 mM K^+ at 25 °C.

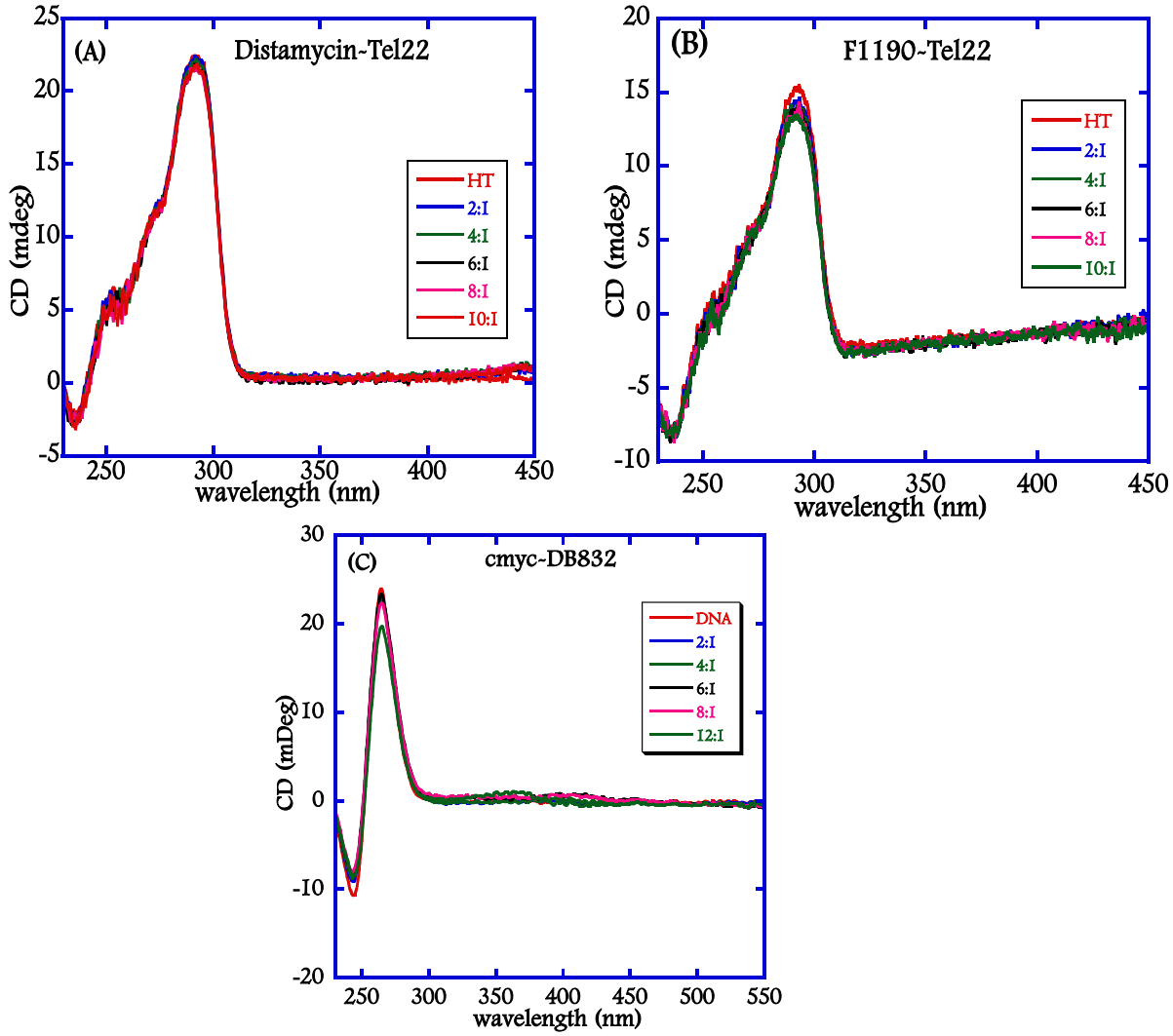


Figure 2.6: CD spectra of Distamycin with Tel22 (A), F1190 with Tel22 (B), and DB832 with *c-myc* (C) in K^+.

Distamycin has been proposed to be the first small molecule to bind in the grooves of quadruplex systems. However, this has been proven only with simple quadruplex systems at high NMR concentrations. CD studies of distamycin with biologically relevant quadruplex systems such as human telomeric sequences do not show any ICD signal, refuting the groove binding mode of this molecule. F1190 was also shown to interact with the grooves of simple quadruplex systems by docking and preliminary NMR studies. However, when F1190 was tested with full length human telomeric sequence, no ICD signal is observed, invalidating the proposed quadruplex groove binding, c-myc quadruplex sequence is shown to fold into an all parallel topology in which the grooves are completely blocked by the connecting loop bases. Groove targeting ligands, such as DB832, should not be able to form stacked species in the grooves of such systems. As expected, no ICD signal is observed even at high ratios of DB832 with c-myc, confirming the groove binding mode of the compound.

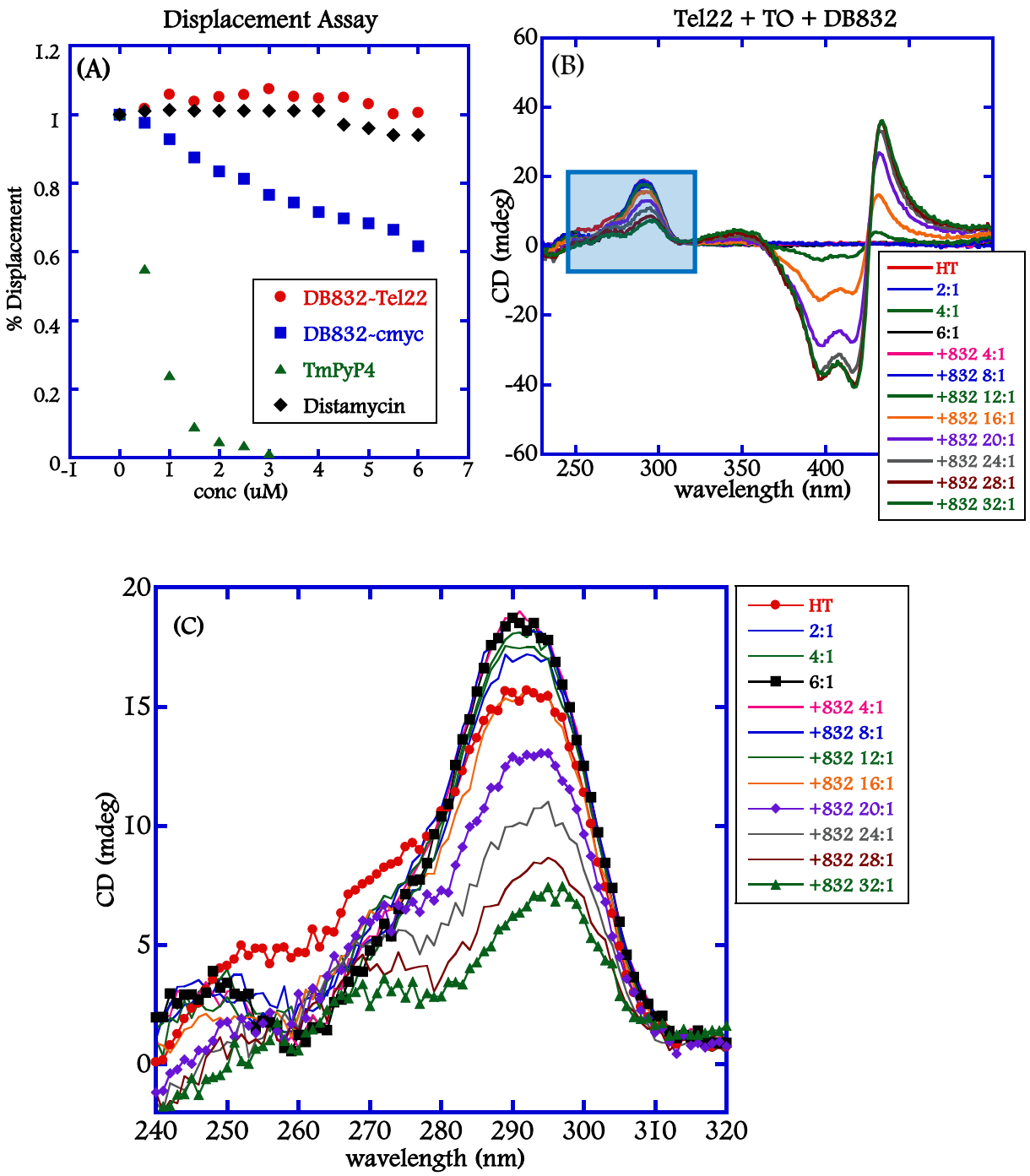


Figure 2.7: Fluorescence Displacement assay (A) and competition CD experiment (B) performed with Tel22 sequence using TO as the fluorescent probe. (C) Close-up of the DNA absorbance region highlighted in blue in (B).

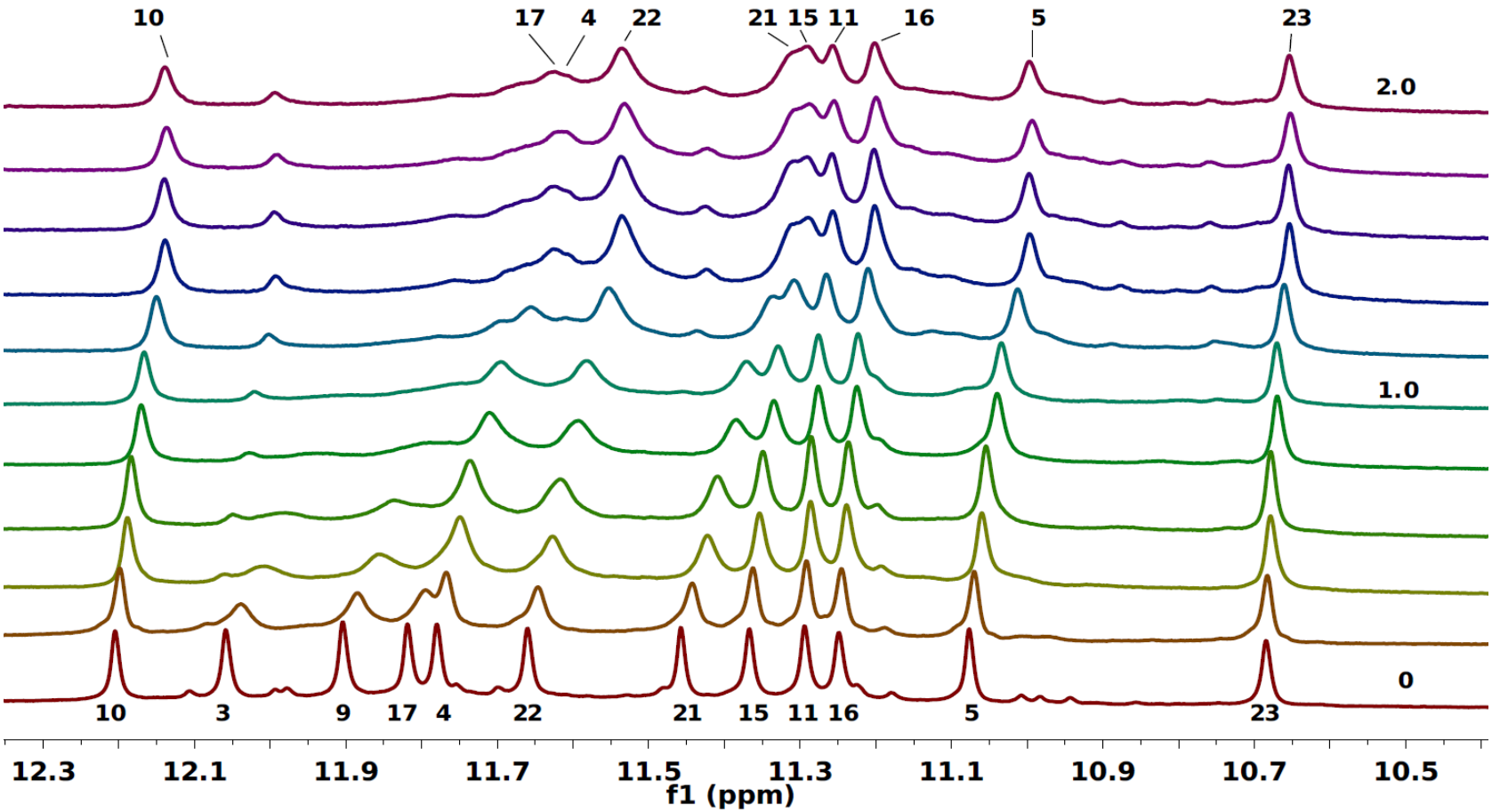


Figure 2.8: Imino proton spectra of Tel24 (hybrid-1) with DB832 at 25 °C in 10 mM K_2HPO_4/80 mM KCl, pH 7.0.

DB832 is titrated in the increments of 0.2 molar equivalents into Tel24 to a final ratio of 2:1 DB832/Tel24. Proton assignments of Tel24 were courteously provided by Dr. Anh T. Phan.

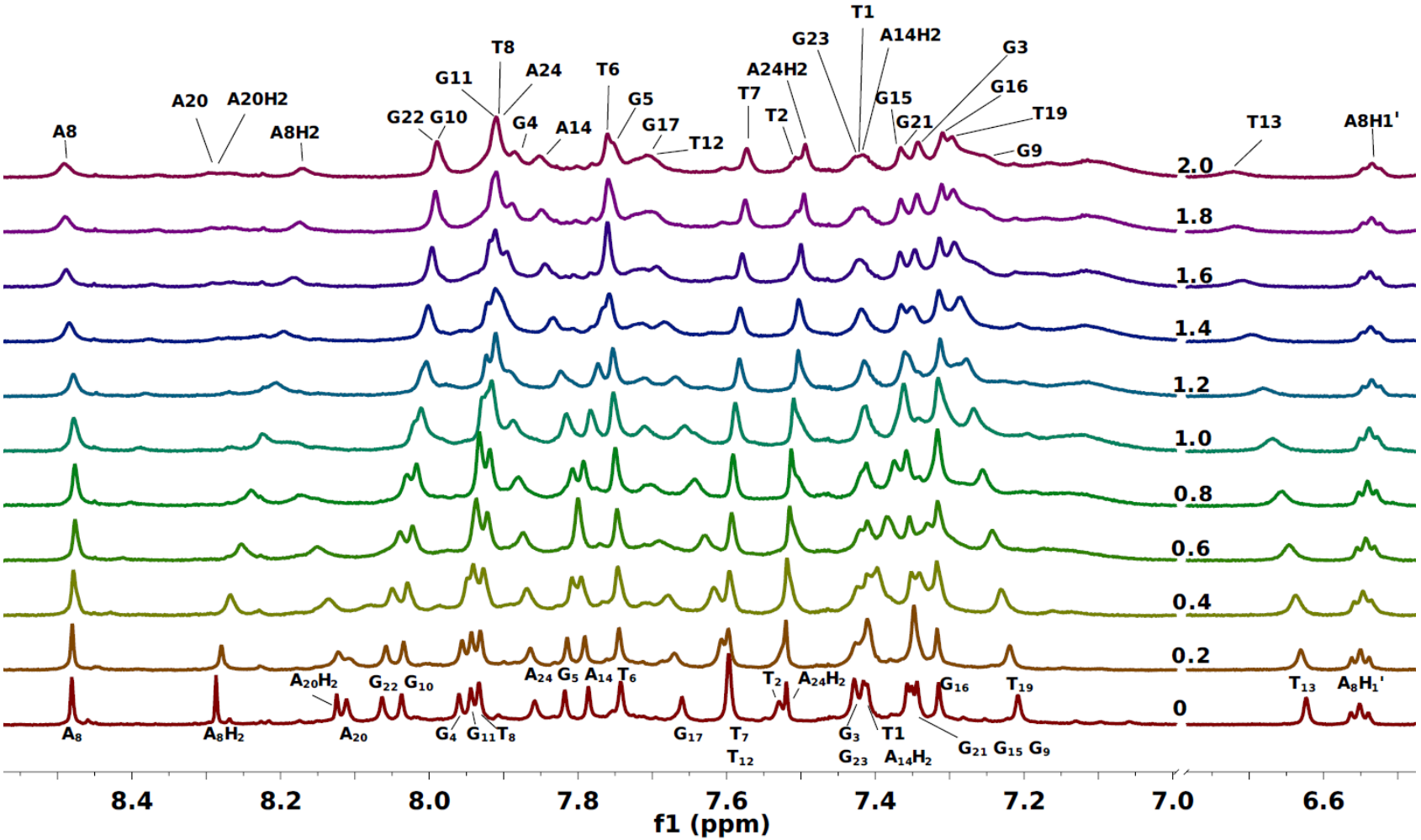


Figure 2.9: Aromatic proton spectra of Tel24 with DB832 at 25 °C in 10 mM K_2HPO_4/80 mM KCl, pH 7.0.

DB832 is titrated in the increments of 0.2 molar equivalents into Tel24 to a final ratio of 2:1 DB832/Tel24. Proton assignments of Tel24 were courteously provided by Dr. Anh T. Phan.

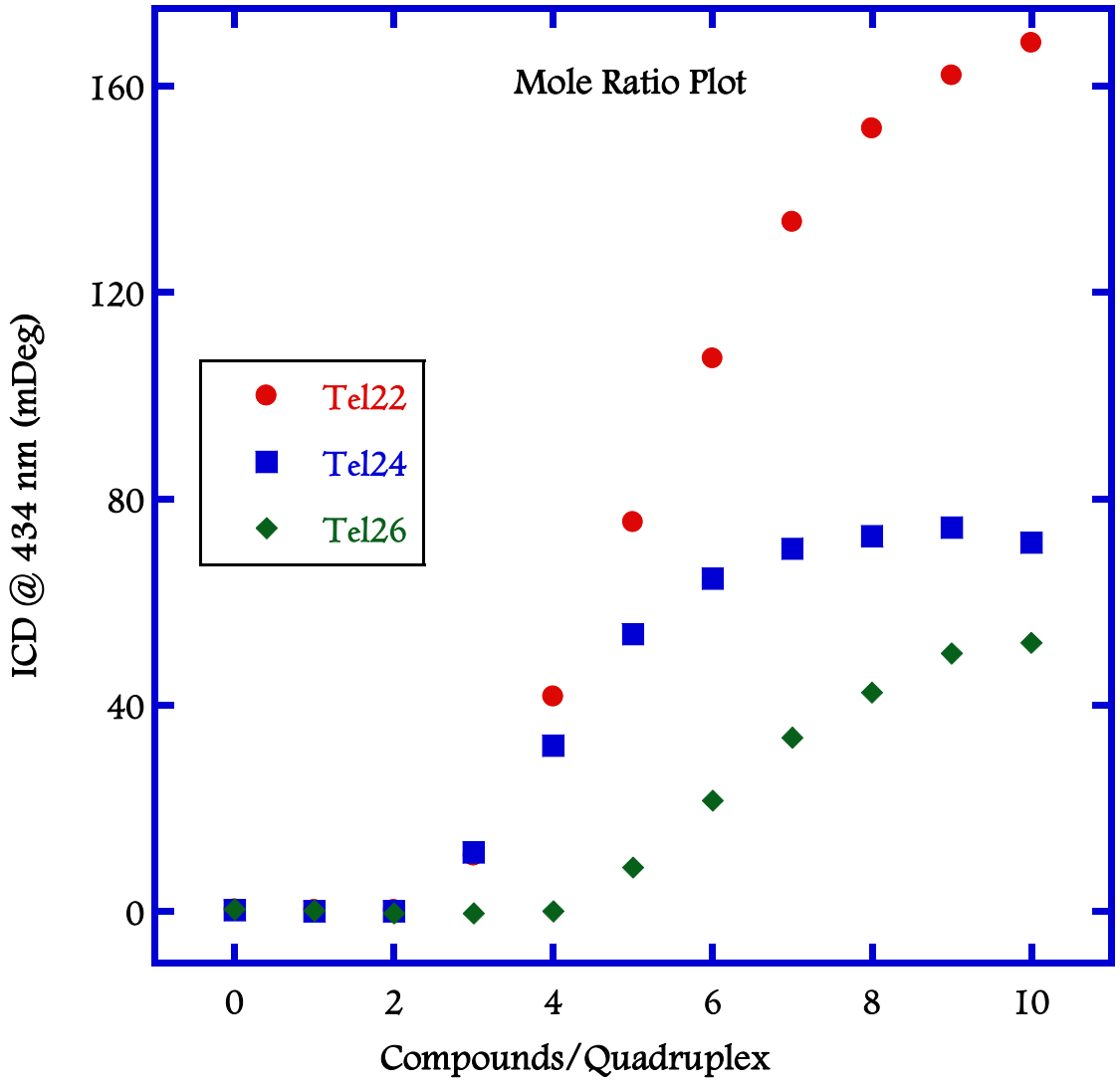


Figure 2.10: Plot of the ICD signal at wavelength corresponding to the maximum absorbance of the bound DB832 (434 nm) as a function of DB832 added molar ratio with different telomeric quadruplex sequences.

From the plot, it is clear that a multiple binding mode is exhibited by DB832. The end-stacking as observed from NMR and CD occurs at the initial ratio, and the subsequent formation of stacked complexes in the grooves occurs at latter ratios (> 2:1 or 3:1)

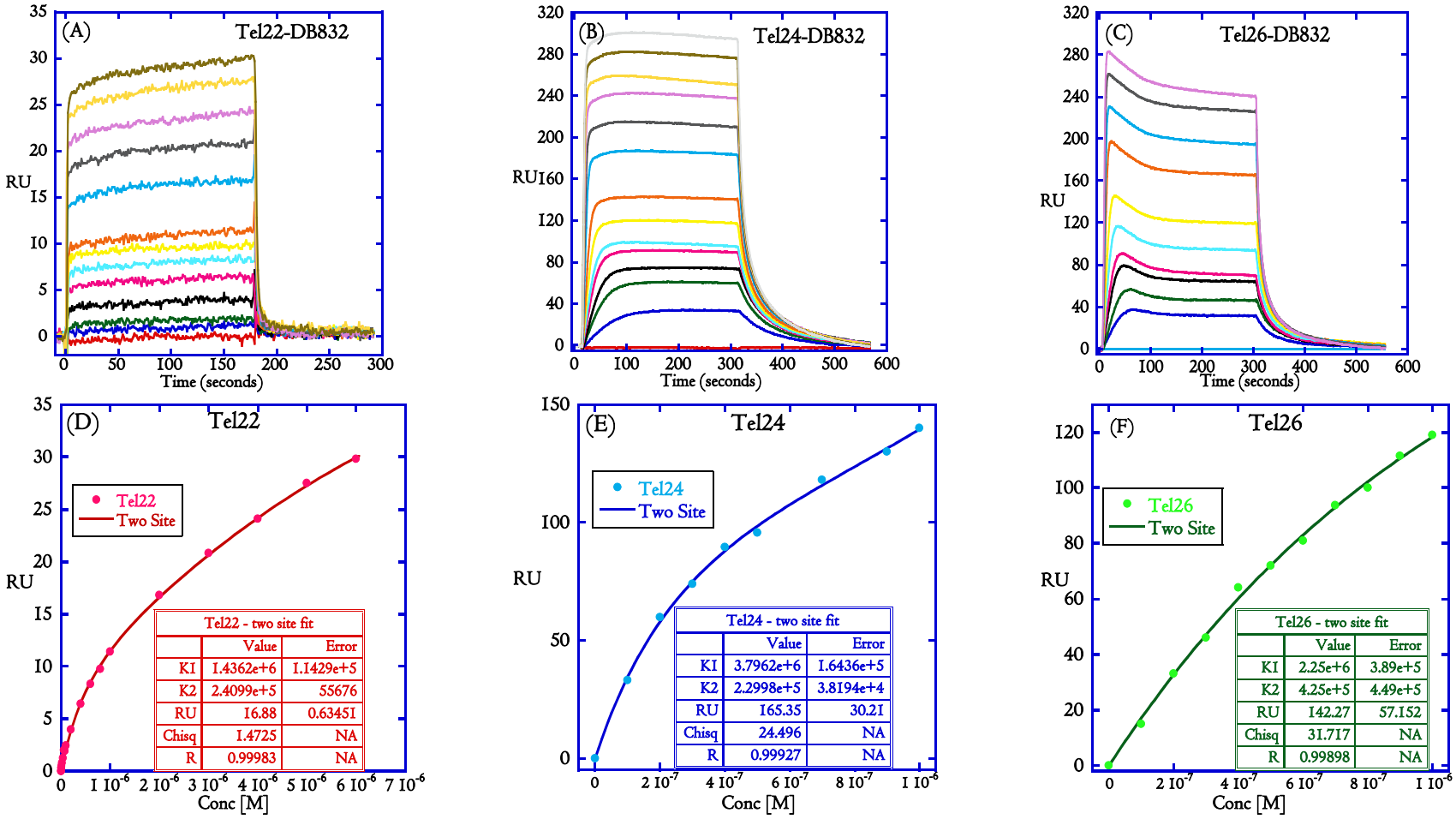


Figure 2.11: SPR sensorgrams for binding of DB832 analogs with human telomeric quadruplex sequences.

(A) Tel22, (B) Tel24, and (C) Tel26 in 10 mM phosphate buffer containing 100 mM K^+ at 25 °C. The quadruplex curves range in ligand concentration from 10 nM for the bottom curve to 10 μM for the top curve. Steady-state binding plots fit to a two-site model (Materials and Methods) and their binding parameters (inset) are listed in D-F. The concentration values are for unbound compound concentration in the flow solution. DB832 exhibits an initial strong binding by stacking at the ends, followed by a weaker stacking in the grooves of these quadruplex systems.

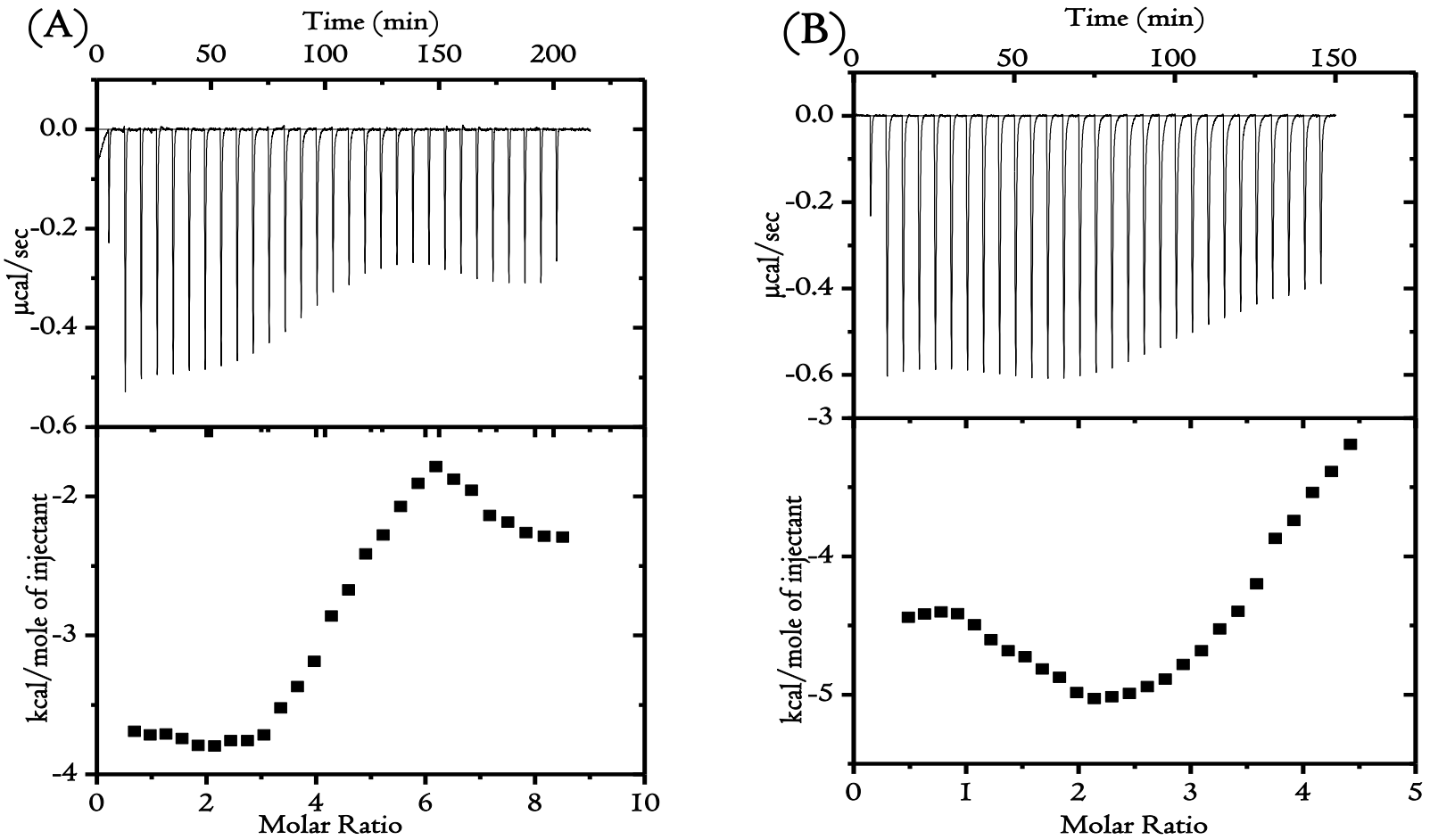


Figure 2.12: Isothermal titration calorimetry plot of DB832 (400 µM) titrated into a 10 µM (A) and 20 µM (B) Tel24 quadruplex sequence in TRIS buffer containing 100 mM K^+.

Plot of heat versus molar ratio (bottom) was obtained by subtracting the integrated peak areas for the blank, buffer titration from the DNA interaction titration. From the ITC plots, it can be clearly seen that multiple binding events are occurring as suggested by CD, NMR and SPR. The binding stoichiometry of DB832 with Tel24 sequence is clearly in the range of 5-6 molecules per quadruplex unit. Accurate thermodynamic and binding parameters could not be obtained due to the high stoichiometry.

2.5 REFERENCES

[1] a)S. Allison, C. Chen, D. Stigter, *Biophys. J.* **2001**, *81*, 2558; b)J. B. Mills, P. J. Hagerman, *Nucl. Acids Res.* **2004**, *32*, 4055.

[2] a)R. D. Wells, *Trends in Biochemical Sciences* **2007**, *32*, 271; b)R. D. Wells, *J. Biol. Chem.* **1988**, *263*, 1095; c)V. Bloomfield, D. Crothers, I. Tinoco, *University Science Books* **2000**.

[3] a)A. T. Phan, V. Kuryavyi, D. J. Patel, *Curr Opin Struct Biol* **2006**, *16*, 288; b)Y. Qin, L. H. Hurley, *Biochimie* **2008**, *90*, 1149; c)S. Burge, G. N. Parkinson, P. Hazel, A. K. Todd, S. Neidle, *Nucleic Acids Res* **2006**, *34*, 5402; d)P. Hazel, J. Huppert, S. Balasubramanian, S. Neidle, *J Am Chem Soc* **2004**, *126*, 16405.

[4] a)J. C. Darnell, K. B. Jensen, P. Jin, V. Brown, S. T. Warren, R. B. Darnell, *Cell* **2001**, *107*, 489; b)M. Fry, L. A. Loeb, *Proc Natl Acad Sci U S A* **1994**, *91*, 4950; c)A. Lew, W. J. Rutter, G. C. Kennedy, *Proc Natl Acad Sci U S A* **2000**, *97*, 12508; d)D. Sun, K. Guo, J. J. Rusche, L. H. Hurley, *Nucleic Acids Res* **2005**, *33*, 6070; e)T. Evans, E. Schon, G. Gora-Maslak, J. Patterson, A. Efstratiadis, *Nucleic Acids Res* **1984**, *12*, 8043; f)R. M. Howell, K. J. Woodford, M. N. Weitzmann, K. Usdin, *J Biol Chem* **1996**, *271*, 5208; g)W. Chai, Q. Du, J. W. Shay, W. E. Wright, *Mol Cell* **2006**, *21*, 427; h)V. L. Makarov, Y. Hirose, J. P. Langmore, *Cell* **1997**, *88*, 657; i)R. McElligott, R. J. Wellinger, *Embo J* **1997**, *16*, 3705; j)R. K. Moyzis, J. M. Buckingham, L. S. Cram, M. Dani, L. L. Deaven, M. D. Jones, J. Meyne, R. L. Ratliff, J. R. Wu, *Proc Natl Acad Sci U S A* **1988**, *85*, 6622.

[5] a)H. Fernando, A. P. Reszka, J. Huppert, S. Ladame, S. Rankin, A. R. Venkitaraman, S. Neidle, S. Balasubramanian, *Biochemistry* **2006**, *45*, 7854; b)J. L. Huppert, S. Balasubramanian, *Nucleic Acids Res* **2005**, *33*, 2908; c)J. L. Huppert, S. Balasubramanian, *Nucleic Acids Res* **2007**, *35*, 406.

[6] M. Gellert, M. N. Lipsett, D. R. Davies, *Proc Natl Acad Sci U S A* **1962**, *48*, 2013.

[7] a)L. H. Hurley, *Nat Rev Cancer* **2002**, *2*, 188; b)J. L. Mergny, J. F. Riou, P. Mailliet, M. P. Teulade-Fichou, E. Gilson, *Nucleic Acids Res* **2002**, *30*, 839; c)S. Neidle, G. Parkinson, *Nat Rev Drug Discov* **2002**, *1*, 383; d)S. Neidle, D. E. Thurston, *Nat Rev Cancer* **2005**, *5*, 285.

[8] a)K. Shin-ya, K. Wierzba, K.-i. Matsuo, T. Ohtani, Y. Yamada, K. Furihata, Y. Hayakawa, H. Seto, *Journal of the American Chemical Society* **2001**, *123*, 1262; b)W. R. Shi DF, Sun D, Hurley LH., *journal of medicinal chemistry* **2001**, *44*, 15; c)E. M. Rezler, D. J. Bearss, L. H. Hurley, *Annu Rev Pharmacol Toxicol* **2003**, *43*, 359.

[9] a)A. Siddiqui-Jain, C. L. Grand, D. J. Bearss, L. H. Hurley, *Proc Natl Acad Sci U S A* **2002**, *99*, 11593; b)C. L. Grand, H. Han, R. M. Muñoz, S. Weitman, D. D. Von Hoff, L. H. Hurley, D. J. Bearss, *Molecular Cancer Therapeutics* **2002**, *1*, 565.

[10] a)O. Tabarrini, V. Cecchetti, A. Fravolini, G. Nocentini, A. Barzi, S. Sabatini, H. Miao, C. Sissi, *journal of medicinal chemistry* **1999**, *42*, 2136; b)J. J. Clement, N. Burres, K.

Jarvis, D. T. W. Chu, J. Swiniarski, J. Alder, *Cancer Research* **1995**, *55*, 830; c)D. T. W. Chu, P. B. Fernandes, R. E. Maleczka, C. W. Nordeen, A. G. Pernet, *journal of medicinal chemistry* **1987**, *30*, 504; d)J. S. Wolfson, D. C. Hooper, *Clin. Microbiol. Rev.* **1989**, *2*, 378.

[11] D. Drygin, A. Siddiqui-Jain, S. O'Brien, M. Schwaebe, A. Lin, J. Bliesath, C. B. Ho, C. Proffitt, K. Trent, J. P. Whitten, J. K. C. Lim, D. Von Hoff, K. Anderes, W. G. Rice, *Cancer Research* **2009**, *69*, 7653.

[12] a)Y. Xu, Y. Noguchi, H. Sugiyama, *Bioorg Med Chem* **2006**, *14*, 5584; b)Y. Wang, D. J. Patel, *Structure* **1993**, *1*, 263; c)A. T. Phan, Y. S. Modi, D. J. Patel, *J Mol Biol* **2004**, *338*, 93; d)A. T. Phan, V. Kuryavyi, K. N. Luu, D. J. Patel, *Nucleic Acids Res* **2007**, *35*, 6517; e)J. Li, J. J. Correia, L. Wang, J. O. Trent, J. B. Chaires, *Nucleic Acids Res* **2005**, *33*, 4649; f)J. Dai, M. Carver, D. Yang, *Biochimie* **2008**, *90*, 1172.

[13] a)Y. Yokoyama, Y. Takahashi, A. Shinohara, Z. Lian, X. Wan, K. Niwa, T. Tamaya, *Cancer Res* **1998**, *58*, 5406; b)D. Sun, L. H. Hurley, *Methods Enzymol* **2001**, *340*, 573.

[14] a)M. Watanabe, S.-k. Yu, M. Sawafuji, M. Kawamura, H. Horinouchi, M. Mukai, K. Kobayashi, *Cancer* **2002**, *94*, 240; b)J. W. Shay, Y. Zou, E. Hiyama, W. E. Wright, *Hum. Mol. Genet.* **2001**, *10*, 677; c)A. G. Bodnar, M. Ouellette, M. Frolkis, S. E. Holt, C.-P. Chiu, G. B. Morin, C. B. Harley, J. W. Shay, S. Lichtsteiner, W. E. Wright, *Science* **1998**, *279*, 349; d)M. A. Blasco, H.-W. Lee, M. P. Hande, E. Samper, P. M. Lansdorp, R. A. DePinho, C. W. Greider, *Cell* **1997**, *91*, 25.

[15] a)S. Taetz, C. Baldes, T. Mürdter, E. Kleideiter, K. Piotrowska, U. Bock, E. Haltner-Ukomadu, J. Mueller, H. Huwer, U. Schaefer, U. Klotz, C. M. Lehr, *Pharmaceutical Research* **2006**, *23*, 1031; b)D. Sun, B. Thompson, B. Cathers, M. Salazar, S. Kerwin, J. Trent, T. Jenkins, S. Neidle, L. Hurley, *J Med Chem* **1997**, *40*, 2113 ; c)C. Sissi, L. Lucatello, A. Paul Krapcho, D. J. Maloney, M. B. Boxer, M. V. Camarasa, G. Pezzoni, E. Menta, M. Palumbo, *Bioorganic & Medicinal Chemistry* **2007**, *15*, 555; d)D. F. Shi, R. T. Wheelhouse, D. Sun, L. H. Hurley, *J Med Chem* **2001**, *44*, 4509; e)P. J. Perry, M. A. Read, R. T. Davies, S. M. Gowan, A. P. Reszka, A. A. Wood, L. R. Kelland, S. Neidle, *journal of medicinal chemistry* **1999**, *42*, 2679.

[16] D. Monchaud, *Organic & biomolecular chemistry* **2008**, *6*, 627.

[17] a)P. J. Perry, S. M. Gowan, A. P. Reszka, P. Polucci, T. C. Jenkins, L. R. Kelland, S. Neidle, *journal of medicinal chemistry* **1998**, *41*, 3253; b)R. Harrison, J. Cuesta, G. Chessari, M. Read, S. Basra, A. Reszka, J. Morrell, S. Gowan, C. Incles, F. Tanious, W. Wilson, L. Kelland, S. Neidle, *J Med Chem* **2003**, *46*, 4463 ; c)M. Gunaratnam, O. Greciano, C. Martins, A. P. Reszka, C. M. Schultes, H. Morjani, J.-F. Riou, S. Neidle, *Biochemical Pharmacology* **2007**, *74*, 679.

[18] a)Y.-T. Fu, B. R. Keppler, J. Soares, M. B. Jarstfer, *Bioorganic & Medicinal Chemistry* **2009**, *17*, 2030; b)A. M. Burger, F. Dai, C. M. Schultes, A. P. Reszka, M. J. Moore, J. A. Double, S. Neidle, *Cancer Research* **2005**, *65*, 1489; c)P. Phatak, J. C. Cookson, F. Dai, V. Smith, R. B. Gartenhaus, M. F. G. Stevens, A. M. Burger, *Br J Cancer* **2007**, *96*, 1223; d)C. Leonetti, S. Amodei, C. D'Angelo, A. Rizzo, B. Benassi, A. Antonelli, R. Elli, M. F. G. Stevens, M. D'Incalci, G. Zupi, A. Biroccio, *Molecular Pharmacology* **2004**, *66*, 1138; e)S. M. Gowan, R. Heald, M. F. G. Stevens, L. R. Kelland, *Molecular Pharmacology* **2001**,

60, 981; f)T. Shalaby, A. O. von Bueren, M.-L. Hürlimann, G. Fiaschetti, D. Castelletti, T. Masayuki, K. Nagasawa, A. Arcaro, I. Jelesarov, K. Shin-ya, M. Grotzer, *Molecular Cancer Therapeutics* **2010**, *9*, 167; g)M.-Y. Kim, H. Vankayalapati, K. Shin-ya, K. Wierzba, L. H. Hurley, *Journal of the American Chemical Society* **2002**, *124*, 2098; h)M.-Y. Kim, M. Gleason-Guzman, E. Izbicka, D. Nishioka, L. H. Hurley, *Cancer Research* **2003**, *63*, 3247.

[19] a)A. T. Phan, Y. S. Modi, D. J. Patel, *J Am Chem Soc* **2004**, *126*, 8710; b)T. S. Dexheimer, D. Sun, L. H. Hurley, *Journal of the American Chemical Society* **2006**, *128*, 5404.

[20] a)L. Martino, A. Virno, B. Pagano, A. Virgilio, S. Di Micco, A. Galeone, C. Giancola, G. Bifulco, L. Mayol, A. Randazzo, *Journal of the American Chemical Society* **2007**, *129*, 16048; b)M. J. Cocco, L. A. Hanakahi, M. D. Huber, N. Maizels, *Nucl. Acids Res.* **2003**, *31*, 2944.

[21] S. Cosconati, L. Marinelli, R. Trotta, A. Virno, S. De Tito, R. Romagnoli, B. Pagano, V. Limongelli, C. Giancola, P. G. Baraldi, L. Mayol, E. Novellino, A. Randazzo, *Journal of the American Chemical Society* **2010**, *132*, 6425.

[22] Q. Chen, I. D. Kuntz, R. H. Shafer, *Proceedings of the National Academy of Sciences of the United States of America* **1996**, *93*, 2635.

[23] a)A. Tomlinson, B. Frezza, M. Kofke, M. Wang, B. A. Armitage, D. Yaron, *Chemical Physics* **2006**, *325*, 36; b)J. L. Seifert, R. E. Connor, S. A. Kushon, M. Wang, B. A. Armitage, *Journal of the American Chemical Society* **1999**, *121*, 2987.

[24] J. Dash, P. S. Shirude, S.-T. D. Hsu, S. Balasubramanian, *Journal of the American Chemical Society* **2008**, *130*, 15950.

[25] E. W. White, F. Tanious, M. A. Ismail, A. P. Reszka, S. Neidle, D. W. Boykin, W. D. Wilson, *Biophysical Chemistry* **2007**, *126*, 140.

[26] a)S. Mazur, F. A. Tanious, D. Ding, A. Kumar, D. W. Boykin, I. J. Simpson, S. Neidle, W. D. Wilson, *Journal of Molecular Biology* **2000**, *300*, 321; b)B. Nguyen, C. Tardy, C. Bailly, P. Colson, C. Houssier, A. Kumar, D. W. Boykin, W. D. Wilson, *Biopolymers* **2002**, *63*, 281; c)L. Wang, C. Bailly, A. Kumar, D. Ding, M. Bajic, D. W. Boykin, W. D. Wilson, *Proceedings of the National Academy of Sciences of the United States of America* **2000**, *97*, 12.

[27] T. M. Davis, W. D. Wilson, *Analytical Biochemistry* **2000**, *284*, 348.

[28] a)B. Nordén, *Applied Spectroscopy Reviews* **1978**, *14*, 157 ; b)R. M. Pagni, *Journal of Chemical Education* **1998**, *75*, 1095.

[29] B. Norden, M. Kubista, T. Kurucsev, *Quarterly Reviews of Biophysics* **1992**, *25*, 51.

[30] a)M. Munde, M. A. Ismail, R. Arafa, P. Peixoto, C. J. Collar, Y. Liu, L. Hu, M.-H. David-Cordonnier, A. Lansiaux, C. Bailly, D. W. Boykin, W. D. Wilson, *Journal of the American Chemical Society* **2007**, *129*, 13732; b)Y. Miao, M. P. H. Lee, G. N.

Parkinson, A. Batista-Parra, M. A. Ismail, S. Neidle, D. W. Boykin, W. D. Wilson, *Biochemistry* **2005**, *44*, 14701.

[31] T. P. Garner, H. E. L. Williams, K. I. Gluszyk, S. Roe, N. J. Oldham, M. F. G. Stevens, J. E. Moses, M. S. Searle, *Organic & biomolecular chemistry* **2009**, *7*.

[32] K. N. Luu, A. T. Phan, V. Kuryavyi, L. Lacroix, D. J. Patel, *J Am Chem Soc* **2006**, *128*, 9963.

[33] a)K. Okamoto, Y. Sannohe, T. Mashimo, H. Sugiyama, M. Terazima, *Bioorg Med Chem* **2008**, *16*, 6873; b)A. Matsugami, Y. Xu, Y. Noguchi, H. Sugiyama, M. Katahira, *Febs J* **2007**, *274*, 3545.

[34] Y. Xu, Y. Noguchi, H. Sugiyama, *Bioorganic & Medicinal Chemistry* **2006**, *14*, 5584.

[35] a)J. Seenisamy, E. M. Rezler, T. J. Powell, D. Tye, V. Gokhale, C. S. Joshi, A. Siddiqui-Jain, L. H. Hurley, *Journal of the American Chemical Society* **2004**, *126*, 8702; b)A. T. Phan, Y. S. Modi, D. J. Patel, *Journal of the American Chemical Society* **2004**, *126*, 8710.

[36] M. L. Kopka, C. Yoon, D. Goodsell, P. Pjura, R. E. Dickerson, *Proceedings of the National Academy of Sciences of the United States of America* **1985**, *82*, 1376.

[37] D. Monchaud, M.-P. Teulade-Fichou*, Vol. 608*, **2008**, pp. 257.

[38] D. Monchaud, C. Allain, M.-P. Teulade-Fichou, *Bioorganic & medicinal chemistry letters* **2006**, *16*, 4842.

[39] D. Monchaud, C. Allain, M. P. Teulade-Fichou, *Nucleosides Nucleotides Nucleic Acids* **2007**, *26*, 1585.

[40] A. T. Phan, V. Kuryavyi, H. Y. Gaw, D. J. Patel, *Nat Chem Biol* **2005**, *1*, 167.

[41] Ö. P. Çetinkol, A. E. Engelhart, R. K. Nanjunda, W. D. Wilson, N. V. Hud, *ChemBioChem* **2008**, *9*, 1889.

[42] G. Scatchard, *Annals of the New York Academy of Sciences* **1949**, *51*, 660.

3 SELECTIVE RECOGNITION OF MIXED PARALLEL/ANTIPARALLEL HYBRID QUADRUPLEX DNA BY DB832 AS A STACKED SPECIES

3.1 INTRODUCTION

G-quadruplex DNA is comprised of a series of stacked guanine tetrads held together in a coplanar cyclic array by Hoogsteen and Watson-Crick hydrogen bonds. The quadruplex structure is stabilized through hydrogen bonds, π-π stacking interactions of the stacked tetrads as well as by coordination with cations located between or within the tetrads. G-Quadruplex DNA is highly polymorphic and can form different structures depending on sequence, length, concentration, nature of cations present, presence of crowding agents and other factors. Potential quadruplex-forming sequences are common in the genome. Telomeric DNA [1], immunoglobulin switch region sequences [2], the fragile-X repeat sequence [3], promoter regions of some oncogenes [4], the insulin gene[5], and other genomic sequences have been shown to form quadruplex structures *in vitro* [1b, 4a, 6]. The extensive variation among quadruplex structures, combined with the fact that some small molecules can induce or lock a particular conformation, may allow the design of compounds that can target particular quadruplex structure with a high degree of selectivity. Selectivity of a compound for quadruplex-DNA over duplex-DNA is also important because nonselective duplex-binding is associated with cytotoxicity and can also result in significant loss of valuable drug. Moreover, selectivity for a certain quadruplex structure over other non-targeted quadruplexes is also a very important consideration in design of quadruplex-interactive agents. Binding to non-targeted quadruplex sequences results in compound loss and may have unintentional effects on regulation of non-targeted genes. Thus, there is a need for new classes of compounds that can selectively discriminate between different quadruplex conformations in order to achieve full therapeutic response and minimal side effects.

Telomeres are guanine-rich regions of non-coding DNA located at the ends of eukaryotic chromosomes. They serve a variety of functions, from protecting the ends of chromosomes from erosion and end-end fusion leading to recombination [7]. In some organisms they play a significant role in gene expression [8]. Telomeric DNA sequences have been shown to fold into four-stranded G-quadruplex structures *in vitro* under physiological conditions [9]. Recent studies suggest that telomeres may also form G-quadruplex structures *in vivo* [10]. Human telomeric sequence is one of the most extensively studied motifs due to its critical role in maintaining chromosomal integrity, and also its potential implication as a critical target for anticancer therapy [11]. The hexanucleotide repeats of TTAGGG can fold into an array of quadruplex topologies *in vitro* under different physiological conditions. Structural information of telomeric quadruplex DNA under *in vivo* conditions is imperative from the drug design standpoint of view.

The earliest solution structure of the human telomeric DNA sequence, $d[AG_3(T_2AG_3)_3]$ solved by NMR in Na^+ shows that this sequence folds into an intramolecular quadruplex conformation termed as an antiparallel basket topology [6h] (Figure 3.1-E). The term "antiparallel" refers to the directionality of a strand with respect to its adjacent connecting strand. Therefore, in the antiparallel topology, the adjacent connecting strands are antiparallel to each other, and the connecting loop bases adopt a "basket-type" conformation. However, since the intracellular K^+ concentration is much higher than Na^+, several attempts have been made to elucidate the folding topology of human telomeric quadruplex in K^+. The first human telomeric quadruplex structure in K^+ by crystallography studies have shown that this sequence folds into a parallel, propeller-type topology [12] (Figure 3.1-D). In a parallel topology, all the G-tracts run parallel to each other, while the connecting residues form a double-chain-reversal loop, resulting in a propeller-type structure. Recent NMR structural studies in K^+ by several groups have showed that the human telomeric quadruplex folds into a mixture of

mixed parallel/antiparallel quadruplex conformations termed hybrid-1 and hybrid-2 structures [13] (Figure 3.1-B, C). Interestingly, a third hybrid structure was also recently found by NMR for the human telomeric sequence [14], adding to the repertoire of conformations this sequence can fold into. These "hybrid-type" structures of human telomeric quadruplexes under physiological K^+ concentrations maybe the predominant conformation *in vivo*.

With the evidence for *in vivo* existence of telomeric quadruplexes becoming more compelling [15], their possible function or functions must be addressed. Small molecules that can bind to and stabilize the quadruplex conformation of telomeres perturb their functions [16]. Telomerase is a reverse-transcriptase enzyme that is expressed in rapidly dividing cells, including cancer cells and eukaryotic parasites [17]. The enzyme exerts its action by adding TTAGGG repeats to the end of the telomere, preventing telomere shortening and allowing the cells to replicate indefinitely, a hallmark of cancer cells. The development of quadruplex-interactive small molecules that can inhibit telomerase is therefore a current area of interest in anticancer and antiparasitic drug design. Small molecules that can selectively bind to the various human telomeric quadruplex conformations or induce a conformational switch from one conformer to the other have significant potential therapeutic impact.

Apart from the telomeres, potential quadruplex forming G-rich sequences are also highly prevalent throughout the genome, particularly in promoter sequences of oncogenes [4a, 6c, 6f, 18]. The 27-nt G-rich sequence in the NHE-III promoter element of c-*myc* oncogene has been shown to exist as multiple quadruplex species [19]. Structural studies on shorter sequences from within the 27-nt G-rich sequence have been shown to form a single, very stable intramolecular parallel-stranded quadruplex conformation in K^+ solution, with the G-tracts joined through propeller loops, and with all guanines involved in the G-tetrads adopting an *anti* conformation [6f, 18a] (Figure 3.1-D). The parallel propeller-type conformation adopted by

c-myc has structural features very similar to the conformation of human telomeric quadruplex solved by crystallography (Figure 3.1-D). The 39-nt G-rich strand in the promoter region of *bcl-2* oncogene has also been shown to form three distinct intramolecular G-quadruplexes in K^+ solution [4c, 6b]. Detailed NMR studies of a dual mutant, 23-nt *bcl-2* (G15T/G16T), showed the formation of a unique intramolecular G-quadruplex structure with mixed parallel/antiparallel G-strands [6a] (Figure 3.1-C). The *bcl-2* quadruplex scaffold contains three G-tetrads with the G-tracts connected by a single-nucleotide double-chain-reversal side loop and two lateral loops, respectively. The first three-nucleotide CGC loop in the *bcl-2* promoter sequence forms a lateral loop, as opposed to a double-chain-reversal side loop observed in a similar sequence in the *c-myc* promoter, and appears to largely determine the overall folding of the *bcl-2* quadruplex conformation. The overall fold of *bcl-2* resembles the hybrid-2 fold adopted by the human telomere (Figure 3.1-C), however, subtle differences in the connecting loops bases makes the two hybrid-2 folds very distinct. Small molecules that can bind and stabilize these quadruplex conformations or, more importantly, discriminate between two very similar structures such as the hybrid-2 folds can have significant therapeutic impact.

In Chapter 2, DB832 (Figure 3.1-A), a heterocyclic diamidine, was shown to selectively bind to human telomeric quadruplex conformation with high affinity over duplex DNA. The multitude of conformations that G-rich sequences could fold into led to elucidate the structure-specific selectivity of DB832 for the different quadruplex architectures throughout the genome. Circular Dichroism (CD) and NMR techniques were used to probe these interactions. The selectivity within the different quadruplex folds exhibited by human telomeric sequences was also investigated. CD studies show that DB832 is highly selective for different human telomeric quadruplex conformations and potentially interacts as a stacked species in the grooves. DB832 binding to non-telomeric sequences such as *c-myc*, *bcl-2*, TBA

and intermolecular quadruplex sequences are non-specific in nature, and most likely interact as monomers by stacking at their terminal tetrads. NMR studies with the wild-type and modified telomeric sequences further suggest that DB832 can bind to a selective quadruplex conformation from a mixture of structures.

3.2 MATERIALS AND METHODS

3.2.1 Sample Preparation

The oligonucleotides Tel22, d[$AG_3(T_2AG_3)_3$]; Tel26, d[$A_3(G_3T_2A)_3G_3A_2$]; wtTel26, d[$T_2G_3(T_2AG_3)_3A$]; TBA, d[$G_2T_2G_2TGTG_2T_2G_2$]; Aptamer, d[$(G_2T_4)_3G_2$]; Tetrahymena d[$T_2G_4]_4$; Oxytricha, d[$G_4(T_4G_4)_3$]; *bcl-2*, d[$G_3CGCG_3AG_2A_2T_2G_3CG_3$]; *c-myc27*, d[$TG_4AG_3TG_4AG_3TG_4AAG_2$]; U6U7, d[TAGGGUUAGGGT]; intramolecular d[TG_4T]; AATT, d[GCGAATTCGC]; mixed duplex d[CGAGATCAAAAGATCTCG], and GC, d[$GC]_7$ were purchased with HPLC purification from Midland Certified Reagent Company or Integrated DNA Technologies, Inc. The G-quadruplex DNA samples were dissolved in buffer to the desired concentrations, heated to 85 ^{0}C and cooled slowly to insure the folding of the quadruplexes prior to each experiment. The concentration of each DNA sample was determined spectrophotometrically at 260 nm using the nearest neighbor extinction coefficient at 80 °C and extrapolated to 25 °C. A stock solution containing 1 mM of each compound was prepared in double distilled water and diluted to working concentrations immediately before use. The synthesis of DB832 will be described elsewhere.

3.2.2 Thermal Denaturation Studies

Thermal denaturation studies were conducted on a Cary Varian 300 BIO UV-visible spectrophotometer in quartz cells with a 1 cm pathlength. A phosphate buffer containing 80 mM KCl, 10 mM K_2HPO_4, and 0.1 mM EDTA was used. A thermistor fixed into a reference

cuvette was used to monitor the temperature. The absorbance of the quadruplex sequences was monitored at 295 nm, while the duplex sequence was monitored at 260 nm. Melting curves were obtained for each DNA sequence in the presence and absence of DB832. For each oligomer the DNA concentration was 3 x 10^{-6} M and the DB832: DNA ratio was 2:1. Data manipulation and plotting were performed using the program Kaleidagraph version 3.6.

3.2.3 Circular Dichroism (CD) Studies

CD measurements were performed at 25 °C in a 10 mM HEPES buffer (pH 7.4) containing 3 mM EDTA and 50 mM KCl, or NaCl. For CD measurements obtained in the presence of 50 mM LiCl, a 10 mM TRIS buffer containing 3 mM EDTA acid and pH adjusted to 7.4 were used. CD spectra were recorded using a Jasco J-810 spectrapolarimeter in a 1-cm cell using a scanning speed of 50 nm/min with a response time of 1 s. The spectra were averaged over four scans. A buffer baseline scan was collected in the same cuvette and subtracted from the average scan for each sample. Appropriate amounts of stock solution of compound were added sequentially to increase the molar ratio. Data manipulation and plotting was performed using the program Kaleidagraph version 3.6.

3.2.4 Nuclear Magnetic Resonance Studies

Tel22, TG_4T, and U6U7 samples were prepared in phosphate buffer containing 50 mM KCl, 10 mM K_2HPO_4 and 0.1 mM EDTA. Tel26, wtTel26 and *bcl-2* samples were prepared in phosphate buffer containing 80 mM KCl, 10 mM K_2HPO_4 and 0.1 mM EDTA and 0.01 mM DSS as an internal reference. Quadruplex DNA concentrations were 0.1-0.2 mM for 1D experiments and 0.5 mM for 2D experiments in 90% H_2O:10% D_2O (Cambridge Isotope Laboratories, Inc.). The final DNA samples were adjusted to pH 7.0 using 1M HCl or 1M KOH solutions and were heated past their transition temperature and annealed to room temperature several times before collecting the spectra. Experiments were performed on a Varian Unity 600 spectrometer equipped with a triple resonance broadband probe. DB832 was titrated with

quadruplex DNA with DB832:DNA ratios varying from 0.5 to 2 for U6U7 sequence and ratios from 1 to 4 for Tel22, Tel26, wtTel26, *bcl-2* and TG_4T sequences. Temperature-dependent 1D spectra were recorded at 35 °C using jump-return and WATERGATE methods for solvent suppression [20]. Homonuclear TOCSY on U6U7 was performed at 35 °C using WATERGATE solvent suppression method with a mixing time of 60 ms [21]. All NMR data were processed and analyzed with a combination of VNMR (Varian Inc.), SPARKY (UCSF) and MestreC (Mestrelab Research) software.

3.3 RESULTS AND DISCUSSION

3.3.1 DB832 is Highly Selective for Quadruplex DNA over Duplex DNA

In order to determine the selectivity of DB832 for human telomeric DNA over duplex DNA, melting curves were obtained for the intramolecular human telomeric sequence Tel22, as well as a random hairpin duplex d(CGAGATCAAAAGATCTCG) (loop underlined) that has a similar Tm value. The melting temperature of the telomeric DNA in the absence of DB832 is 53.6 °C (Figure 3.2-B). Upon addition of DB832 at 2:1, the melting temperature of the DNA increases by approximately 9 °C, indicating that the compound is highly stabilizing the quadruplex conformation of the DNA. This is comparable to the ΔTm for a wide range of small molecules known to stabilize the human telomeric quadruplex conformation under similar conditions [22]. In contrast, when DB832 is added to the duplex sequence at the same ratio and conditions, there is an insignificant change in Tm (Figure 3.2-A). This indicates that DB832 binds much more strongly to quadruplex than to duplex DNAs, making it ideal for further study as a highly selective G-quadruplex interactive compound. CD spectra were also obtained with two duplex sequences: an AT-rich sequence, d(GCGAATTCGC), and a GC-rich sequence, $d(GC)_7$ (Figure 3.2 C, D respectively). AT-rich duplexes have been shown to have narrower grooves than GC-rich sequences [23]. The wider grooves of GC can mimic the different groove

dimensions of telomeric quadruplex. In both cases, only a small induced signal was observed upon titration of DB832, and the small with no exciton-type splitting suggests that binding as an extensively stacked species is not occurring.

3.3.2 Effect of Salt on DB832 Binding to Human Telomeric Quadruplex DNA

The CD spectrum for the human telomeric sequence has previously been shown to differ in the presence of sodium ions versus potassium ions [24]. This suggests that this DNA sequence exhibits a different conformation in the presence of each of these counterions. Differing NMR and crystal structures which were obtained in the presence of sodium ions and potassium ions, respectively, support this idea [12, 14, 25]. The NMR structure of $d[AG_3(T_2AG_3)_3]$ in the presence of sodium is an antiparallel basket-like structure [6h]. The crystal structure of this sequence obtained in the presence of potassium ions consists of a parallel, propeller-like structure [12]. Other structures have been proposed to exist in solution in the presence of potassium [26]. Lithium has been shown to have a destabilizing effect on quadruplex structure [27]. To investigate the effect of DNA structure on the binding of DB832, CD titrations were performed (by Dr. Elizabeth White in Dr. Wilson's laboratory) with Tel22 under different salt conditions to determine if DB832 would bind as stacked species to different conformations of this DNA sequence. Figures 3.3 A-C shows the CD spectra for Tel22 titrated with DB832 in the presence of 50 mM potassium, 50 mM sodium, and 50 mM lithium, respectively. Regardless of the cation present, upon addition of DB832 exciton splitting occurs in the induced wavelength region of the CD spectra and peaks occur in the DNA region at 265 and 295 nm. The final CD spectra in different salt conditions had very similar parallel/antiparallel characteristics (Figure 3.3 D-F). This suggests that regardless of the starting conformation of the DNA, DB832 induces the formation of a specific conformation of DNA to form an optimum binding site for stacked species.

The large exciton-type ICD signal observed in Figure 3.3 can be compared to that obtained with a demonstrated end-stacking compound, RHPS4 (Chapter 2, Figure 2.3-C), and a cyanine dye (Chapter 2, Figure 2.3-A) that is of the groove-binding type. Therefore, DB832 is binding as a stacked species in the grooves of the telomeric quadruplex conformation regardless of the type of the cation present.

3.3.3 DB832 Binds Selectively as Stacked Species to the Hybrid Conformation of the Human Telomere

Formation of hybrid quadruplex conformations is highly dependent on the flanking bases of the G-rich sequences [28]. To investigate the effect of flanking bases on the binding of DB832, CD titrations were performed with Tel22, Tel26 and wtTel26 conformations to determine if DB832 would bind as a stacked species to different conformations of this DNA sequence. Tel26, composed of d(AA$\underline{A_3}$GGGTT$\underline{A_9}$GGGTTAGGGTT$\underline{A_{21}}$GGGAA), adopts a hybrid-1 intramolecular G-quadruplex (Figure 3.1 B) consisting of three G-tetrads linked with mixed parallel/antiparallel G-strands in K^+ solution [28b]. The three TTA loop segments sequentially adopt diagonal, lateral, and lateral loop conformations. The three highlighted adenines form a stable triple capping structure on top of the 5' terminal tetrad of the motif further stabilizing this conformation. wtTel26, composed of d(TTAGGGT$\underline{T_8A_9}$GGGTTAGGGTTAGGG$\underline{T_{25}}$T); adopts a hybrid-2 intramolecular G-quadruplex (Figure 3.1 C) consisting of three G-tetrads linked with mixed parallel/antiparallel G-strands in K^+ solution [13c]; however, this hybrid-type structure is different from the hybrid-1 structure in loop arrangement. The three TTA loop segments sequentially adopt lateral, lateral, and diagonal orientations. The three highlighted residues form a very well defined T•A•T triple platform at the 3' terminal tetrad of the motif. Tel22, composed of d(AGGGTTAGGGTTAGGGTTAGGG) as previously described, exists as multiple conformers in K^+.

CD studies were performed to evaluate the effect of flanking bases on the binding of DB832. The CD spectrum of Tel26 sequence (Figure 3.4 B) has peaks at both 260 and 295 nm, suggesting the DNA has a hybrid structure, which is consistent with the NMR data. Upon addition of DB832, the CD signal at both peaks initially increases slightly (up to a 2:1 compound:DNA ratio), then decreases, resulting in a final conformation similar in shape to the initial conformation, indicating that DB832 does not cause significant rearrangement of the DNA conformation upon binding. This behavior is also commonly observed with the titration of DB832 into the wild type human telomeric sequence, Tel22 (Figure 3.4 A). In the wavelength region of DB832 absorbance induced exciton splitting is observed, indicating that the compound is binding to the Tel26 sequence as a stacked species.

The CD spectrum of wtTel26 sequence (Figure 3.4 C) also exhibits peaks at both 260 and 295 nm, suggesting that this DNA is folding into a hybrid structure. However, based on CD, it cannot be firmly established the particular type of the hybrid conformation the sequence folds into since, as aforementioned, this sequence folds into a hybrid-2 conformation. Upon titrating DB832, the CD signal at both peaks also exhibit a behavior similar to Tel26, indicating that DB832 does not affect the DNA conformation upon binding. Also, DB832 exhibits a similar ICD signal indicating the same preferential binding towards this sequence. However, a significantly larger concentration of DB832 (~ 6:1 ratio) is required to observe the induced exciton splitting in this case. This may be due to the requirement of additional DB832 molecules probably to induce a conformational switch from hybrid-2 to hybrid-1. However, a conformational switch between two very similar hybrid structures with very similar energy profiles is rather difficult to confirm based only on CD pattern analysis.

A CD titration was also performed with the human telomere in potassium, in a buffer containing polyethylene glycol (PEG 400). PEG is a crowding agent that reduces accessible

volume. It has been used in *in vitro* experiments to mimic crowded intracellular conditions where biomacromolecules can take up as much as 40% of the cellular volume [29]. PEG has been shown *in vitro* to induce conformational changes in quadruplex DNA by driving the DNA to a conformation with reduced excluded volume [30]. Oxytricha telomeric DNA has been shown to transition from an antiparallel to a parallel conformation upon addition of PEG [31]. The human telomeric sequence has been shown to adopt a parallel conformation in the presence of PEG 400, even in the presence of potassium [30a] (Figure 3.1-D). The crystal structure also confirms the parallel topology adopted by the human telomeric sequence, further suggesting that the crystal packing effect closely mimics the molecular crowding conditions. When PEG 400 is added to human telomeric DNA in the presence of potassium (Figure 3.4 D), the CD spectrum has a dominant peak at 260 nm, consistent with the presence of a parallel structure, as expected. When DB832 is added, the CD signal does not change in either the DNA region or the induced region, further confirming the selectivity for the quadruplex conformation formed by the human telomere in the presence of potassium and in the absence of PEG.

3.3.4 Selectivity of DB832 to Diverse Quadruplex-Forming Sequences

To investigate the selectivity of DB832 for the different quadruplex forming sequences, CD studies of DB832 were performed with telomeric DNA sequences from different organisms that are shown to fold into a variety of structures, including antiparallel chair and basket quadruplexes, and also an intermolecular quadruplex sequence (Figure 3.1). Quadruplex forming sequences from oncogene promoter regions of *c-myc* and *bcl-2*, and G-rich aptamers capable of forming stable G-quadruplexes were also evaluated to gain insight into selectivity of DB832.

The tetrahymena telomeric sequence, $d[T_2G_4]_4$, has also been shown by NMR to form a mixed parallel/antiparallel quadruplex structure with loops oriented lateral/lateral/diagonal from 5' to 3' [6i] (Figure 3.1-C). Before the addition of DB832, the spectrum in the DNA wavelength region has peaks at both 260 and 295 nm, consistent with the formation of a mixed parallel/antiparallel structure (Figure 3.5-A). Upon addition of DB832, the peak at 260 nm shifts to a higher wavelength and has a large decrease in magnitude. The presence of a shoulder peak at 269 nm suggests that the final structure is still a mixed hybrid conformation. An ICD signal is seen in the wavelength region of DB832 absorbance. However, it is much smaller in magnitude than the signal seen when DB832 is titrated with either Tel22 or Tel26. DB832 may be inducing a conformation change of the tetrahymena telomere from a 5'-lateral/lateral/diagonal-3' hybrid to a 5'-diagonal/lateral/lateral-3' hybrid quadruplex. The small ICD signal seen with this sequence may be due to DB832 binding as an end-stacking species to the rearranged conformation.

The dual mutant *bcl2*-G15T/G16T, $d(G_3CGCG_3AG_2A_2T_2G_3CG_3)$, in the promoter region of the *bcl-2* proto-oncogene sequence has been shown to fold into a stable mixed parallel/antiparallel hybrid-2 quadruplex conformation [32] (Figure 3.1-C). The loops of the *bcl2*-G15T/G16T sequence are oriented lateral/lateral/diagonal from 5' to 3'. Before the addition of DB832, the spectrum in the DNA wavelength region has peaks at both 260 and 295 nm (Figure 3.5-C) suggesting the DNA has a mixed parallel/antiparallel structure, which is consistent with the NMR structure and previously published CD data [33]. The addition of DB832 does not significantly change the shape of the spectra in the DNA region, indicating that the compound does not significantly change the conformation of the DNA upon binding. In the wavelength region of DB832 absorbance, a very weak induced CD signal is observed and the absence of exciton-type splitting indicates that the compound is likely binding as a weak, end-stacked monomer. Proton NMR studies of the *bcl-2* sequence with DB832 are

shown in Figure 3.8-A. Significant line broadening is observed for all the guanine imino protons upon titration of DB832, suggesting the binding of DB832 to this sequence is non-specific in nature. The 5'-lateral/lateral/diagonal-3' loop orientation of this sequence is clearly not the favored conformation for DB832 binding.

Oxytricha telomeric sequence, $d[G_4(T_4G_4)_3]$, is shown to fold into an antiparallel-basket quadruplex conformation in Na^+ [34] (Figure 3.1-E); whereas, it exists as a mixed parallel/antiparallel conformer in K^+ [35]. CD studies of this sequence, in the absence of DB832, reveal a major peak at 290 nm and a very small shoulder at 260 nm, suggesting a more antiparallel character of this sequence (Figure 3.5-B). Surprisingly, a weak exciton-type ICD signal is observed suggesting that DB832 might be binding as a stacked species to this sequence. The magnitude of the ICD signal observed can be classified as weak when compared to the ICD of DB832 with Tel22. Upon titration, DB832 does not induce any change in the DNA conformation of this sequence as indicated by the absence of any changes in peak shapes in the DNA absorbance region, which suggests that the preformed mixed parallel/antiparallel quadruplex conformation is the preferred fold for the DB832. Tel22, which is shown to fold into an antiparallel basket conformation in Na^+, undergoes major conformational changes upon titration of DB832 and results in a final structure that has both parallel/antiparallel characteristics as observed in K^+. The fact that DB832 does not induce conformational switch in the oxytricha telomeric sequence, unlike Tel22 in Na^+, but still exhibits a weak exciton-type ICD signal, highlights the subtle differences that the ligands can identify between very similar conformers. In this case the antiparallel-basket makes it important to thoroughly understand the smallest differences that exist in a DNA microenvironment in order to develop promising therapeutic candidates.

Even though the above discussed three sequences form mixed parallel/antiparallel hybrid quadruplex structures, they are not identical to the hybrid conformation formed by the human telomeric sequences Tel22 and Tel26. The three loops of the Tel22 and Tel26 sequences are oriented diagonal/lateral/lateral as read from the 5' to 3' direction. The loops of the tetrahymena, bcl2, and oxytricha sequences are oriented lateral/lateral/diagonal from 5' to 3'. These three sequences also have a different pattern of glycosidic orientation of the sugars than Tel22 and Tel26. The relative difference in magnitude of the CD signals at 260 nm vs. 295 nm for tetrahymena, *bcl-2* and oxytricha, as compared to Tel22 and Tel26, may be indicative of this difference in structure. If this is the case, the fact that the addition of DB832 to tetrahymena results in a smaller signal at 260 nm, as compared to 295 nm, suggests that DB832 may be inducing a conformation change of tetrahymena from a 5'-lateral/lateral/diagonal-3' hybrid to a 5'-diagonal/lateral/lateral-3' hybrid quadruplex. The small amount of exciton splitting seen with this sequence may be due to DB832 binding as a stacked species to the rearranged conformation.

Aptamers have proven to be therapeutically significant due to their high stability, the ease of synthesis and several clinical applications [36]. In fact, a G-rich aptamer, AS1411, which is shown to fold into diverse quadruplex conformers, is currently in clinical trials as a potential anticancer agent [37]. Small molecules that can selectively target aptamers can also have potential therapeutic applications. Thrombin Binding Aptamer, $d(G_2T_2G_2TGTG_2T_2G_2)$, forms an antiparallel basket-type quadruplex as determined by NMR [6g] (Figure 3.1-E). The G-rich oligonucleotide, $d[(G_2T_4)_3G_2]$, has been characterized as an antiparallel chair-type quadruplex by 1-D NMR [38] (Figure 3.1-F). CD studies were performed to determine the interaction of DB832 with the TBA and G-rich oligonucleotide. When DB832 is titrated into this DNA no exciton splitting similar to that with human telomeric DNA occurs, and the small ICD signal is observed with TBA is characteristic of an end-stacking mode (Figure 3.6 A). Before the

addition of DB832, the spectrum in the DNA wavelength region of TBA has a maximum at 295 nm, with a minimum at 260 nm, confirming the presence of the antiparallel-chair conformation. The addition of DB832 increases the signal at 260 nm and decreases the signal at 295 nm, indicating a conformation change in the DNA as the compound binds. CD of the G-rich oligonucleotide, d[$(G_2T_4)_3G_2$], confirms a predominant antiparallel-chair topology of this sequence (Figure 3.6 B). The addition of DB832 does not cause a significant change in the shape of the spectra in the DNA region, indicating that DB832 does not significantly rearrange the conformation of the DNA. No induced exciton splitting is observed in the wavelength region of DB832 absorbance, indicating that binding is occurring as a monomer and not a stacked species.

The G-rich strand of the promoter region of the *c-myc* oncogene, d[$TG_4AG_3TG_4AG_3TG_4AAG_2$], forms a parallel propeller-type quadruplex in solution [39] (Figure 3.1-D). Before the addition of DB832, the spectrum in the DNA wavelength region has a maximum at 260 nm, consistent with a parallel conformation of DNA (Figure 3.6 C). The addition of DB832 causes a decrease in signal at 260 nm, but still results in a positive peak at that wavelength, indicating that DB832 does not significantly rearrange the conformation of the DNA. No induced exciton splitting is observed in the wavelength region of DB832 absorbance, indicating that binding is occurring as a monomer and not a stacked species. d(TG_4T) forms a simple four-stranded intermolecular quadruplex with easily accessible grooves and exposed terminal tetrads [40] (Figure 3.1-G). The addition of DB832 does not change the CD spectrum of this DNA, suggesting that this compound does not bind to intermolecular quadruplexes (Figure 3.6 D). Intermolecular G-quadruplexes may have less biological significance than intramolecular G-quadruplexes, which can be formed at the 3' overhang of telomeric DNA, but are likely to play a role *in vivo* in phenomena such as

recombination and end to end pairing. Binding to intermolecular quadruplexes may result in undesirable genomic effects including end-end fusion of chromosomes.

The absence of any significant exciton-type ICD signals when DB832 is titrated into these DNA sequences suggests that DB832 binds as a stacked species in the grooves of mixed parallel/antiparallel hybrid quadruplexes with a high degree of selectivity.

3.3.5 DB832 Induces the Formation of a Single Quadruplex Structure

To test the idea that DB832 is inducing the formation of a single mixed parallel/antiparallel quadruplex structure in the human telomere and is selective for hybrid type conformations, ^{1}H-NMR studies were performed on several telomeric quadruplex sequences. Tel22 exists as a mixture of conformations in K^{+} solution as previously shown by NMR studies; whereas, Tel26 and wtTel26 are shown to fold predominantly into hybrid-1 and hybrid-2 conformations respectively (discussed in Section 3.3.3). The imino proton spectra of Tel22 show the presence of more than one conformer in solution judging by the number of imino proton signals (Figure 3.7 Top, ratio 0). When DB832 is added, significant changes in the imino proton region can be readily observed, with spectral broadening and shifts indicative of an intermediate exchange of ligands on NMR timescale. At higher DB832 ratios, the intensity of many imino protons is significantly reduced and a new set of signals is observed, suggesting selectivity for a particular quadruplex conformation from an ensemble.

The imino proton spectra of Tel26 show the presence of a distinct number of imino protons (Figure 3.7 Middle, ratio 0) corresponding to the total number of guanines in that sequence. This is consistent with an intramolecular monomeric quadruplex formation, hybrid-1. Upon titration of DB832, a very similar spectral broadening and shifts, as in the case of

Tel22, are observed revealing an intermediate exchange of ligands on NMR timescale. A number of imino resonances broaden in a site-specific fashion consistent with the local chemical environment of different residues being perturbed upon ligand binding. This indicates that the compound is selectively binding to the hybrid-1 conformation, as seen from CD studies (Section 3.3.3). A similar effect is also observed with the wtTel26 (Figure 3.7-Bottom) upon ligand binding, indicating preferential binding to the preformed hybrid conformation.

The sequence d[TG_4T] is shown to fold into a simple intermolecular quadruplex topology by NMR and confirmed by CD studies (Section 3.3.4) (Figure 3.1-G). Current CD results have shown that DB832 does not induce a change in quadruplex conformation even at high ratios of the ligand, and does not have any ICD signals. ^{1}H NMR studies of TG_4T were performed with DB832, as shown in Figure 3.8-B. In the absence of any ligand (Figure 3.8-B, ratio 0), the imino proton spectra confirms the parallel topology adopted by this sequence. Upon titration of DB832, significant upfield shifts are observed for certain imino resonances indicative of the ligand stacking on the terminal tetrads. This is consistent with the current CD studies, which also suggest end-stacking as the only possible binding mode. This is also a commonly observed phenomenon between distamycin and TG_4T sequence, where the ligand initially stacks at the ends followed by stacking in the grooves [41].

The sequence, U6U7, d(TAGGGUUTAGGGT), has been shown to exist in solution as an ensemble of parallel and antiparallel dimeric hairpin quadruplexes [42] (Figure 3.1-H). This sequence is suitable as a model system for studying interactions with DB832, since the CD spectra of DB832 titrated into this DNA exhibits the same patterns in both the induced wavelength region and the DNA region as observed with the intramolecular sequence d[$AG_3(T_2AG_3)_3$] (CD spectra not shown). The imino proton spectra of DB832 titrated into

U6U7 is shown in Figure 3.9-A. As the molar ratio of DB832 is increased, the NMR peaks become sharper, suggesting that the compound is selectively binding to a single DNA conformation and stabilizing the structure. This is further supported by TOCSY spectra of the U6U7 sequence with DB832 at a 2:1 ratio (Figure 3.9-B). The four peaks correspond to the two uracil H5-H6 interaction peaks from the parallel and antiparallel conformations. In the absence of DB832, all the peaks are of similar intensity indicating equilibrium of two different conformations. However, upon DB832 titration, there is a significant decrease in the intensity of the two most upfield peaks suggesting that the compound is preferably binding to a particular conformation from the mixture.

3.4 CONCLUSION

Many DNA-targeted anti-cancer drugs act by intercalating between base pairs of duplex DNA, which causes a disruption in transcription and replication and leads to cell death. A difficulty with these types of small molecules is that they typically exhibit non-selective binding to many sites. Drugs in use today against cancer and eukaryotic parasites have significant toxicity to the host and can cause deleterious side-effects and death. Telomeres and some oncogenes have been shown to form quadruplex structures *in vitro*, and may serve as structure-specific targets for anticancer as well as antiparasitic therapeutics. The unique structural features of quadruplexes offer a way to target DNA in a structure-specific manner, leading to increased selectivity for quadruplex over duplex DNA, as well as selectivity for a particular quadruplex structure over other quadruplexes. DB832 is a heterocyclic diamidine that binds to human telomeric DNA as a groove-specific, stacked species. It is highly selective for quadruplex binding, showing virtually no binding to duplex DNA. Regardless of the starting conformation of the human telomeric DNA, DB832 is capable of inducing the formation of a mixed parallel/antiparallel hybrid quadruplex structure, and can even induce

formation of this structure in the absence of salt. DB832 also binds very selectively as a stacked species to the 5'-diagonal/lateral/lateral-3' hybrid quadruplex structure which is preformed or induced by DB832 in human telomere and modified human telomeric DNA sequences. Table 3.1 summaries the interaction of DB832 with different quadruplex conformations in this study. The unique binding mode of DB832 as well as its selectivity for its DNA target may allow it to serve as the starting point for the design of unique compounds with high affinity and selectivity for human telomeric DNA, leading to improved therapeutics for treating cancer and parasitic infections with decreased side effects.

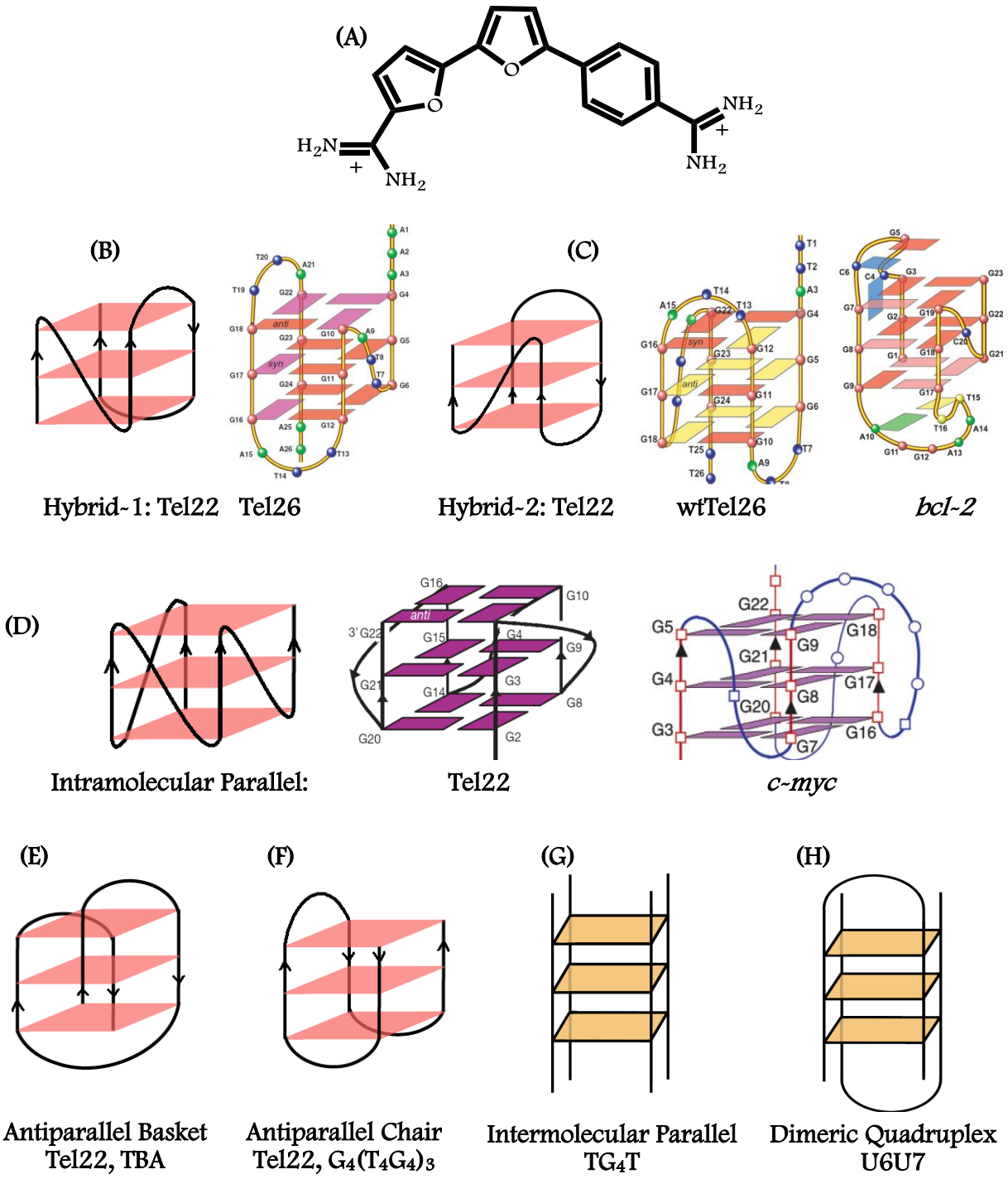


Figure 3.1: (A) Chemical structure of DB832. (B-H) Major folding topologies of quadruplex forming sequences used in this study.

References for the different quadruplex folds are provided within the main text.

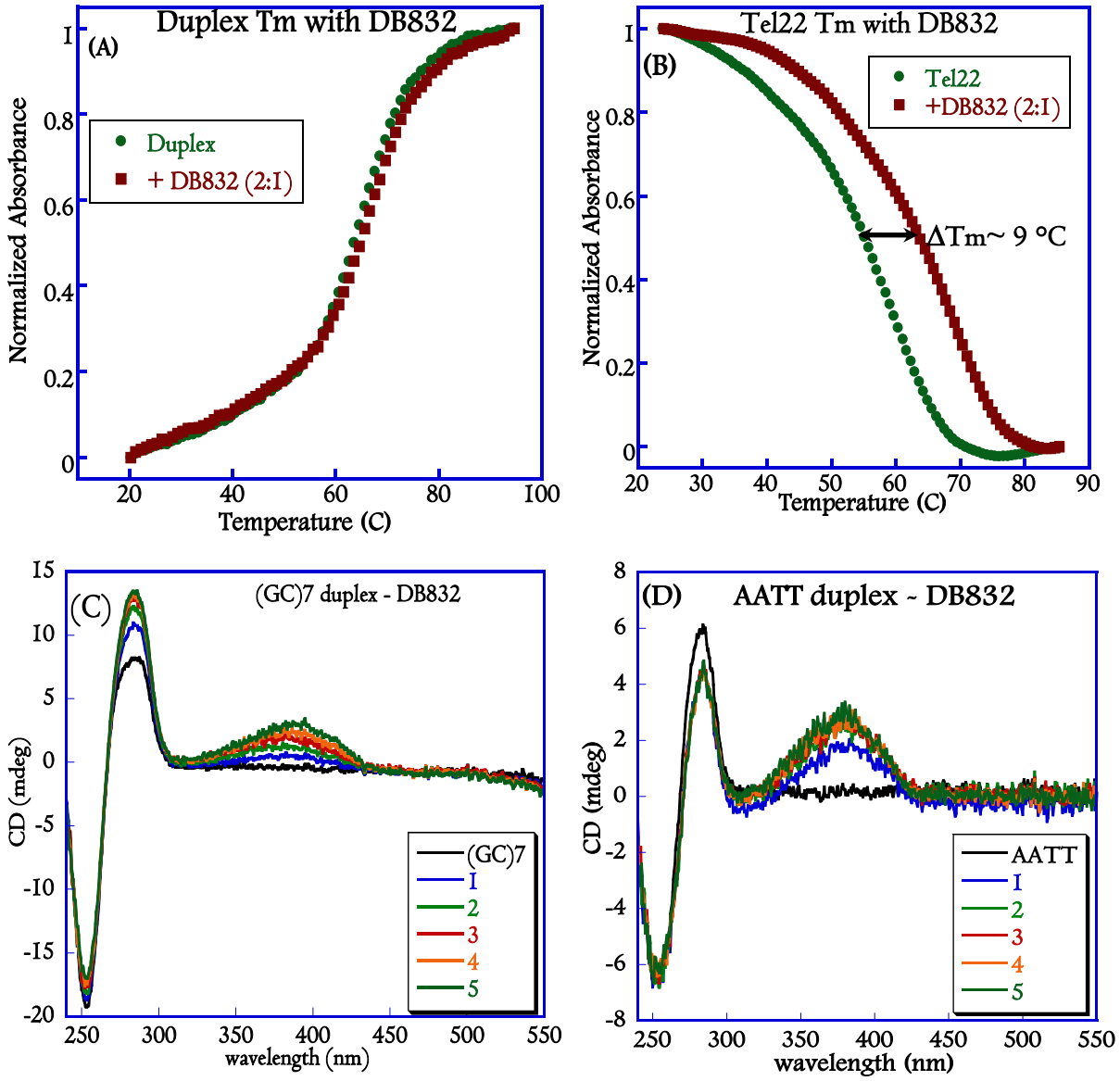


Figure 3.2: UV melting profiles of the hairpin duplex, d(CGAGATCAAAAGATCTCG), monitored at 260 nm (A) and the human telomere sequence, Tel22, monitored at 295 nm (B) in the absence and presence of DB832 in phosphate buffer containing 100 mM K^+. (C) CD spectra of DB832 titrated into d[$(GC)_7$] in HEPES buffer containing 50 mM KCl. Compound:DNA ratios ranged from 1:1 to 5:1. (D) CD spectra of DB832 titrated into d(GCGAATTCGC) in HEPES buffer containing 50 mM KCl.

Thermal melting compound:DNA ratios ranged from 1:1 to 5:1. DB832:DNA ratio was 2:1 for both DNAs. DB832 increases the melting temperature of the Tel22 by approximately 9 °C, while it has a negligible effect on the melting temperature of the duplex sequence. DNA/quadruplex concentration for Tm and CD experiments is ~3-5 μM unless otherwise mentioned.

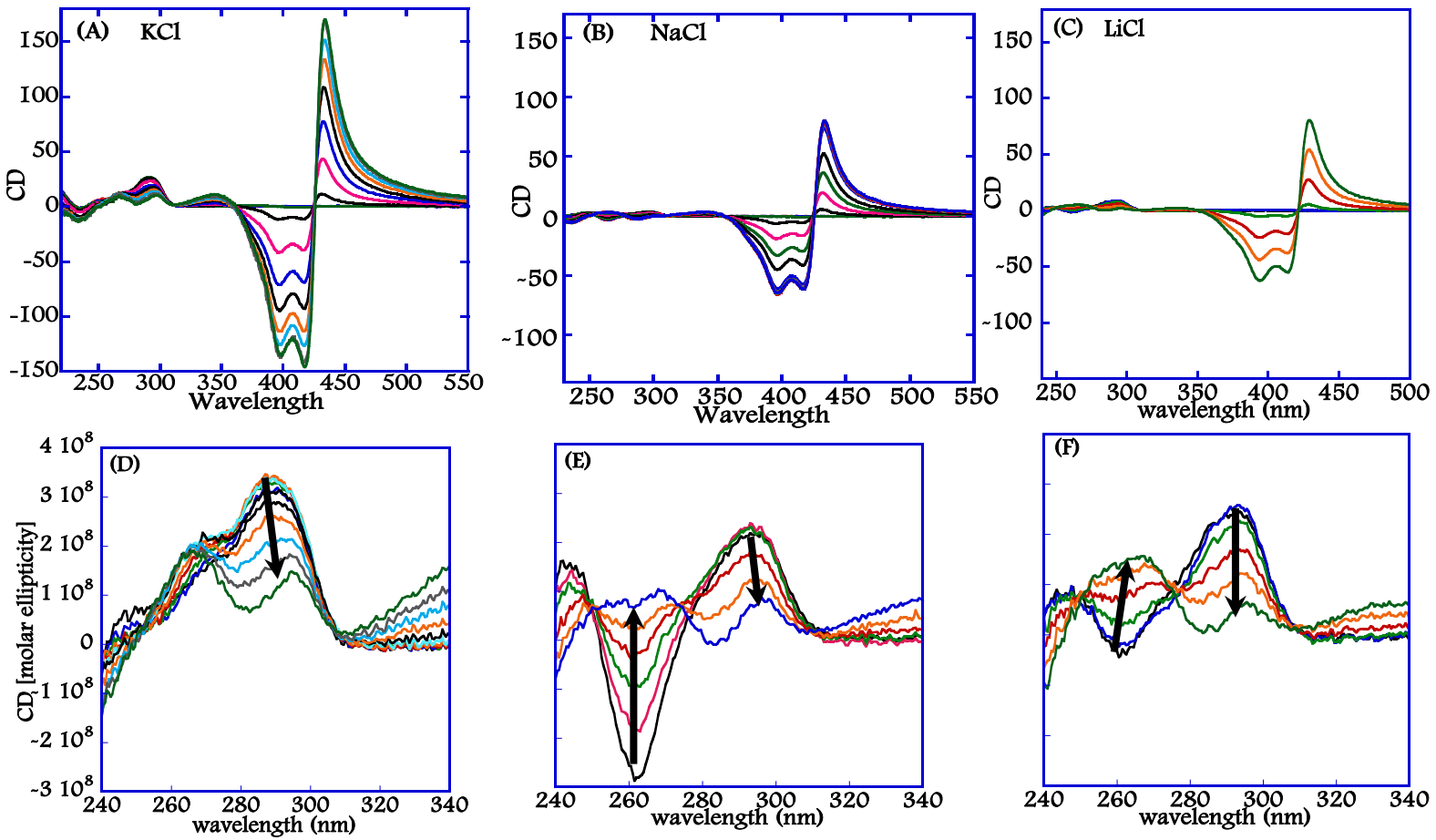


Figure 3.3: CD spectra of DB832 titrated into Tel22 in HEPES buffer containing 50 mM KCl (A), NaCl (B), or LiCl (C). (D-F) Close-up of the wavelength region of DNA absorbance for the spectra shown in A-C respectively.

In each case (top panels), the spectra show ICD signals, suggesting that DB832 is binding as a stacked species to human telomeric DNA under each of these conditions. The arrows (bottom panels) indicate increasing DB832 concentration. Although the initial spectra of the free DNA are quite different in each of these salts, upon addition of DB832, the CD spectra all exhibit peaks near both 265 and 295 nm. This suggests that regardless of the starting DNA conformation, DB832 induces the same final quadruplex conformation, which has both parallel and antiparallel characteristics.

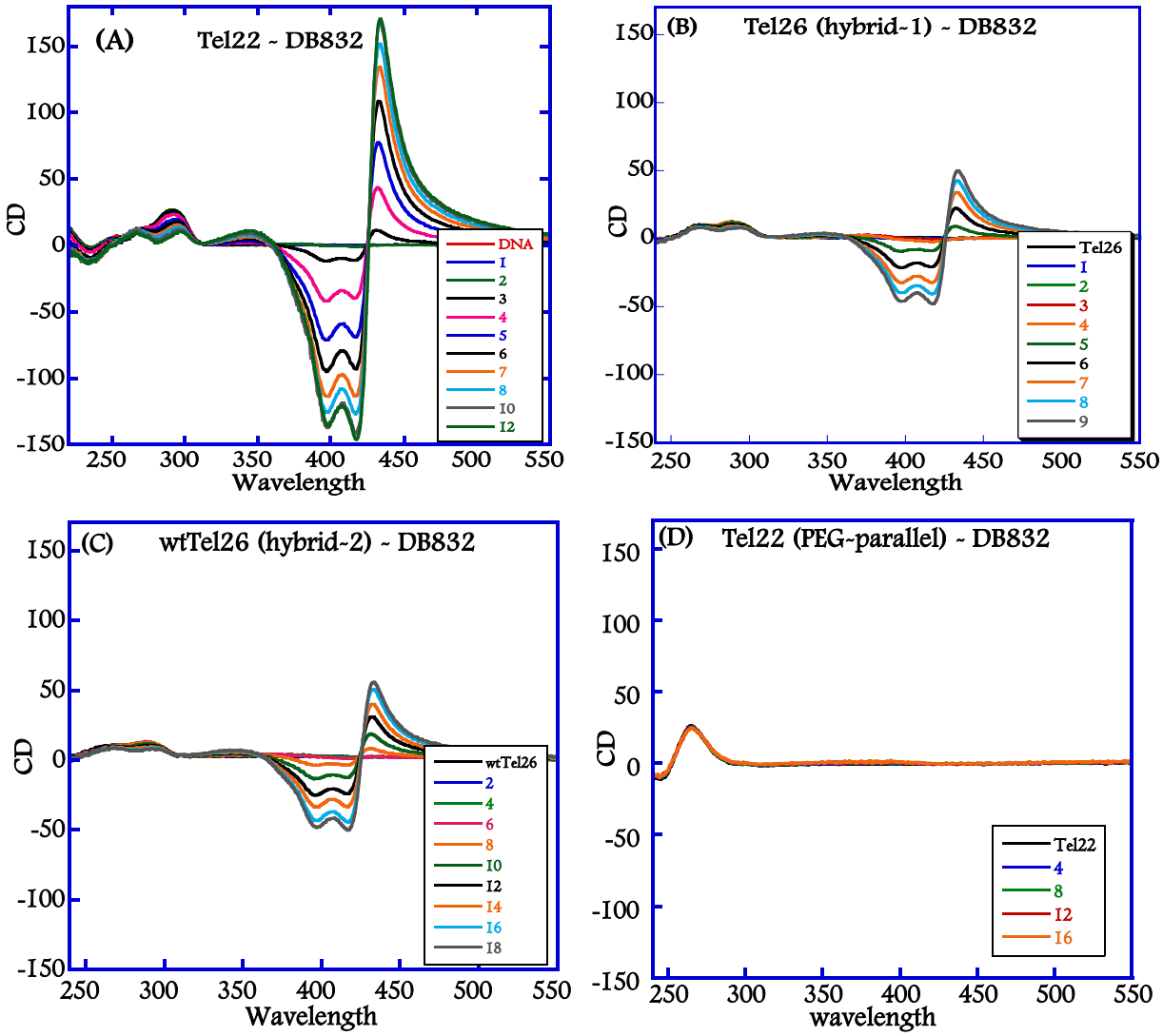


Figure 3.4: CD spectra of DB832 titrated into (A) Tel22, (B) Tel26, (C) wtTel26, and (D) Tel22 (with PEG 400).

Buffer conditions: HEPES buffer containing 50 mM KCl, 25 °C. DB832 was titrated into DNA solutions until saturation point is reached in the ICD region. Quadruplex concentration is ~ 3-5 μM unless otherwise mentioned. Exciton splitting is observed for Tel22, Tel26 and wtTel26 in the induced region, indicating that DB832 binds as a stacked species to the mixed parallel/antiparallel hybrid structure formed by these sequences. Tel22, with PEG 400, predominantly folds into a parallel conformation. Absence of any exciton-type ICD signal under these conditions suggests that the parallel conformation is not a preferred fold for DB832.

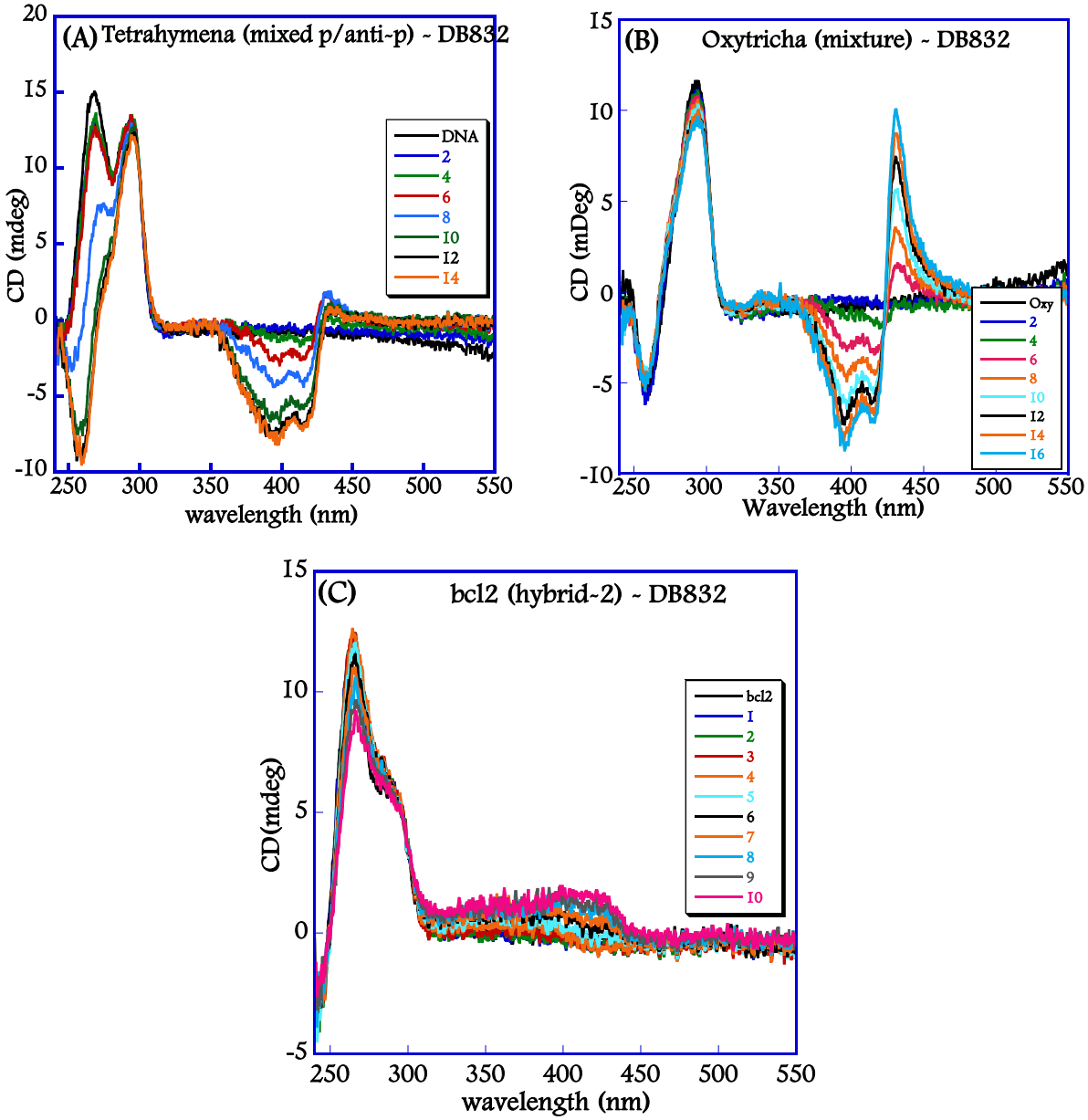


Figure 3.5: CD spectra of DB832 titrated into (A) Tetrahymena telomeric sequence, $d(T_2G_4)_4$, (B) Oxytricha, $d[G_4(T_4G_4)_3]$, and (C) *bcl-2*, $d(G_3CGCG_3AG_2A_2T_2G_3CG_3)$ quadruplex sequences.

Buffer conditions: HEPES buffer containing 50 mM KCl, 25 °C. DB832 was titrated into DNA solutions until saturation point is reached in the ICD region. Quadruplex concentration is ~ 3-5 μM unless otherwise mentioned. DB832 does not exhibit any exciton-type splitting with the bcl-2 sequence, and only exhibits small ICD signal at high concentrations with the Tetrahymena telomere. The large change in the CD signal in the DNA region of Tetrahymena suggests that the DNA undergoes a major conformation change upon the addition of DB832. Oxytricha exists as a mixture of quadruplex conformations in K^+. DB832 might be selectively binding to a single conformation from this ensemble, but exhibits a low ICD signal when compared to Tel22.

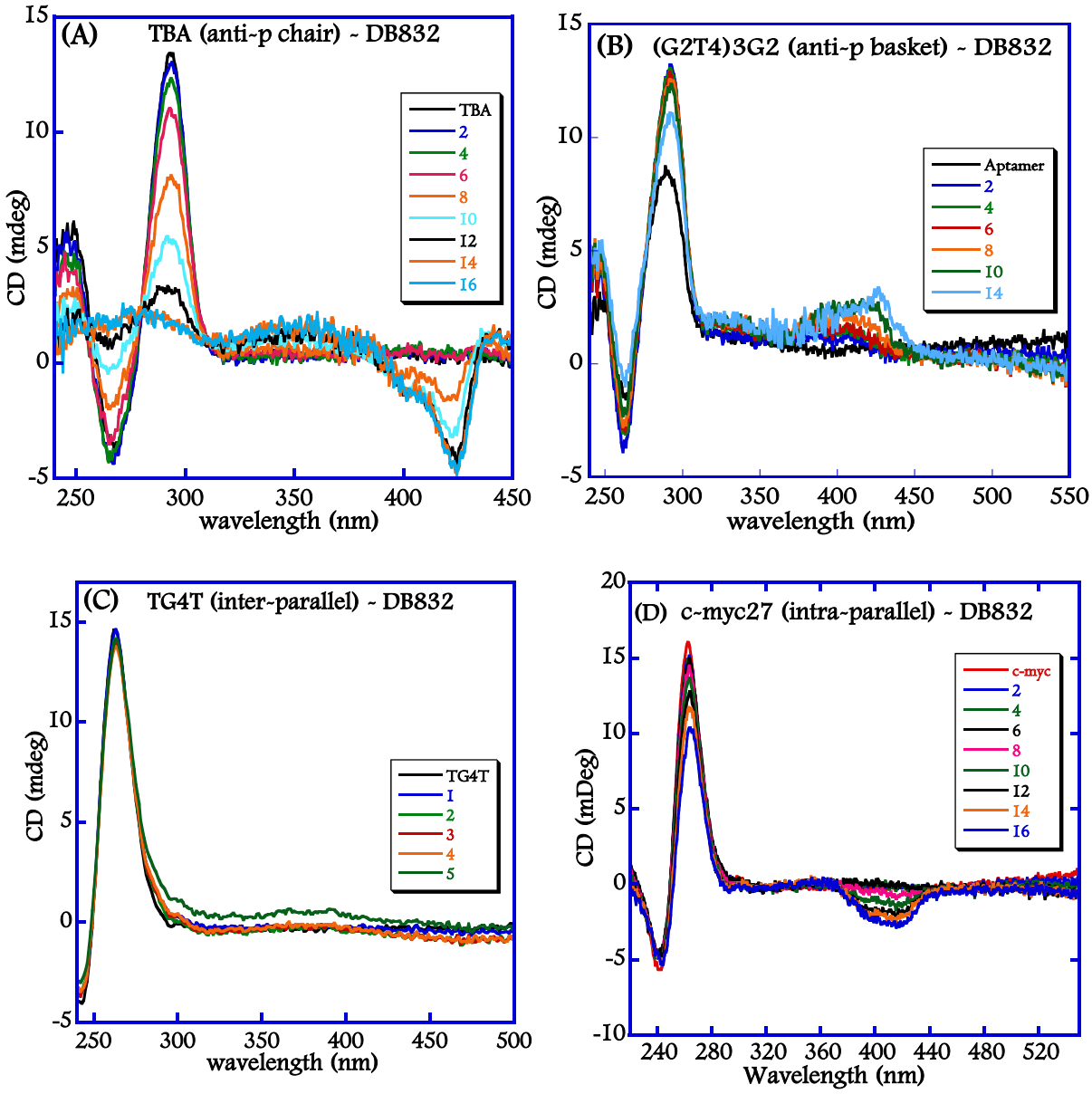


Figure 3.6: CD spectra of DB832 titrated into (A) TBA, d($G_2T_2G_2TGTG_2T_2G_2$), (B) d$(G_2T_4)_3G_2$, (C) TG_4T, and (D) *c-myc27*, d($TG_4AG_3TG_4AG_3TG_4AAG_2$) quadruplex sequences.

Buffer conditions: HEPES buffer containing 50 mM KCl, 25 °C. DB832 was titrated into DNA solutions until saturation point is reached in the ICD region. Quadruplex concentration is ~ 3-5 μM unless otherwise mentioned.

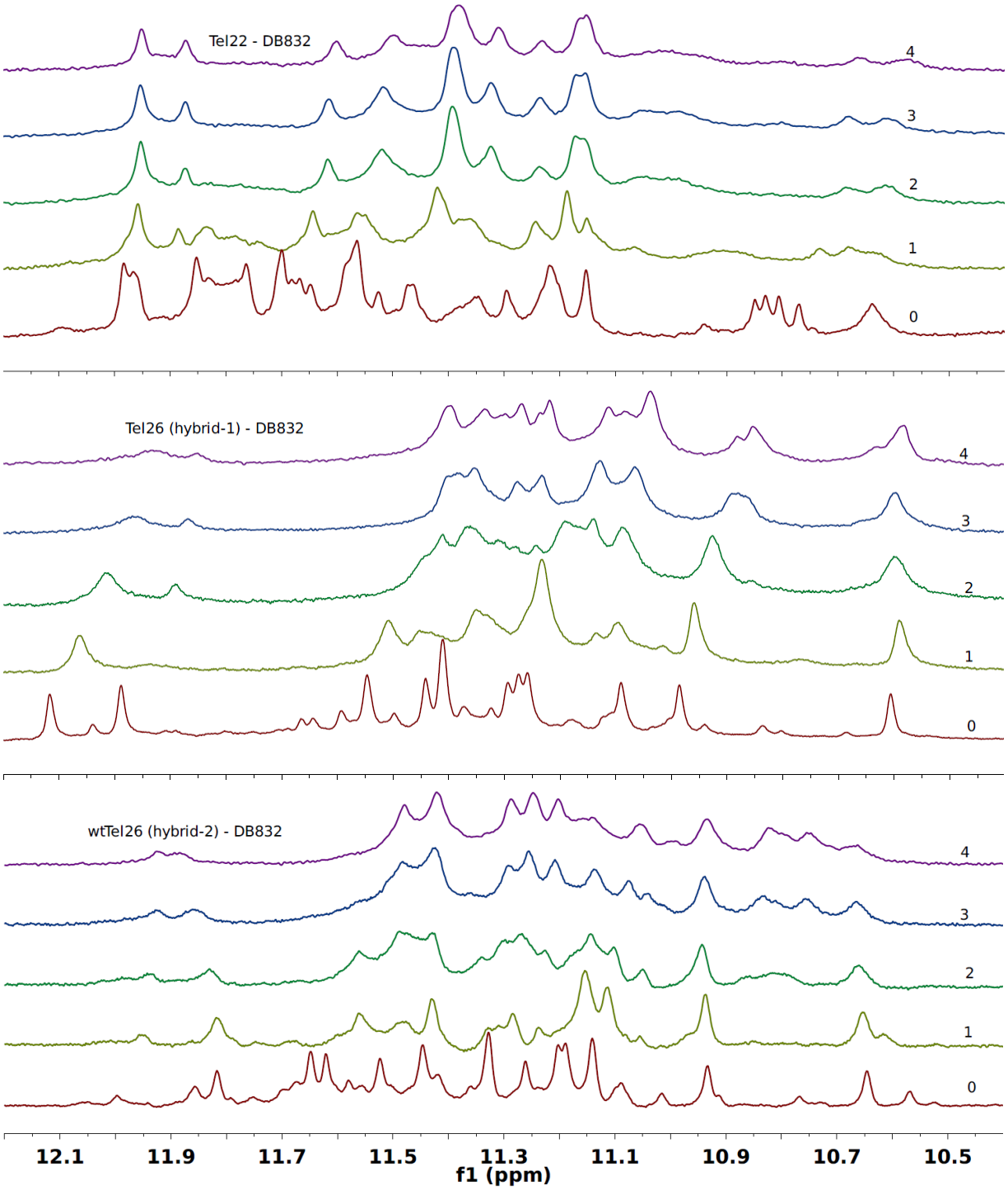


Figure 3.7: Imino proton spectra of DB832 with Tel22 (top), Tel26 (middle), and wtTel26 (bottom) at 25 °C in 10 mM K_2HPO_4/80 mM KCl, pH 7.0.

As the molar ratio of DB832 is increased, distinct chemical shifts for some guanine imino protons are observed for all the sequences indicating that the compound is binding to these quadruplex conformations. These results show that DB832 is selective for the diagonal/lateral/lateral-type hybrid quadruplex structure.

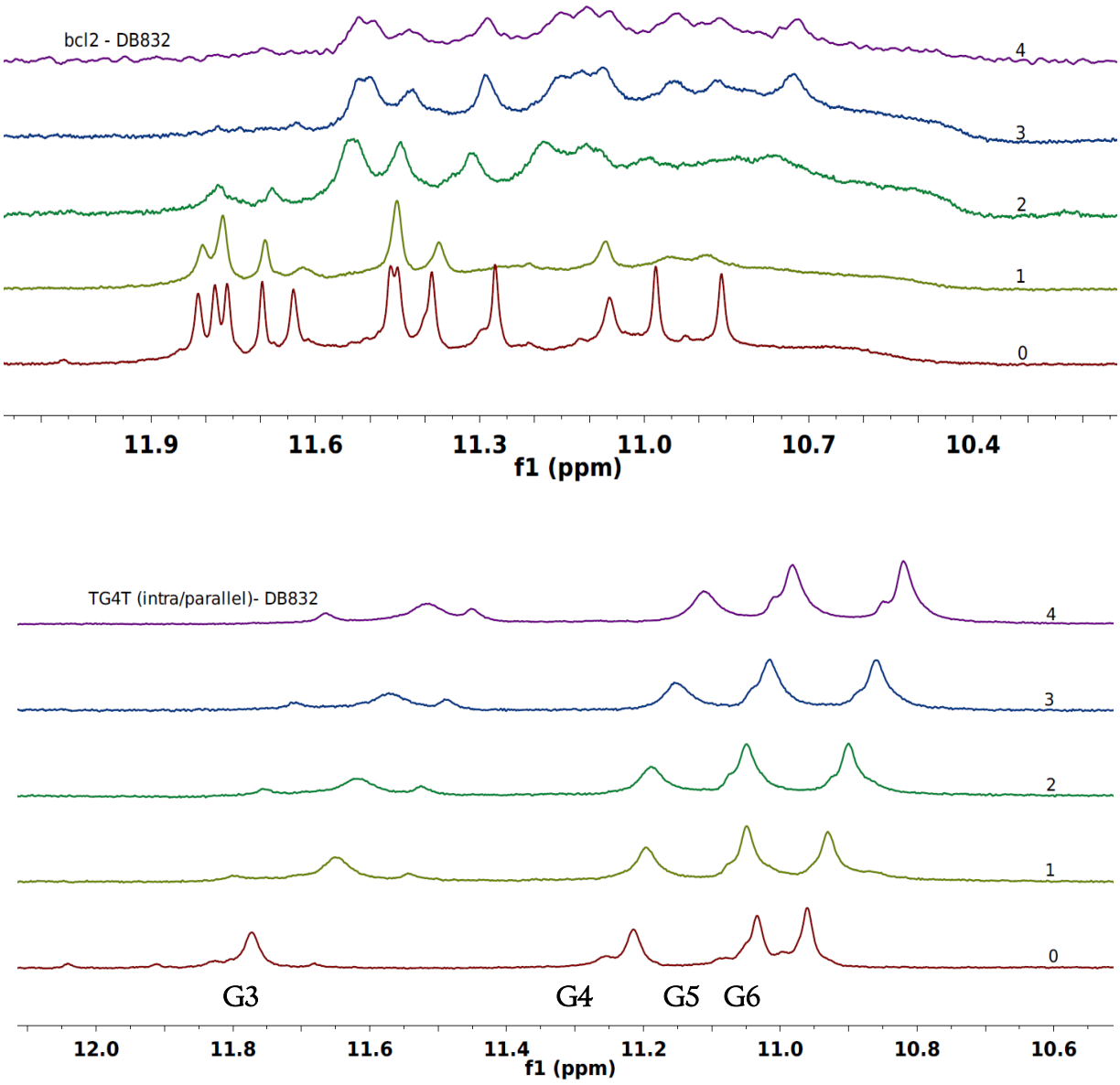


Figure 3.8: NMR imino proton titrations of (A) *bcl-2* promoter sequence and (B) TG_4T with DB832.

As the molar ratio of DB832 is increased, the guanine imino peaks of bcl-2 undergoes significant broadening, indicating nonspecific binding. TG_4T exhibits significant upfield shifts for some imino protons. This is consistent with the stacking of the compound at the terminal tetrads of this sequence: a commonly recurring theme with most of the small molecules interacting with this conformation. Imino proton assignments are for $d[TAGGGGTT]_4$ intermolecular quadruplex obtained from [41].

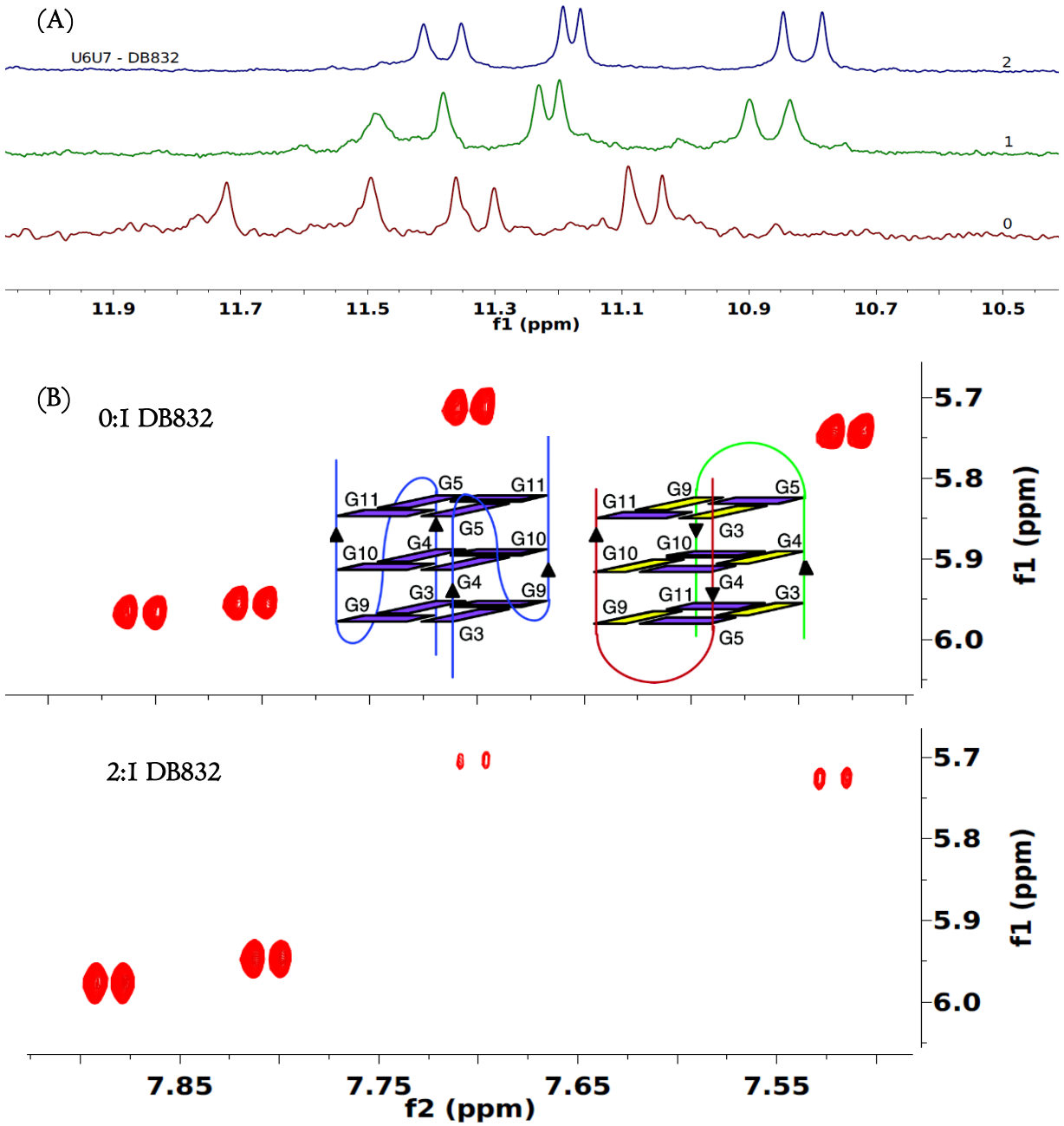


Figure 3.9: NMR imino proton titration of U6U7, d(TAGGGUUAGGGT) dimeric hairpin quadruplex with DB832 (A). (B) TOCSY spectra of uracil H5-H6 cross-peaks for 0.5 mM U6U7 at 0:1 (top), and 2:1 (bottom) DB832 molar ratios at 308K.

Buffer conditions: 80 mM KCl/10 mM K_2HPO_4, 0.1 mM EDTA, pH 7.0 at 35 °C. As the molar ratio of DB832 is increased, the peaks exhibit upfield shifting and become sharper, (A). At a ratio of 2:1, only six peaks are seen, indicating that the compound is binding to a single DNA conformation for each DNA sequence. In the absence of the compound, U6U7 exists as a mixture of conformations (shown in inset). Upon DB832 titration, the intensity of upfield peaks significantly decreases (B, 2:1), suggesting that the compound is binding to a single hybrid quadruplex conformation from a mixture.

Table 3.1: DNA sequences selected for CD studies with DB832.

DNA	Sequence	Conformation[a]	Exciton Splitting
Tel22	d[$AG_3(T_2AG_3)_3$]	Mixture	Yes
Tel26	d[$A_3(G_3T_2A)_3G_3A_2$]	Hybrid-1	Yes
wtTel26	d[$T_2G_3(T_2AG_3)_3A$]	Hybrid-2	Yes
c-myc27	d[$TG_4AG_3TG_4AG_3TG_4AAG_2$]	Intra-Parallel	No
Oxytricha	d[$G_4(T_4G_4)_3$]	Mixed parallel/antiparallel	No
Tetrahymena	d[$T_2G_4]_4$	Mixed parallel/antiparallel	No
bcl-2	d[$G_3CGCG_3AG_2A_2T_2G_3CG_3$]	Mixed parallel/antiparallel	No
Thrombin Binding Aptamer	d[$G_2T_2G_2TGTG_2T_2G_2$]	Antiparallel-chair	No
G-rich oligonucleotide	d[$(G_2T_4)_3G_2$]	Antiparallel-basket	No
TG_4T	d[TG_4T]	Inter-Parallel	No
AATT	d[GCGAATTCGC]	Duplex	No
$(GC)_7$	d[CGAGATCAAAAGATCTCG]	Duplex	No
Mixed duplex	d[$GC]_7$	Hairpin	No

[a] *Quadruplex conformations reported are determined in K^+.*

3.5 REFERENCES

[1] a)T. R. Cech, *Nature* **1988**, *332*, 777; b)E. Henderson, C. C. Hardin, S. K. Walk, I. Tinoco, E. H. Blackburn, *Cell* **1987**, *51*, 899.

[2] D. Sen, W. Gilbert, *Nature* **1988**, *334*, 364.

[3] M. Fry, L. A. Loeb, *Proceedings of the National Academy of Sciences of the United States of America* **1994**, *91*, 4950.

[4] a)G. C. Siddiqui-Jain A, Bearss DJ, Hurley LH., *Proceedings of the National Academy of Sciences of the United States of America* **2002**, *99*, 6; b)T. Simonsson, P. Pecinka, M. Kubista, *Nucl. Acids Res.* **1998**, *26*, 1167; c)T. S. Dexheimer, D. Sun, L. H. Hurley, *Journal of the American Chemical Society* **2006**, *128*, 5404.

[5] P. Catasti, X. Chen, R. K. Moyzis, E. M. Bradbury, G. Gupta, *Journal of Molecular Biology* **1996**, *264*, 534.

[6] a)J. Dai, D. Chen, R. A. Jones, L. H. Hurley, D. Yang, *Nucl. Acids Res.* **2006**, *34*, 5133; b)J. Dai, T. S. Dexheimer, D. Chen, M. Carver, A. Ambrus, R. A. Jones, D. Yang, *Journal of the American Chemical Society* **2006**, *128*, 1096; c)H. Fernando, A. P. Reszka, J. Huppert, S. Ladame, S. Rankin, A. R. Venkitaraman, S. Neidle, S. Balasubramanian, *Biochemistry* **2006**, *45*, 7854; d)J. Li, J. J. Correia, L. Wang, J. O. Trent, J. B. Chaires, *Nucl. Acids Res.* **2005**, *33*, 4649; e)G. N. Parkinson, M. P. H. Lee, S. Neidle, *Nature* **2002**, *417*, 876; f)A. T. Phan, Y. S. Modi, D. J. Patel, *Journal of the American Chemical Society* **2004**, *126*, 8710; g)P. Schultze, R. F. Macaya, J. Feigon, *Journal of Molecular Biology* **1994**, *235*, 1532; h)Y. Wang, D. J. Patel, *Structure* **1993**, *1*, 263; i)Y. Wang, D. J. Patel, *Structure* **1994**, *2*, 1141.

[7] a)E. H. Blackburn, *Nature* **1991**, *350*, 569; b)D. Rhodes, R. Giraldo, *Curr Opin Struct Biol* **1995**, *5*, 311; c)L. L. Sandell, V. A. Zakian, *Cell* **1993**, *75*, 729.

[8] a)H. Arthanari, P. H. Bolton, *Chem Biol* **2001**, *8*, 221; b)P. Borst, S. Ulbert, *Mol Biochem Parasitol* **2001**, *114*, 17; c)G. Ghosal, K. Muniyappa, *Biochem Biophys Res Commun* **2006**, *343*, 1.

[9] a)F. W. Smith, P. Schultze, J. Feigon, *Structure* **1995**, *3*, 997; b)Y. Wang, D. J. Patel, *Structure* **1994**, *2*, 1141; c)Y. Wang, D. J. Patel, *J Mol Biol* **1995**, *251*, 76.

[10] a)G. Fang, T. R. Cech, *Cell* **1993**, *74*, 875; b)G. Fang, T. R. Cech, *Biochemistry* **1993**, *32*, 11646; c)R. Giraldo, M. Suzuki, L. Chapman, D. Rhodes, *Proc Natl Acad Sci U S A* **1994**, *91*, 7658; d)C. Granotier, G. Pennarun, L. Riou, F. Hoffschir, L. R. Gauthier, A. de Cian, D. Gomez, E. Mandine, J.-F. Riou, J.-L. Mergny, P. Mailliet, B. Dutrillaux, F. D. Boussin, *Nucleic Acids Research* **2005**, *33*, 4182.

[11] a)Y. Xu, Y. Suzuki, K. Kaminaga, M. Komiyama, *NUCLEIC ACIDS SYMP SER (OXF)* **2009**, *53*, 63; b)S. Neidle, G. Parkinson, *Nat Rev Drug Discov* **2002**, *1*, 383.

[12] G. N. Parkinson, M. P. Lee, S. Neidle, *Nature* **2002**, *417*, 876.

[13] a)A. Ambrus, D. Chen, J. Dai, T. Bialis, R. A. Jones, D. Yang, *Nucl. Acids Res.* **2006**, *34*, 2723; b)K. N. Luu, A. T. Phan, V. Kuryavyi, L. Lacroix, D. J. Patel, *Journal of the American Chemical Society* **2006**, *128*, 9963; c)J. Dai, M. Carver, D. Yang, *Biochimie* **2008**, *90*, 1172.

[14] K. W. Lim, S. Amrane, S. Bouaziz, W. Xu, Y. Mu, D. J. Patel, K. N. Luu, A. T. n. Phan, *Journal of the American Chemical Society* **2009**, *131*, 4301.

[15] a)G. Fang, T. R. Cech, *Cell* **1993**, *74*, 875; b)R. Giraldo, M. Suzuki, L. Chapman, D. Rhodes, *Proceedings of the National Academy of Sciences of the United States of America* **1994**, *91*, 7658; c)C. Granotier, G. Pennarun, L. Riou, F. Hoffschir, L. R. Gauthier, A. De Cian, D. Gomez, E. Mandine, J.-F. Riou, J.-L. Mergny, P. Mailliet, B. Dutrillaux, F. D. Boussin, *Nucl. Acids Res.* **2005**, *33*, 4182; d)K. Paeschke, T. Simonsson, J. Postberg, D. Rhodes, H. J. Lipps, *Nat Struct Mol Biol* **2005**, *12*, 847.

[16] a)C. M. Incles, C. M. Schultes, S. Neidle, *Curr Opin Investig Drugs* **2003**, *4*, 675; b)L. R. Kelland, *Eur J Cancer* **2005**, *41*, 971; c)E. M. Rezler, D. J. Bearss, L. H. Hurley, *Curr Opin Pharmacol* **2002**, *2*, 415; d)G. Saretzki, *Cancer Lett* **2003**, *194*, 209; e)D. Sun, B. Thompson, B. E. Cathers, M. Salazar, S. M. Kerwin, J. O. Trent, T. C. Jenkins, S. Neidle, L. H. Hurley, *J Med Chem* **1997**, *40*, 2113.

[17] a)E. M. Rezler, D. J. Bearss, L. H. Hurley, *Annu Rev Pharmacol Toxicol* **2003**, *43*, 359; b)K. A. Olaussen, K. Dubrana, J. Domont, J. P. Spano, L. Sabatier, J. C. Soria, *Crit Rev Oncol Hematol* **2006**, *57*, 191.

[18] a)A. Ambrus, D. Chen, J. Dai, R. A. Jones, D. Yang, *Biochemistry* **2005**, *44*, 2048; b)A. T. Phan, V. Kuryavyi, S. Burge, S. Neidle, D. J. Patel, *Journal of the American Chemical Society* **2007**, *129*, 4386; c)S. Rankin, A. P. Reszka, J. Huppert, M. Zloh, G. N. Parkinson, A. K. Todd, S. Ladame, S. Balasubramanian, S. Neidle, *Journal of the American Chemical Society* **2005**, *127*, 10584.

[19] a)Y. Qin, L. H. Hurley, *Biochimie* **2008**, *90*, 1149; b)J. Seenisamy, E. M. Rezler, T. J. Powell, D. Tye, V. Gokhale, C. S. Joshi, A. Siddiqui-Jain, L. H. Hurley, *Journal of the American Chemical Society* **2004**, *126*, 8702.

[20] a)G. S. H. Lee, M. A. Wilson, B. R. Young, *Organic Geochemistry* **1998**, *28*, 549; b)M. Liu, X.-a. Mao, C. Ye, H. Huang, J. K. Nicholson, J. C. Lindon, *Journal of Magnetic Resonance* **1998**, *132*, 125; c)A. Louis-Joseph, D. Abergel, I. Lebars, J.-Y. Lallemand, *Chemical Physics Letters* **2001**, *337*, 92; d)P. Plateau, M. Gueron, *Journal of the American Chemical Society* **1982**, *104*, 7310.

[21] a)L. R. Brown, B. C. Sanctuary, *Journal of Magnetic Resonance (1969)* **1991**, *91*, 413; b)C. Griesinger, G. Otting, K. Wuethrich, R. R. Ernst, *Journal of the American Chemical Society* **1988**, *110*, 7870; c)M. Kadkhodaei, T. L. Hwang, J. Tang, A. J. Shaka, *Journal of Magnetic Resonance, Series A* **1993**, *105*, 104.

[22] F. X. Han, R. T. Wheelhouse, L. H. Hurley, *Journal of the American Chemical Society* **1999**, *121*, 3561.

[23] a)D. G. Alexeev, A. A. Lipanov, I. Ya. Skuratovskii, *Nature* **1987**, *325*, 821; b)D. S. Goodsell, M. L. Kopka, D. Cascio, R. E. Dickerson, *Proceedings of the National Academy*

of Sciences of the United States of America **1993**, *90*, 2930; c)R. H. Sarma, M. H. Sarma, L. Dai, K. Umemoto, *FEBS Letters* **1997**, *418*, 76.

[24] a)Y. Xu, Y. Noguchi, H. Sugiyama, *Bioorganic & Medicinal Chemistry* **2006**, *14*, 5584; b)W. Li, P. Wu, T. Ohmichi, N. Sugimoto, *FEBS Letters* **2002**, *526*, 77; c)P. Balagurumoorthy, S. K. Brahmachari, D. Mohanty, M. Bansal, V. Sasisekharan, *Nucl. Acids Res.* **1992**, *20*, 4061.

[25] Y. Wang, D. J. Patel, *Structure* **1993**, *1*, 263.

[26] a)Y. He, R. D. Neumann, I. G. Panyutin, *Nucleic Acids Res* **2004**, *32*, 5359; b)E. M. Rezler, J. Seenisamy, S. Bashyam, M. Y. Kim, E. White, W. D. Wilson, L. H. Hurley, *J Am Chem Soc* **2005**, *127*, 9439; c)J. Li, J. J. Correia, L. Wang, J. O. Trent, J. B. Chaires, *Nucleic Acids Res* **2005**, *33*, 4649.

[27] a)J. Y. Lee, J. Yoon, H. W. Kihm, D. S. Kim, *Biochemistry* **2008**, *47*, 3389; b)D. Renciuk, M. Zemánek, I. Kejnovská, M. Vorlícková, *Biochimie* **2009**, *91*, 416.

[28] a)K. N. Luu, A. T. Phan, V. Kuryavyi, L. Lacroix, D. J. Patel, *J Am Chem Soc* **2006**, *128*, 9963; b)A. Ambrus, D. Chen, J. Dai, T. Bialis, R. A. Jones, D. Yang, *Nucleic Acids Res* **2006**, *34*, 2723; c)J. Dai, C. Punchihewa, A. Ambrus, D. Chen, R. A. Jones, D. Yang, *Nucleic Acids Res* **2007**, *35*, 2440.

[29] R. J. Ellis, *Trends in Biochemical Sciences* **2001**, *26*, 597.

[30] a)Y. Xue, Z.-y. Kan, Q. Wang, Y. Yao, J. Liu, Y.-h. Hao, Z. Tan, *Journal of the American Chemical Society* **2007**, *129*, 11185; b)J. Zhou, C. Wei, G. Jia, X. Wang, Q. Tang, Z. Feng, C. Li, *Biophysical Chemistry* **2008**, *136*, 124.

[31] D. Miyoshi, A. Nakao, N. Sugimoto, *Biochemistry* **2002**, *41*, 15017.

[32] J. Dai, D. Chen, R. A. Jones, L. H. Hurley, D. Yang, *Nucleic Acids Res* **2006**, *34*, 5133.

[33] a)J. Dai, T. S. Dexheimer, D. Chen, M. Carver, A. Ambrus, R. A. Jones, D. Yang, *J Am Chem Soc* **2006**, *128*, 1096; b)T. S. Dexheimer, D. Sun, L. H. Hurley, *J Am Chem Soc* **2006**, *128*, 5404.

[34] Y. Wang, D. J. Patel, *Journal of Molecular Biology* **1995**, *251*, 76.

[35] S. Haider, G. N. Parkinson, S. Neidle, *J Mol Biol* **2002**, *320*, 189.

[36] a)E. N. Brody, L. Gold, *Reviews in Molecular Biotechnology* **2000**, *74*, 5; b)E. N. Brody, M. C. Willis, J. D. Smith, S. Jayasena, D. Zichi, L. Gold, *Molecular Diagnosis* **1999**, *4*, 381; c)D. Jellinek, L. S. Green, C. Bell, C. K. Lynott, N. Gill, C. Vargeese, G. Kirschenheuter, D. P. C. McGee, P. Abesinghe, *Biochemistry* **1995**, *34*, 11363.

[37] a)P. J. Bates, D. A. Laber, D. M. Miller, S. D. Thomas, J. O. Trent, *Experimental and Molecular Pathology* **2009**, *86*, 151; b)S. Soundararajan, L. Wang, V. Sridharan, W. Chen, N. Courtenay-Luck, D. Jones, E. K. Spicer, D. J. Fernandes, *Molecular Pharmacology* **2009**, *76*, 984; c)Y. Teng, A. C. Girvan, L. K. Casson, W. M. Pierce, M. Qian, S. D. Thomas, P. J. Bates, *Cancer Research* **2007**, *67*, 10491.

[38] V. Dapic, V. Abdomerovic, R. Marrington, J. Peberdy, A. Rodger, J. O. Trent, P. J. Bates, *Nucleic Acids Res* **2003**, *31*, 2097.

[39] A. T. Phan, Y. S. Modi, D. J. Patel, *J Am Chem Soc* **2004**, *126*, 8710.

[40] K. Phillips, Z. Dauter, A. I. H. Murchie, D. M. J. Lilley, B. Luisi, *Journal of Molecular Biology* **1997**, *273*, 171.

[41] L. Martino, A. Virno, B. Pagano, A. Virgilio, S. Di Micco, A. Galeone, C. Giancola, G. Bifulco, L. Mayol, A. Randazzo, *Journal of the American Chemical Society* **2007**, *129*, 16048.

[42] A. T. Phan, D. J. Patel, *Journal of the American Chemical Society* **2003**, *125*, 15021.

4 HETEROCYCLIC DIAMIDINES AS POTENTIAL G-QUADRUPLEX TARGETING AGENTS

4.1 INTRODUCTION

DNA is an important drug target in anticancer therapies and development of successful anticancer agents is highly dependent on the discovery and understanding of the DNA duplex and its associated processes. Unfortunately, many of the currently available drugs that target DNA are cytotoxic and nonspecific. As such, extensive efforts have been directed toward the discovery of new agents with improved selectivity and less cytotoxicity [1]. Apart from the typical right-handed double helix, DNA can adopt other biologically relevant structures, such as G-quadruplexes [2]. The unique structural characteristics of G-quadruplexes and their polymorphism can contribute to their varied biological roles. The major roles of the G-quadruplexes may be their ability to "turn-on" or "turn-off" some physiological events by the transcriptional regulation of genes or telomere length. The particular geometry of the G-quadruplex structure is thought to allow specific recognition by small molecules through various binding modes in a manner corresponding to that of duplex-DNA binders. The G-quadruplex architecture is recognized as a significant drug target for cancer and other diseases and extensive efforts have been directed toward the discovery of promising lead compounds capable of stabilizing G-quadruplexes [3]. Despite the emergence of a wide array of quadruplex-interactive small molecules, the interaction of most of these small molecules with G-quadruplexes *in vitro* have been shown to be primarily through by end-stacking mode. The intercalation mode between the tetrads of quadruplexes is considered to be highly improbable, not only because the G-quadruplex is an extremely stable and rigid structure, but also because distortion of quadruplex integrity requires a very high energy cost. Electrostatic interactions between cationic ligands and the anionic backbone can also be important for improving their binding affinity. Most importantly, quadruplex groove-binding has been shown to be a viable

strategy to develop a completely new spectrum of small molecules [4]. Moreover, opportunities do exist to combine any of these binding modes to get enhanced degrees of binding specificity.

DNA duplex binders constitute an important class of small molecules that have exhibited great success in antiparasitic and anticancer therapy [5]. The minor groove of double helical B-DNA has been a site of greatest interest for developing new drugs since it is the site of non-covalent high sequence specific interactions for a large number of small molecules. The DNA-binding properties of these agents have been examined in great detail and have provided insights into the structural and functional features that contribute to binding selectivity. Polyamides and diamidines have garnered the most attention as the most successful duplex groove-binding agents [5e, 6]. Much of the facts that are established about DNA minor groove binders are based on detailed studies of polyamides such as distamycin and netropsin (Figure 4.1: A, B). These compounds are characterized by repeating pyrrole units connected by amide bonds and ending with one or more positively charged nitrogen atoms. They are capable of interacting with the DNA minor groove based on their curved shape, which matches well with the topology of double-stranded DNA. Distamycin and netropsin are potent inhibitors of Topoisomerase I, II and DNA helicases [1f, 7]. Diamidines such as DAPI, berenil, and pentamidine (Figure 4.1: C-E) were proved to be therapeutically successful agents against several protozoan diseases [8]. DAPI binds to the minor groove with the phenyl and indole rings parallel to the groove walls, covering the four basepair sequence, AATT [9]. The amidine groups at the termini of the molecule seemed to contribute significantly to the complex stability between the dications and the DNA through H-bonding and electrostatic interactions. A weaker intercalation mode at GC sites is also commonly observed for DAPI [10]. Berenil, a diaryl diamidine, is widely used for the treatment of trypanosomal infections in animals [8d, 8e]. Its biological activity has been related to a selective inhibition of kinetoplasts, specialized DNA-containing regions of the mitochondrion [11]. This compound is also characterized by a high

affinity for the AT-rich sequences [12]. Pentamidine, also an aromatic diamidine, is particularly effective in the treatment of *P. carinii* pneumonia (PCP) in patients with AIDS [13]. Footprinting analysis indicates a strong preference of pentamidine for the AT-rich region of the DNA minor groove containing at least five consecutive AT base pairs [14]. Although, the mechanism of pentamidine action against *P. carinii* is unknown, there is considerable evidence that direct interaction with the pathogenic genome is important for its activity [15]. Although pentamidine is one of the drugs of choice used for the treatment of PCP, this compound is associated with serious toxicities, and has prompted the search for new and safer agents to combat this fatal infection, which afflicts the majority of AIDS patients. Although a wide number of pentamidine analogs have been developed, there is still an ongoing need for more potent and less toxic derivatives of pentamidine [1b, 16]. The replacement of the linker in either berenil or pentamidine with a furan ring system resulted in furamidine (DB75) (Figure 4.1: F) shown to be less toxic and more effective against PCP in immunosupressed rat models [17]. In addition, DB75 is active against diverse highly infectious parasites such as *Giardia lamblia, Plasmodium falcifarum,* and *Trypanosoma rhodesience* [17b, 18]. Interestingly, replacement of a phenyl ring with a benzimidazole system resulted in DB293 (Figure 4.1: G), the first unfused aromatic dication capable of forming stacked dimer in the DNA minor groove of GC-containing sequence [19]. In particular, a 13-bp sequence including two contiguous ATGA motifs provided a highly preferential recognition site for this compound [20]. This is particularly important considering the dearth of small molecules that can selectively recognize sequences containing guanine residues, and the possible potential of these ligands to recognize guanine-rich quadruplex forming sequences as well.

Owing to the success as potent antiparasitic and anticancer duplex-interactive agents, it is not highly inconceivable to use polyamides and heterocyclic diamidines also as potential anticancer quadruplex-interactive agents. However, polyamides that can interact with

quadruplex systems have not been reported so far. In fact, Distamycin-A (Dist-A) is the only polyamide that was shown to recognize simple quadruplex systems [21]. Dist-A was shown to bind with a 4:1 stoichiometry as an asymmetric dimer in the opposite grooves of $[d(TG_4T)]_4$ sequence [22]. An end-stacking interaction mode has however, been shown to be most important in more complex quadruplex systems [21]. The lack of selectivity of distamycin for quadruplex systems over duplex systems and the lack of evidence supporting targeting biologically significant quadruplex systems, such as those formed in telomeres and proto-oncogenes, suggests that polyamides are not good paradigms for quadruplex-interactive agents.

In a recent exciting finding from our laboratory, DB832 (Figure 4.1: H), a heterocyclic diamidine, was shown to strongly and selectively interact with quadruplex forming sequences of telomeric DNA. Even though the compound encompassed all those structural elements that are explicitly used to characterize duplex minor-groove binders, DB832 exhibited a very high degree of selectivity for telomeric quadruplex-DNA and virtually zero affinity for a wide range of duplex-DNAs. Moreover, DB832 exhibited a unique CD signature that has been reported for small molecules that can form stacked complexes with quadruplex-DNA [23]. DB832 was shown in Chapter 2 to selectively interact with hybrid-type telomeric quadruplexes as stacked species and most likely stacked in the multiple accessible grooves. The oxygen atoms of the two furan ring systems were hypothesized to potentially participate in H-bonding interactions with the G-NH_2 groups positioned in the quadruplex grooves, and the amidine moieties of DB832 to potentially recognize the N3 atoms of guanine, creating additional quadruplex recognition sites. To gain a better understanding of all the structural elements of DB832 contributing towards quadruplex groove recognition, Dr. David Boykin and coworkers from Georgia State University developed a series of structurally similar aromatic diamidines (Figure 4.2) using DB832 as the prototype. Particular emphasis was directed towards compounds that maintained

the core structure of DB832: the 5-5-6 ring systems. Systematic atom-wise and group-wise modifications were undertaken on the 5-5-6 ring systems of DB832 that were hypothesized to be important for the compound binding. Switching the order of components in DB1093, for example, significantly changes the molecular shape. Converting one or two furans to thiophenes (DB1463, DB1438 and DB1450) makes subtle shape changes but also can affect stacking. The furan-thiophene change has been shown to have a significant effect in duplex groove binding systems [24]. Ortho or meta nitrogen substitution on the phenyl ring (DB934 and DB1693) or on the central furan ring (DB1999) of DB832 modifies stacking ability and can have potential to participate in hydrogen-bonding interaction in the grooves of G-quadruplex. Insertion of a methyl group to the terminal furan (DB1949) provides changes in stacking ability but also a steric restraint to that region of the molecule. A freely rotatable triple bond addition between the phenyl and the furan (DB1694) is expected to further weaken duplex interactions and potentially enhance quadruplex binding. Finally, the role of the amidine groups can be probed by converting them to imidazolines in DB832 (DB1972) as well as in a thiophene (DB2037). The imidazoline substitution will provide planarity to the central system and maintain a doubly charged state at physiological pH. To help understand the quadruplex interactions of the new compounds, two central molecules, which have been reported in literature, were used as reference compounds. One, RHPS4 is known to bind to the human telomeric quadruplex conformation at the terminal G-tetrads of the human telomeric sequence [25]. The other, distamycin-A, was recently assigned as a G-quadruplex groove binder after its NMR evaluation with a parallel four-stranded intermolecular quadruplex [22].

An array of biophysical techniques was used to characterize the binding of these ligands to the telomeric quadruplex-DNA and also duplex-DNA. Thermal melting results showed these compounds bind to quadruplex-DNA with significant selectivity over duplex sequences. Interestingly, NMR and CD results indicated a mixed binding mode with an initial strong binding, through stacking at either one or both ends of the terminal G-quartets,

followed by a weaker binding, through groove-recognition. Surface plasmon resonance studies on selective compounds show that the compounds exhibited binding affinities that are commonly observed for most of the quadruplex end-stacking compounds: in the order of 10^6 - 10^7 M^{-1}. The binding affinities for the subsequent binding phenomena, proposed to be groove-recognition, could not be accurately determined with the available simple binding models, but were generally found to be over 10-fold weaker. Nevertheless, we have identified a series of heterocyclic diamidines, derived from duplex groove binders, which can potentially recognize quadruplex architecture and can be used as lead candidates to design compounds with improved selectivity and affinity.

4.2 MATERIALS AND METHODS

4.2.1 Sample Preparation

The biotin-free and 5′-biotin labeled DNA sequences Tel22, d[AGGG(TTAGGG)$_3$]; Tel24, d[TTGGG(TTAGGG)$_3$A]; Dickerson, d[CGAATTCGT$_4$CGAATTCG]; GC20, [(CG)$_4$TTTT(CG)$_4$]; were purchased with HPLC purification and mass spectrometry characterization from Integrated DNA Technologies (Coralville, IA) and from Midland Certified Reagent Company. NMR analysis performed on these sequences further confirmed the purity of these sequences. The concentration of oligonucleotides was determined from absorbance at 260 nm using molar extinction coefficient obtained from nearest-neighbor principle. The syntheses of DB compounds were performed by Dr. David Boykin's group at Georgia State University and will be reported elsewhere. Appropriate stock solution of each compound was prepared in double deionized water and in deuterated water for NMR experiments. This stock solution was diluted to required concentrations with appropriate buffer right before their usage.

4.2.2 Circular Dichroism Experiments

All CD experiments were performed at 25 °C in 10 mM K_2HPO_4 buffer containing 80 mM KCl and 1 mM EDTA. CD spectra were recorded using a Jasco J-810 spectrophotometer with a 1-cm pathlength quartz cell at a scan speed of 50 nm/min and response time of 1 second. Appropriate amount of compounds were sequentially titrated from the stock solution into the DNA solution in the cuvette until the desired mole ratios of compound to quadruplex were obtained. The spectra were averaged over four scans. A buffer baseline scan was collected in the same cuvette and subtracted from the average scan of each ratio. Data were processed and plotted using Kaleidagraph 4.0 software.

4.2.3 Thermal Melting Studies

Thermal denaturation studies were conducted on a Cary 300 BIO UV-visible spectrophotometer in quartz cuvettes of 1 cm pathlength. Compound-DNA solutions were prepared in low salt buffer containing 10 mM TRIS buffer (pH 7.4), 10 mM KCl and 1 mM EDTA. The absorbance of the oligonucleotides was monitored at the recommended wavelength of 295 nm for quadruplex sequences and 260 nm for duplex sequences as a function of temperature, and DNA without compound was used as a control. Samples of compound to DNA ratios ranging from 0:1 to 4:1 were prepared. Cuvettes were mounted in a thermal block, and the solution temperatures were monitored by a thermistor in a reference cuvette with a computer-controlled heating rate of 0.5 °C/min. Experiments were generally conducted at quadruplex concentrations in the range of 2-3 μM in TRIS buffer containing 10 mM KCl. Data were analyzed and plotted using Kaleidagraph 4.0 software.

4.2.4 Surface Plasmon Resonance Studies

Biosensor SPR experiments were performed with a four-channel BIAcore 2000 optical biosensor system (BIAcore, Inc.) and streptavidin-coated sensor chips (BIAcore SA with linked streptavidin). All DNA samples, for either duplex- or quadruplex-binding experiments, were

used as single strands to prevent dissociation in the SPR flow system. The chips were prepared for use by conditioning with a series of 1 min injections of 1 M NaCl in 50 mM NaOH followed by extensive washing with buffer. 5'-Biotinylated DNA samples (25-50 nM) in HBS buffer were immobilized on the flow cell surface by non-covalent capture as previously described [26]. Three flow cells were used to immobilize DNA samples, and the first flow cell was left blank as a control. Interaction analysis was performed by using steady-state methods with multiple injections of increasing compound concentrations over the immobilized DNA surface at 25 °C. Biosensor experiments were conducted in filtered, degassed HEPES buffer (10 mM HEPES, 100 mM KCl, 3 mM EDTA, 0.005 v/v of 10% P20 BIACORE surfactant, pH 7.3) at 25 °C. Flow cell 1 was left blank as a reference, while flow cells 2-4 were immobilized with DNA on a streptavidin-derivatized gold chip (SA chip from BIAcore) by manual injection of DNA stock solutions (flow rate of 1 μL/min) until the desired value of DNA response was obtained (350-400 RU). Compound solutions were prepared with the running buffer by serial dilutions from stock solution. Typically, a series of different ligand concentrations (1 nM to 10 μM from 20 mM H_2O stock) were injected onto the chip (flow rate of 50 μL/min, 5-10 min) until a constant steady-state response was obtained followed by a dissociation period (buffer, 10 min). After every cycle, the chip surface was regenerated (20 s injection of 10 mM glycine solution, pH 2.0) followed by multiple buffer injections.

The instrument response (RU) in the steady-state region is proportional to the amount of bound drug and was typically determined by linear averaging over a 10-20 s or longer time span, depending on the length of the steady-state plateau. The predicted maximum response per bound compound in the steady-state region (RUmax) was determined from the DNA molecular weight, the amount of DNA on the flow cell, the compound molecular weight, and the refractive index gradient ratio of the compound and DNA, as previously described [27]. In most of the cases, the observed RU values at high concentrations were greater than RUmax,

pointing to more than one binding site in these DNA sequences. The number of binding sites was estimated fitting plots of RU versus C_{free}. These methods can also be used to determine an empirical RUmax value. The RUmax value is required to convert the observed response (RU) to the standard binding parameter *r* (moles of drug bound per moles of DNA hairpin)

$$r = RU/RUmax$$

which is useful for comparison of a compound binding to different DNAs to obtain the binding constants, the data were evaluated with different interaction models to obtain an optimal fit using BIAevaluation (BIAcore Inc.) and Kaleidagraph (Synergy Software) software for nonlinear least-squares optimization of the binding parameters:

One site: $r = (K_1C_{free})/(1 + K_1C_{free})$

Two site: $r = (K_1C_{free} + 2K_1K_2C_{free}^2)/(1 + K_1C_{free} + K_1K_2C_{free}^2)$

Three site: $r = (K_1C_{free} + 2K_1K_2C_{free}^2 + 3K_1K_2K_3C_{free}^3)/(1 + K_1C_{free} + K_1K_2C_{free}^2 + K_1K_2K_3C_{free}^3)$

where K_1, K_2 and K_3 are equilibrium constants for three types of binding sites and C_{free} is the concentration of the compound in equilibrium with the complex and is fixed by the concentration in the flow solution.

4.2.5 Nuclear Magnetic Resonance Studies

Quadruplex DNA samples were prepared in degassed phosphate buffer, containing 80 mM KCl, 10 mM K_2HPO_4 and 0.1 mM EDTA, and reconstituted in 90% H_2O:10% D_2O. 0.01 mM DSS was employed as an internal reference. DNA concentrations were in the range of 0.1 mM to 0.3 mM unless otherwise mentioned. The final DNA samples were adjusted to pH 7.0 using 1M HCl or 1M KOH solutions. Finally, the NMR samples were heated past their transition temperature and annealed to room temperature several times before collecting the spectra. Experiments were performed on a Varian Unity 600 spectrometer. DB compounds were titrated into the quadruplex DNA with compound to DNA ratios varying from 1 to 3.

Temperature-dependent ^{1}H-spectra were recorded from 15 ^{o}C to 45 ^{o}C using jump-return and WATERGATE methods for solvent suppression [28]. All NMR data were processed and analyzed with a combination of VNMR (Varian Inc.), NMRPipe (NIH) and MNova (Mestrelab Research) software.

4.3 RESULTS AND DISCUSSION

4.3.1 Compounds are Highly Selective for Quadruplex-DNA over Duplex-DNA

With the abundance of duplex DNA in the genome, targeting relatively few non-duplex structures such as quadruplexes becomes a complex task. Non-selective binding to secondary sites might result in deleterious effects and cytotoxicity, therefore, structure and sequence selectivity is of utmost importance to achieve desired results. In order to determine the selectivity of DB compounds for human telomeric quadruplex sequence over duplex DNA sequences, thermal melting studies were performed with telomeric quadruplex DNA (Tel22: Materials and Methods, 4.2.1) and two different duplex sequences (AATT, GC20: Materials and Methods, 4.2.1). The AT-rich duplex sequence was chosen since small molecules show a preferential binding towards sequences with a narrow minor groove, and a GC-rich duplex sequence was chosen to mimic the wider grooves of quadruplex DNA and, also, DB293, as aforementioned, has been shown to cooperatively bind as a stacked dimer in GC- containing sequences. This would also eliminate any bias towards any particular groove width for the tested compounds. Melting curves were obtained for quadruplex sequences up to ratios as high as 6:1 compound per quadruplex and ratios as high as 4:1 for duplex sequences. The ΔTm values for quadruplex (2:1, 4:1 ratios) and duplex sequences (2:1 ratio) are reported in Table 4.1; ratios higher than the listed range resulted in the signification complex aggregation and produced inconsistent melting curves. This high stoichiometry was used because, as aforementioned, the recently solved NMR structure of distamycin bound to an intermolecular

quadruplex structure shows that up to four distamycin molecules are bound in the opposite grooves of the quadruplex. Therefore, it is a likely possibility that relatively small-sized molecules that mimic distamycin can bind to quadruplex DNA with a high stoichiometry. The melting temperature of Tel22 in 100 mM K+ is 48.8 °C. **DB832**, the paradigm compound, showed a ΔTm of 19.1 °C at 2:1 and an impressive 26.1 °C at 4:1 ratios indicating signification stabilization of the quadruplex conformation. Also, with the two duplex sequences, Dickerson and GC20, DB832 shows very little increase in the Tm (3 °C and 3.7 °C respectively at 2:1 ratio) suggesting high selectivity over duplex DNA. Interestingly, **DB1093**, a more linear structural isomer of DB832, showed almost no change in ΔTm at 2:1 ratio and a small increase (5 °C) at 4:1 ratio highlighting the importance of the neighboring furan ring systems for effective recognition of quadruplex architecture. DB1093 also exhibited very little change in ΔTm for duplex sequences (Table 4.1) suggesting that the compound is too linear to interact with duplex DNA minor groove. The substitution of the two furan rings with thiophenes, **DB1450**, results in very similar quadruplex stabilization (2:1-17.6 °C, 4:1-24.2 °C) as DB832 but, unfortunately, this modification decreases the specificity, with a ΔT_m of almost 9 °C, for the duplex sequences. Affinity for duplex sequences decreases by an impressive 6°C, relative to DB1450, when the terminal furan is replaced by a thiophene (**DB1438**), while maintaining strong quadruplex stabilization potential (4:1-25.5 °C). Interestingly, the substitution of the central furan by a thiophene, **DB1463**, significantly increases the binding to AATT sequence (8.5 °C), but not the GC20 sequence (3.4 °C), suggesting the effect of compound curvature on recognizing different groove dimensions. However, the quadruplex stabilization for **DB1463** is not very different (2:1-17.3 °C, 4:1-23.1 °C) from what was observed for DB832. The addition of a freely rotatable triple bond between the phenyl and furan ring systems, **DB1694**, also significantly decreases the affinity for duplex sequences (AATT-1.0 °C, GC20-2.1 °C) while maintaining good stabilization properties with quadruplex DNA. This has been previously

observed in the case of diaryl amides where the addition of the acetylene group has increased quadruplex recognition potential [29]. The nitrogen substitution on the phenyl ring of DB832 exhibited enhanced stabilization of the quadruplex arrangement when it is placed in the meta position (**DB1693**) but slightly decreased affinity when it is in ortho (**DB934**). The conversion of amidine groups to imidazolines (**DB1972** and **DB2037**) yields a more planar arrangement while still maintaining a dicationic charge system. These modifications provide an increased stacking surface available for ligand interaction with the terminal tetrads, and might have higher quadruplex recognition potential. The imidazoline analogs, however, exhibited generally lower ΔTm than amidine counterparts. The oxazole derivative of DB832, **DB1999**, and the addition of a methyl group on terminal furan, **DB1949**, also had high quadruplex stabilization properties while and very little affinity for duplex sequences under low salt conditions. The reference compound, **distamycin**, recently proposed as a G-quadruplex groove binder with the four-stranded d(TG_4T) at NMR concentrations, did not show any ΔT_m with the human telomere quadruplex sequence. As it is well known, however, that the polyamide is a strong binder to AT sequences in duplex DNA, distamycin is the only compound of this group that has reverse selectivity for duplex over quadruplex. Finally, **RHPS4**, a well-studied quadruplex stabilizing compound that has been reported to bind by π-π interactions with the external tetrad, shows melting values lower than to DB832 in terms of stabilization properties and selectivity for quadruplex over duplex DNA. The quadruplex stabilization potential observed for some of the DB compounds ranks among the strongest for quadruplex-interacting small molecules reported so far. It is highly apparent that the ligands are significantly stabilizing the quadruplex conformation, while maintaining a low degree of selectivity for duplex sequences. Therefore, these compounds make ideal candidates for further study as highly selective quadruplex interacting agents.

4.3.2 DB Compounds Bind to the Human Telomeric DNA with Multiple Binding Modes

Circular dichroism (CD) studies were performed to evaluate how modifications on the DB832 system influence the DNA conformation in the complex and the ligand interaction mode with the human telomere. Circular dichroism has emerged as an important non-invasive technique for determination of the conformation of biomolecules and also provides insights about the binding modes of small molecules with DNA using pattern recognition [30]. When an achiral ligand, which has no CD by itself in solution, binds to a chiral macromolecule such as DNA, an induced CD signal is observed in the wavelength region corresponding to the bound achiral ligand [31]. This induced CD (ICD) signal can be a small positive or negative, as in the case of duplex intercalators, or can be a large positive signal, as in the case of duplex groove binders [32]. A combination of both positive and negative CD signals, coupled with different ICD signal shapes indicates formation of stacked complexes, as was previously shown with DODC with quadruplex systems [33] (Chapter 2, Figure 2.3-A). Recent CD studies have shown that DB832 binds to the human telomeric quadruplex DNA possibly in the grooves of the quadruplex conformation with large positive (435 nm) and negative (390 and 418 nm) ICD signals at the wavelength region corresponding to the absorbance of bound DB832 [23]. Moreover modifications in the DNA region driven by the compound indicate stabilization of a quadruplex conformation.

Using the ICD pattern exhibited by DB832 as a model for quadruplex multiple-site recognition (Figure 4.3-A), CD studies were conducted for the compounds in Figure 4.2 with Tel24 to investigate if a pattern similar to DB832 exists. Figures 4.3-4.5 show CD spectra of the DB compounds with the Tel24 sequence. The addition of a methyl group on the terminal furan (**DB1949**, Figure 4.3-B) and the substitution of a central furan with an oxazole (**DB1999**, Figure 4.3-C) yielded ICD patterns similar to DB832. These modifications apparently did not significantly affect the interaction mode with Tel24. The only difference

that emerges by comparing the spectra is the magnitude of the signal at the ICD signal saturation ratio. However, **DB1093**, the linear isomer of DB832, exhibits a different pattern with Tel24 sequence (Figure 4.3-D). The positive band below 350 nm and negative ICD signal above 350 nm with relatively low intensity is characteristic of ligands intercalating with duplex-DNA or end-stacking with quadruplex-DNA. Therefore, DB1093 might be an end-stacking molecule with quadruplex systems with relatively low ΔTm (Section 4.3.1, Table 4.1).

Tm studies of thiophene containing diamidines (DB1438, DB1450 and DB1463: Section 4.3.1) showed that all the compounds have quadruplex stabilization potential with varying degrees of duplex selectivity. Sulfur is not as electronegative as oxygen and therefore has weaker hydrogen bonding interactions than oxygen. The two monosubstituted thiophene analogs, **DB1438** (Figure 4.4-B) and **DB1463** (Figure 4.4-C) exhibit a smaller magnitude of ICD signals relative to DB832, which may be due to decreased stacking interactions due to a small change in the curvature of the compounds. However, the disubstituted **DB1450** has a very similar exciton splitting as DB832 suggesting a very similar binding mode.

DB1972 and **DB2037** have structural characteristics that make them unique: the presence of two imidazoline rings instead of the two amidine groups. These rings affect the equilibrium structure of the compound because the imidazoline relative to an amidine reduces the torsional angle with the neighbor ring. This feature provides a more planar shape to the molecule. This kind of shape suggests a more favorable end stacking interaction with the external tetrads of the quadruplex. The ICD shape looks different from all the other compounds spectra and it shows small positive and negative peaks (Figure 4.5: A, B). Interestingly, both the compounds exhibit very similar saturation ratios indicating very similar interaction modes and possibly with a very similar stoichiometry. **DB934**, with ortho nitrogen substitution on the phenyl ring, exhibited virtually no ICD even at very high ratios (Figure 4.5-

C), and, **DB1693**, with meta nitrogen substitution, exhibited a weak excition-type splitting (Figure 4.5-D). The fact that substitution of a single atom makes a large change in compound ICD signal shows that recognition of the quadruplex as a stacked species is extremely sensitive to the compound structure. Tm studies with these two compounds (Section 4.3.1) have shown that these compounds also have great quadruplex stabilization potential, suggesting an end-stacking mode of interaction. Most of the compounds in this study had a very little effect on the quadruplex conformation as can be seen from the CD signals corresponding to DNA regions (Figure 4.3-4.5). NMR studies have shown that the sequence, Tel24, folds into a mixed parallel/antiparallel hybrid-1 conformation in K^+ [34]. The very small effect on DNA conformation upon ligand binding indicates that the DB compounds are preferentially binding to an already preformed hybrid structure.

Interestingly, closer inspection of the ICD signal exhibited by most of the compounds reveals a pattern emerging upon complex formation. Figure 4.6 shows a plot of ICD signal at the wavelength corresponding to the maximum CD signal of different compounds versus mole ratio of the complex. A very weak ICD signal is observed for all of the compounds until the ratio 3:1, followed by a dramatic increase in ICD above ratios 3:1 for DB832, DB1949 and DB1450 (Figure 4.6). Other compounds have weaker ICD signals that are less than the original DNA CD signal. A weak induced CD signal with duplex DNA is generally observed for intercalation-type mechanisms [32f]. However, with quadruplexes, intercalation is not commonly observed due to the high energetic cost required to replace the coordinating cations along the quadruplex helical axis and also due to the conformational restraints posed by the terminal diamidine units of the DB compounds [35]. Therefore, intercalation has been ruled out as the possible binding mode for the compounds. The observed weak ICD signal at low ratios is consistent with the interaction of the ligands at either one or both the ends of the quadruplex. NMR studies (discussed later in Section 4.3.3) further confirm the primary end stacking mode

observed for most of the compounds at low ratios. Also, the aromatic moieties of these ligands can favorably stack on the G-tetrads to form optimal π-π interactions. At ratios higher than 3:1, some compounds appear to be forming stacked complexes probably in different grooves of the quadruplexes, characterized by exciton-type ICD signals. To summarize, the CD results, at ratios until 3:1 compound per quadruplex, ligands exhibit initial mode of binding by stacking on either one or both ends of the quadruplex structure, whereas, at ratios higher than 3:1, DB832, DB1949 and DB1450, form stacked species probably in the grooves of the quadruplex.

4.3.3 Multiple Binding Modes Observed in NMR

Telomeric quadruplex sequences in K^+ exists as a mixture of conformations, with the hybrid conformation as a predominant structure [34, 36]. An NMR structure of Tel24 show that this sequence folds into a mixed parallel/antiparallel hybrid-1 type structure [34, 37]. To test the idea that multiple binding mode are exhibited by some of the compounds from CD, ^{1}H-NMR titration studies were conducted with Tel24. Imino protons of guanines are excellent probes to determine if a G-rich sequence folds into a quadruplex conformation [38]. These imino protons resonate between 10.0 to 12.5 ppm in a quadruplex structure and are readily observable in water sample. Moreover, the presence of multiple quadruplex conformations can also be readily answered by monitoring the imino protons. Binding of a small molecule to a quadruplex structure affects the chemical environment of imino protons which can also be identified from simple titration experiments. Figures 4.7-4.8 show the guanine imino proton NMR spectra of different DB compounds with the Tel24 sequence at 25 °C. At higher compound to quadruplex ratios, particularly at NMR concentrations, solubilities of the compound and complex become an issue. Considerable aggregation is observed at high compound concentrations resulting in significant line broadening and rendering it difficult to obtain spectra for analysis. In the absence of any ligands (Figures 4.7-4.8, 0:1 ratio), the

spectra show distinct numbers of imino protons corresponding to the total number of guanines in the Tel24 sequence. This shows that the sequence forms primarily a single structure in K^+, and in this case it is the hybrid-1 type structure. Titration of the compounds with Tel24 resulted in significant changes in the imino proton spectra and spectral broadening and shifts revealed that the ligands were in intermediate exchange on the NMR timescale. A number of imino resonances broaden in a site-specific fashion, consistent with the local chemical environment of various residues being perturbed differentially by ligand binding. Distinct chemical shift changes can be readily observed for most of the imino protons. Imino protons of G3, G9, G17 and G21, that constitute the 5' terminal tetrad of the quadruplex, are the most perturbed even at the lowest compound to quadruplex ratio. This suggests the inherent sensitivity of this technique even for the smallest changes produced, at relatively low ratios, upon complex formation.

The 5' G-tetrad imino protons broadened considerably and disappeared completely at high ratios. The high perturbation of the imino protons at the top end is consistent with the stacking of the ligand on the 5' G-quartet. NMR studies of distamycin, Hoechst-33258, and ethidium with several quadruplex forming sequences has revealed similar perturbations of the imino protons of the top tetrad of the quadruplex and can stack on the top end of the quadruplex indicating stacking on the top end of the quadruplex [38-39]. Small molecules that have excellent duplex-DNA groove binding properties, such as Hoechst-33258 and distamycin, can preferentially bind as end-stackers with quadruplex systems. Therefore, it is equally conceivable for the currently studied ligands to exhibit a similar end-stacking as the primary mode of binding. Interestingly, broadening occurs for the aromatic residues of the 5' G-tetrads (G3, G9, G17 and G21 of 5'-tetrad) and A20·T1 base pair (spectra not shown), providing support for a mixed intercalation-exterior stacking mode of binding (i.e., between

the 5'-tetrad and the capping A20·T1 base pair of the loops). Therefore the results agree with the stacking of the ligands on the top quartet at low ratios as suggested from CD studies (Section 4.3.2). In most of the cases imino protons of G11and G23, that constitutes the bottom tetrad, exhibited very little change. This is probably due to the presence of a stable capping A24·T13 base pair at the 3' terminal, offering significant protection from solvent exchange. Also, relatively unchanged imino proton chemical shifts of the bottom tetrad indicates that 3' end of the Tel24 quadruplex structure is not likely the binding site for these compounds. In all the cases, imino protons of G4 and G22 that correspond to the adjacent guanine of the middle tetrad exhibit small downfield shifts accompanied by slow broadening; however, imino protons of G10 and G16 of the same tetrad exhibit very little changes. G4 and G22 are located in a groove that has optimal dimensions for compounds to form stacked species. Therefore, after the initial end-stacking mode at the 5' terminal, there is a very likely possibility that the ligands are targeting the "medium" groove encompassed by the G3, G5, G23, G21 plane. The groove adjacent to the medium groove is classified as the narrowest. Therefore, it is highly unlikely that the compounds can efficiently form stacked species in the narrow groove. Absence of any major changes in imino protons of guanines (G16 and G15) further confirms this. Therefore NMR results are consistent with the recognition of multiple sites by DB compounds at high ratios, thereby, agreeing with the CD results.

4.3.4 Surface Plasmon Resonance Study Reveals Multiple Binding Events

Surface plasmon resonance (SPR) has developed into an effective tool to measure the affinity and study the kinetics of most biomolecular interactions [26]. A recent literature survey has shown that more and more quadruplex-small molecule interactions are being evaluated using SPR techniques due to the high quality results obtained using SPR methodology [40]. In order to obtain qualitative binding affinities of the compounds in Figure 4.1 with the telomeric

quadruplex sequence, a series of steady-state SPR experiments were conducted (Materials and Methods, 4.2.4). Figures 4.9 and 4.10 show sensorgrams for several compounds with Tel22 sequence. From the sensorgrams, the large response obtained suggests that the compounds are binding to the Tel22 sequence, and both the compounds exhibit fast association and dissociation kinetics. Table 4.2 lists the binding constants obtained using a two-site model by plotting the steady-state RU values as a function of free ligand concentration for different compounds. From the table, we can readily notice that the primary binding constant (K_1) of different compounds fall in the range of 8 x 10^5 to 3 x 10^6 M^{-1}, and is followed by a weak binding constant (K_2) that is 10-100 fold weaker in most cases. Although the primary binding constants obtained for different compounds are not high, it does fall in the range of binding constants that is generally observed for compounds that interact at the ends of the quadruplex[23].

As discussed earlier, NMR and CD results have shown that the recognition of quadruplex conformation by the DB compounds occur by more than one type of binding mechanism. The initial binding observed at low ratios from NMR and CD is due to the end-stacking of the ligands, whereas, binding events occurring at high ratios is possibly due to groove recognition. Excellent results can be obtained using the SPR technique at low ligand concentrations and so the first binding event was easily detectable. As the concentration range is increased to follow the second binding, the free compound adsorbs on the chip surface resulting in curves that are difficult to fit. This can be readily observed in the sensorgrams of DB1450 (Figure 4.8, top right) and DB1463 (Figure 4.8, top middle), and their corresponding fitting. The cooperative binding that gives the large ICD signal could not be accurately detected by this method for most compounds. The K_1 and K_2 binding constants obtained from SPR were determined by fitting only the initial parts of the binding curve with a two-site model. The entire binding curve could not be fitted with either of the models due to the complex shape of

the curve. The binding affinities obtained by fitting the initial part of the binding curve might be due to the binding of one or more ligands to the ends of the quadruplex structure at low ratios.

4.4 CONCLUSION

The current studies suggest that the binding mode of these compounds is very sensitive to the subtle changes in compound. Here, we have identified a small subset of heterocyclic diamidines that can recognize quadruplex DNA conformation with high selectivity and the recognition of telomeric quadruplex DNA may occur by more than one type of binding mechanism. Although end stacking is what was found to be the most common binding mechanism for these compounds, the secondary binding that we believe is occurring, by stacking in the grooves of the quadruplex, is a much more interesting phenomenon. It is this groove binding phenomenon that we are more actively pursuing, since targeting grooves of the quadruplex with small molecules has become an important strategy and has posed great challenges in recent years for drug design and development. Stacked species are of particular interest for recognition of quadruplexes since studies with duplex DNAs show that compounds which bind as stacked dimers have increased binding affinity and selectivity over similar compounds that bind as monomers. Similar stacking in the grooves of quadruplex DNA structures would appear to be a favorable way to selectively recognize quadruplexes with optimum interactions, perhaps employing an induced fit component between the stacked heterocycles and the guanine bases of the quadruplex tetrads. Potentially, the individual monomer units could be covalently linked, dramatically increasing the affinity and selectivity of the compound for human telomeric DNA, leading to enhanced telomerase inhibition with decreased cytotoxic side effects.

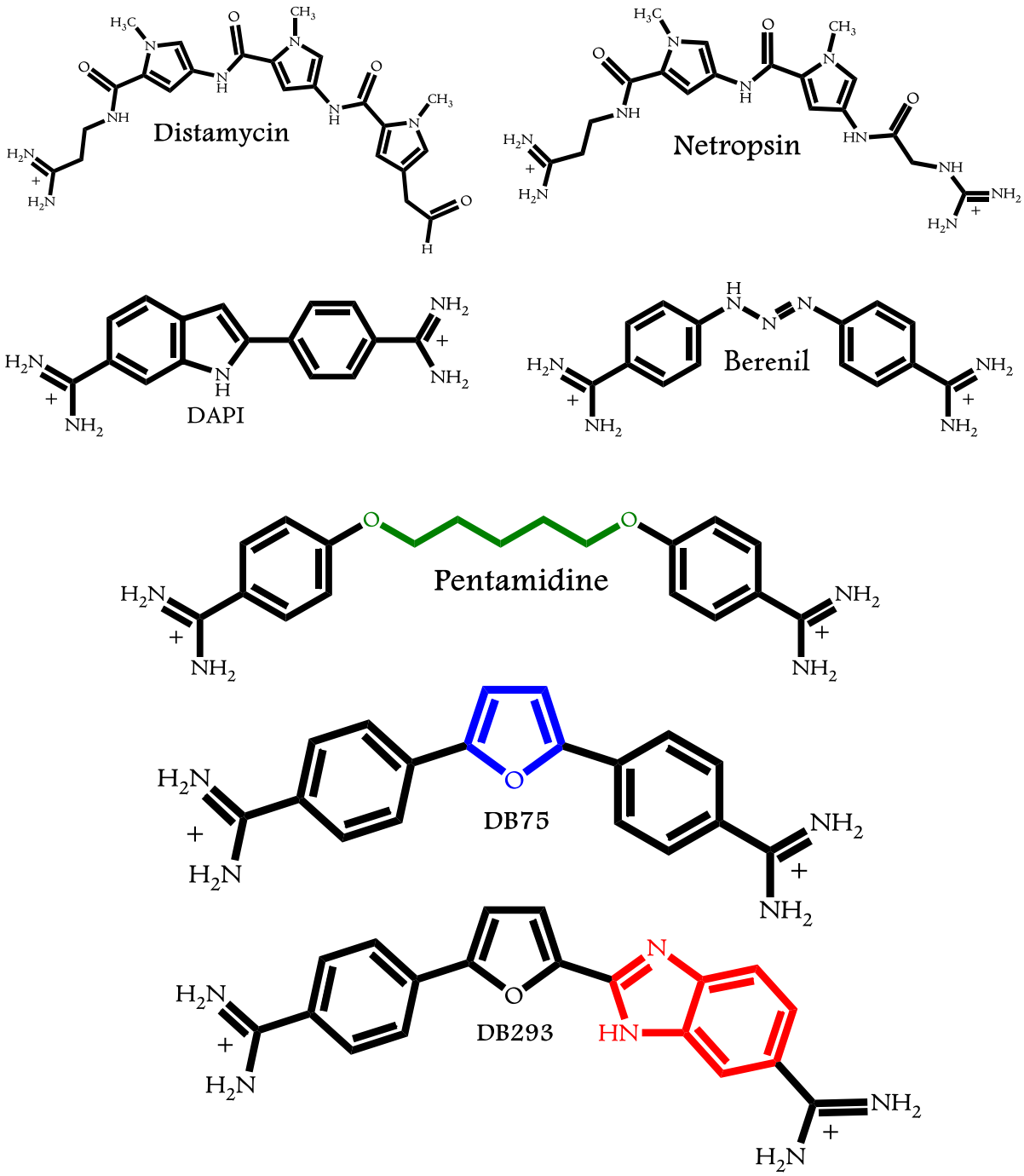


Figure 4.1: Chemical structures of representative Duplex-DNA minor groove binders. (A) Distamycin, (B) Netropsin, (C) DAPI, (D) Berenil, (E) Pentamidine, (F) DB75, and (G) DB293.

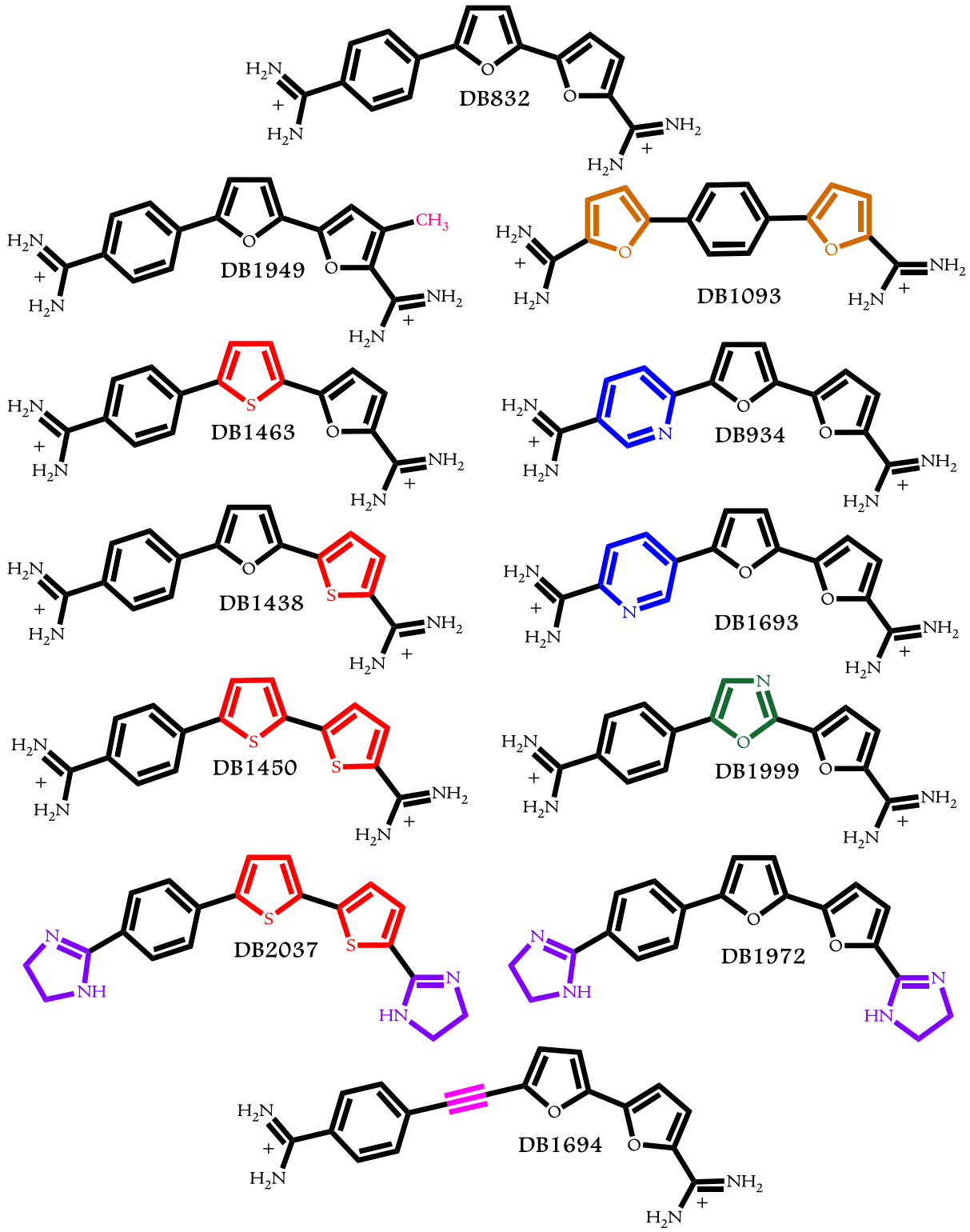


Figure 4.2: Chemical structures of 5-5-6 and related ring systems used in this study.

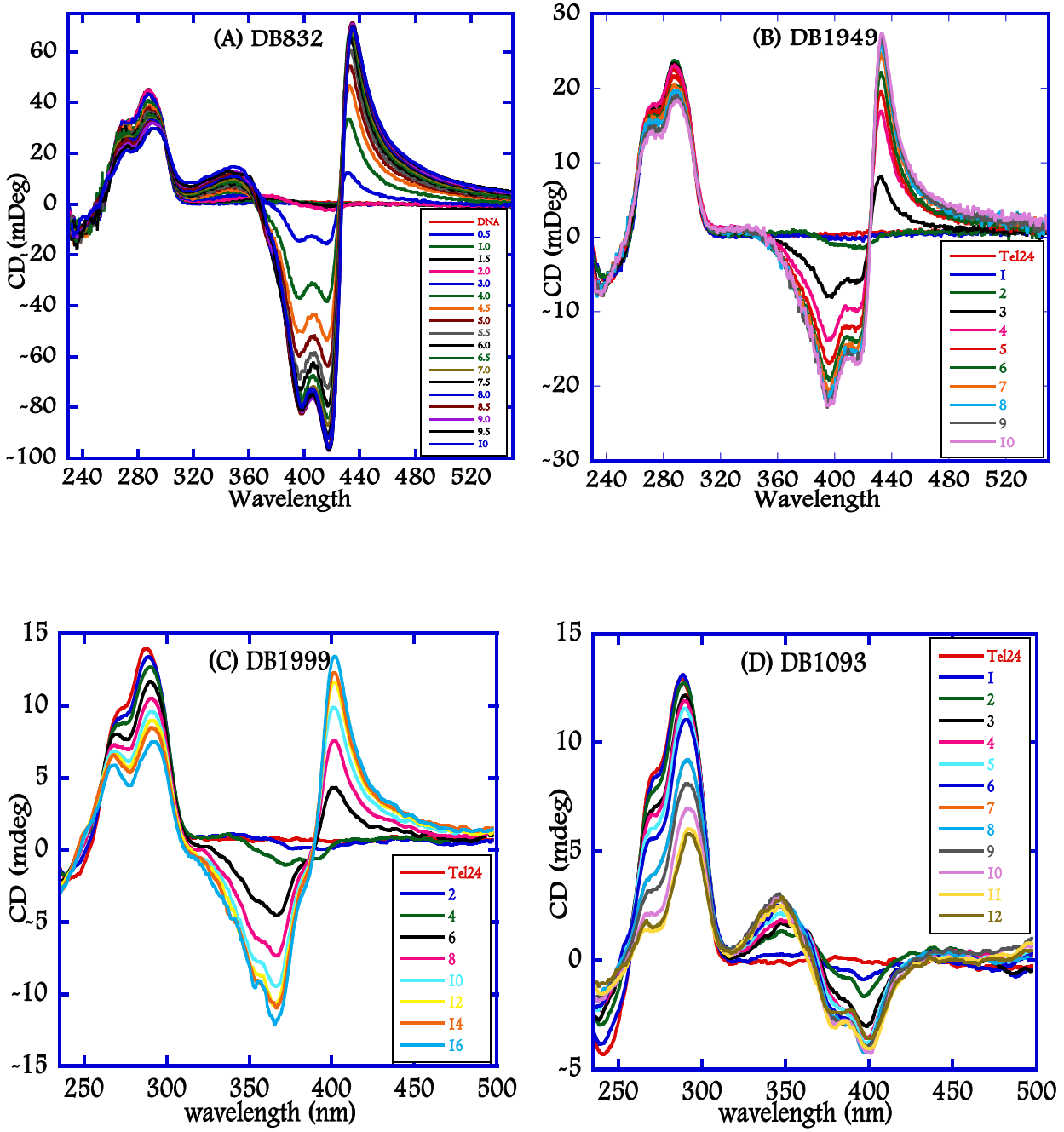


Figure 4.3: CD spectra of DB compounds (A) DB832, (B) DB1949, (C) DB1999, and (D) DB1093 with Tel24 quadruplex sequence.

80 mM KCl/10 mM K_2HPO_4 buffer at 25 °C. Single strand quadruplex concentration is 4-5 μM unless otherwise mentioned. Inset in each case shows the compound/quadruplex ratio. (A) [Tel24] = 10 μM

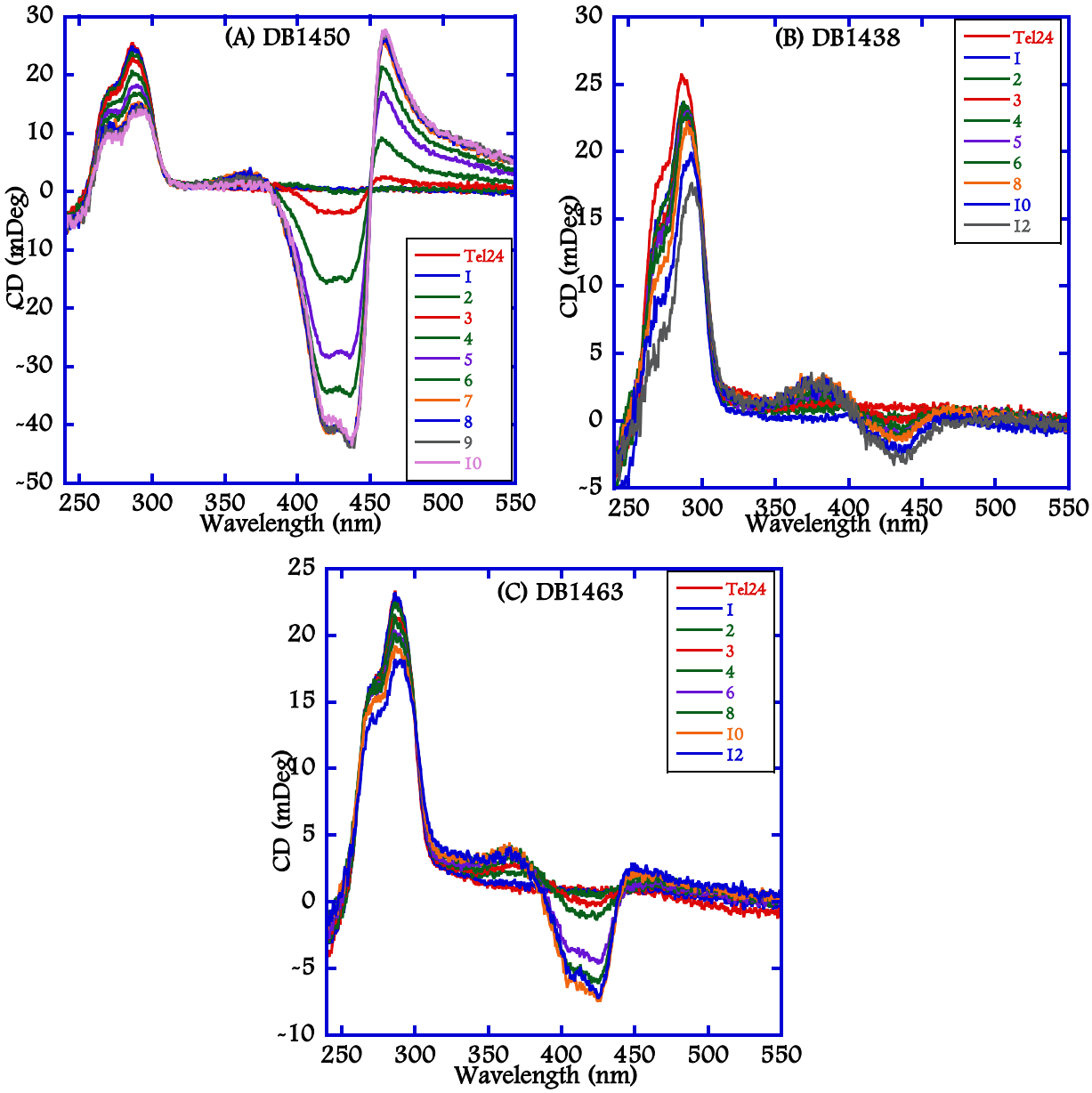


Figure 4.4: CD spectra of thiophene containing DB compounds (A) DB1450, (B) DB1438, and (C) DB1463 with Tel24 quadruplex sequence.

80 mM KCl/10 mM K_2HPO_4 buffer at 25 °C. Single strand quadruplex concentration is 4-5 μM unless otherwise mentioned. Inset in each case shows the compound/quadruplex ratio.

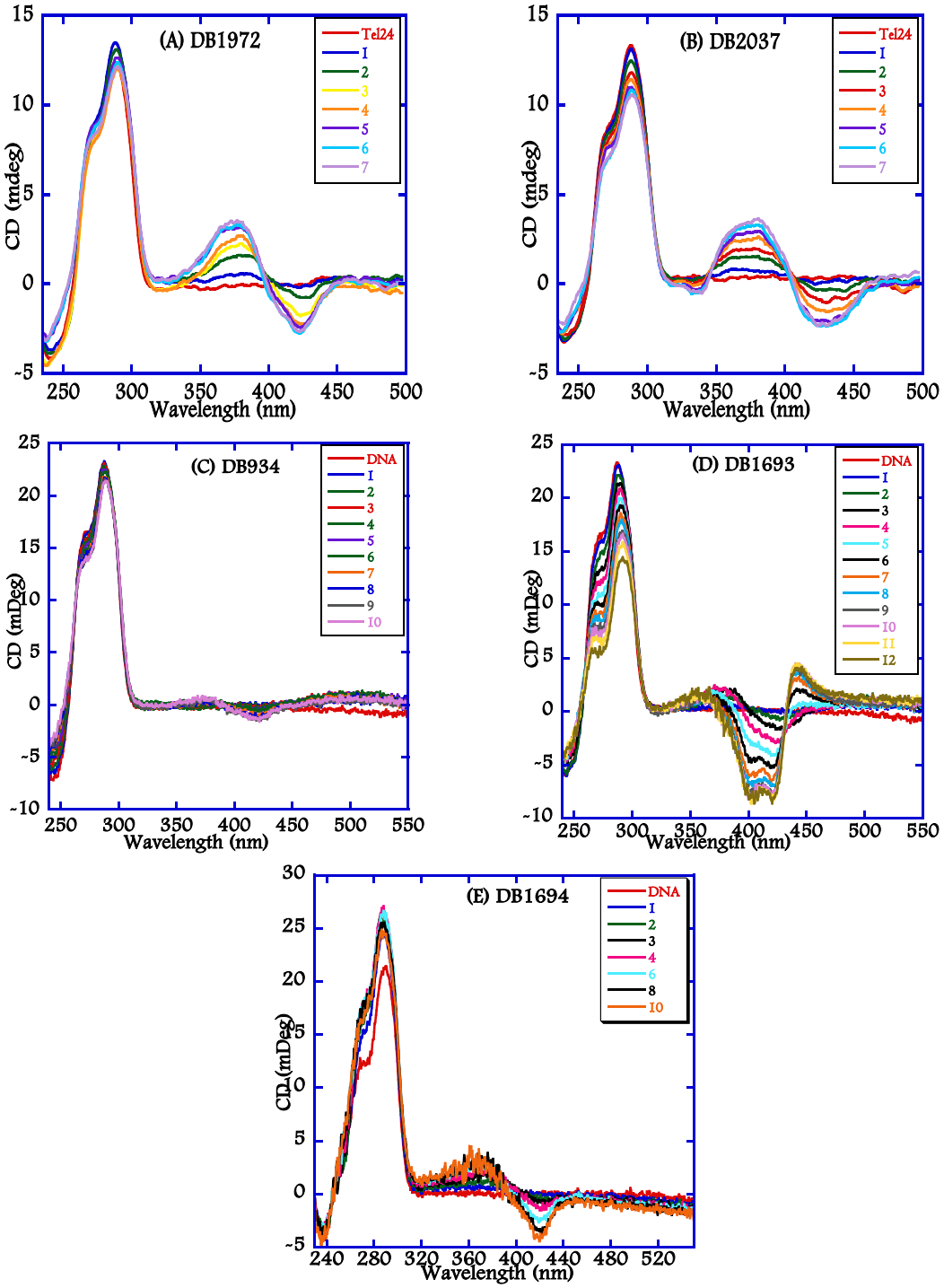


Figure 4.5: CD spectra of DB compounds (A) DB1972, (B) DB2037, (C) DB934, (D) DB1693, and (E) DB1694 with Tel24 sequence.

80 mM KCl/10 mM K_2HPO_4 buffer at 25 °C. Single strand quadruplex concentration is 4-5 μM unless otherwise mentioned. Inset in each case shows the compound/quadruplex ratio.

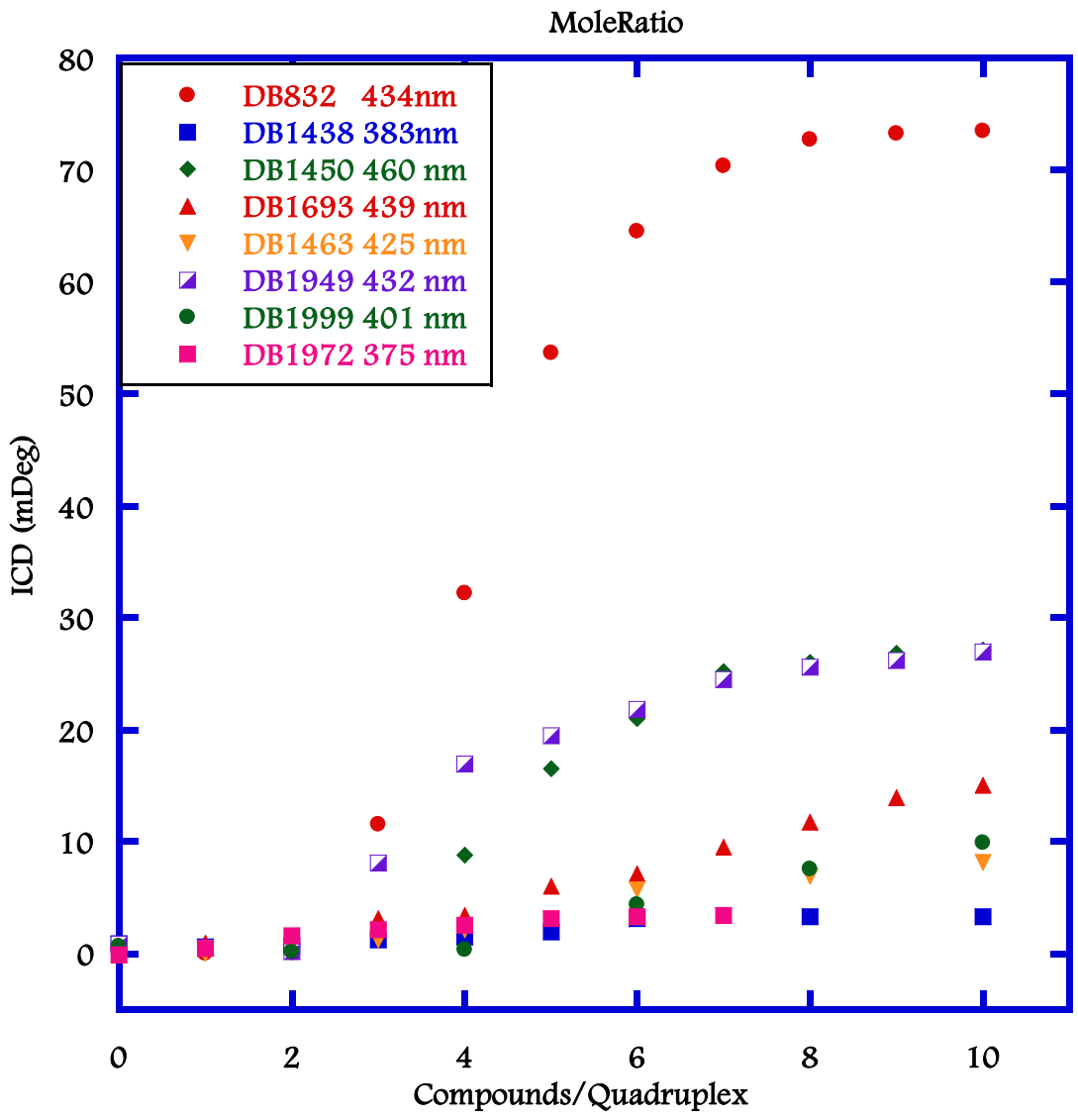


Figure 4.6: Plot of mole ratio versus the ICD signal at wavelength corresponding to the maximum absorbance of the bound ligand with Tel24.

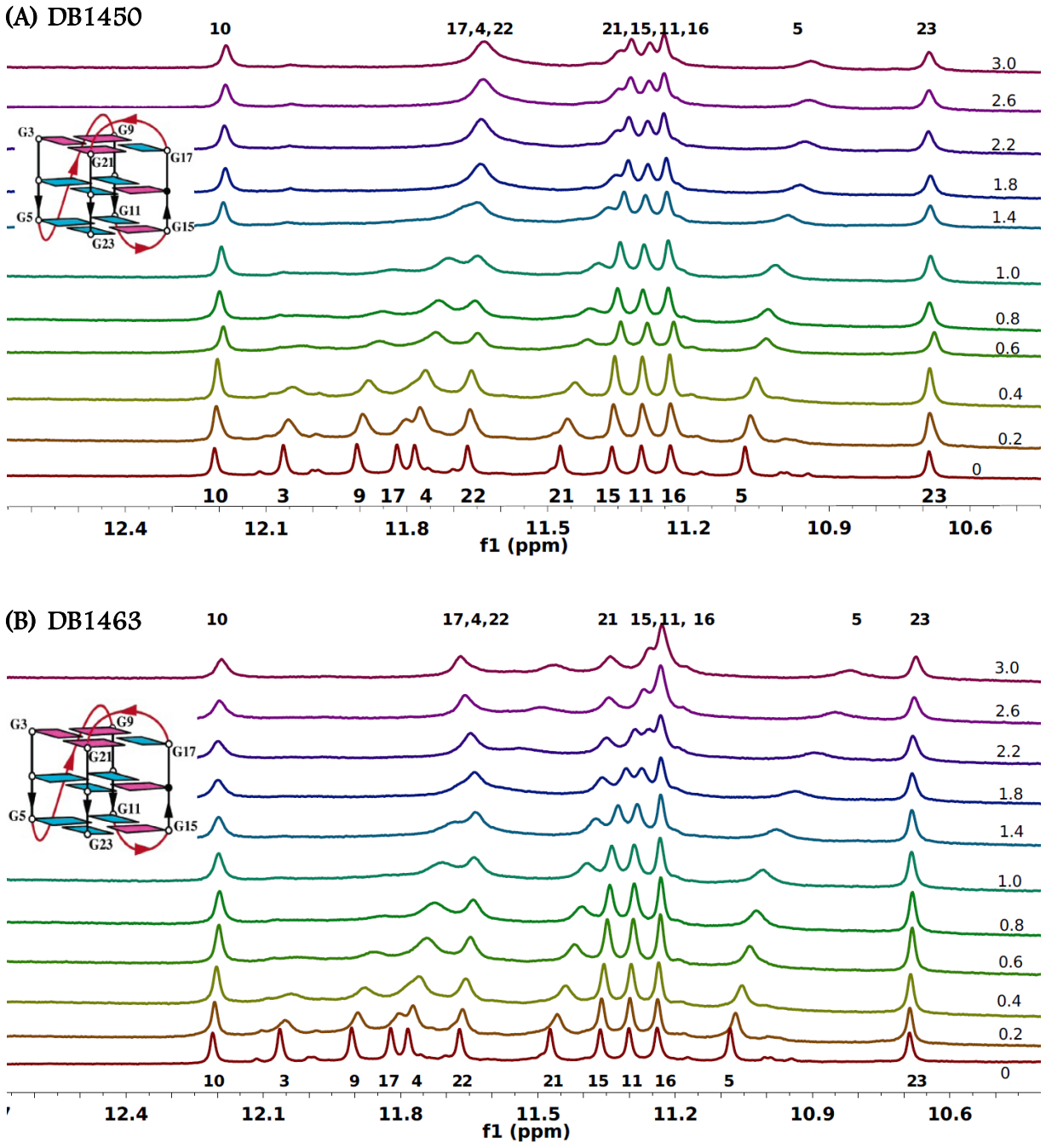


Figure 4.7: Imino proton spectra of Tel24 with (A) DB1450 and (B) DB1463 at 25 °C in 10 mM K_2HPO_4/80 mM KCl, pH 7.0.

Inset shows the hybrid-1fold of Tel24 reported in Luu et al [34]. Proton assignments of Tel24 were provided by Dr. Anh T. Phan

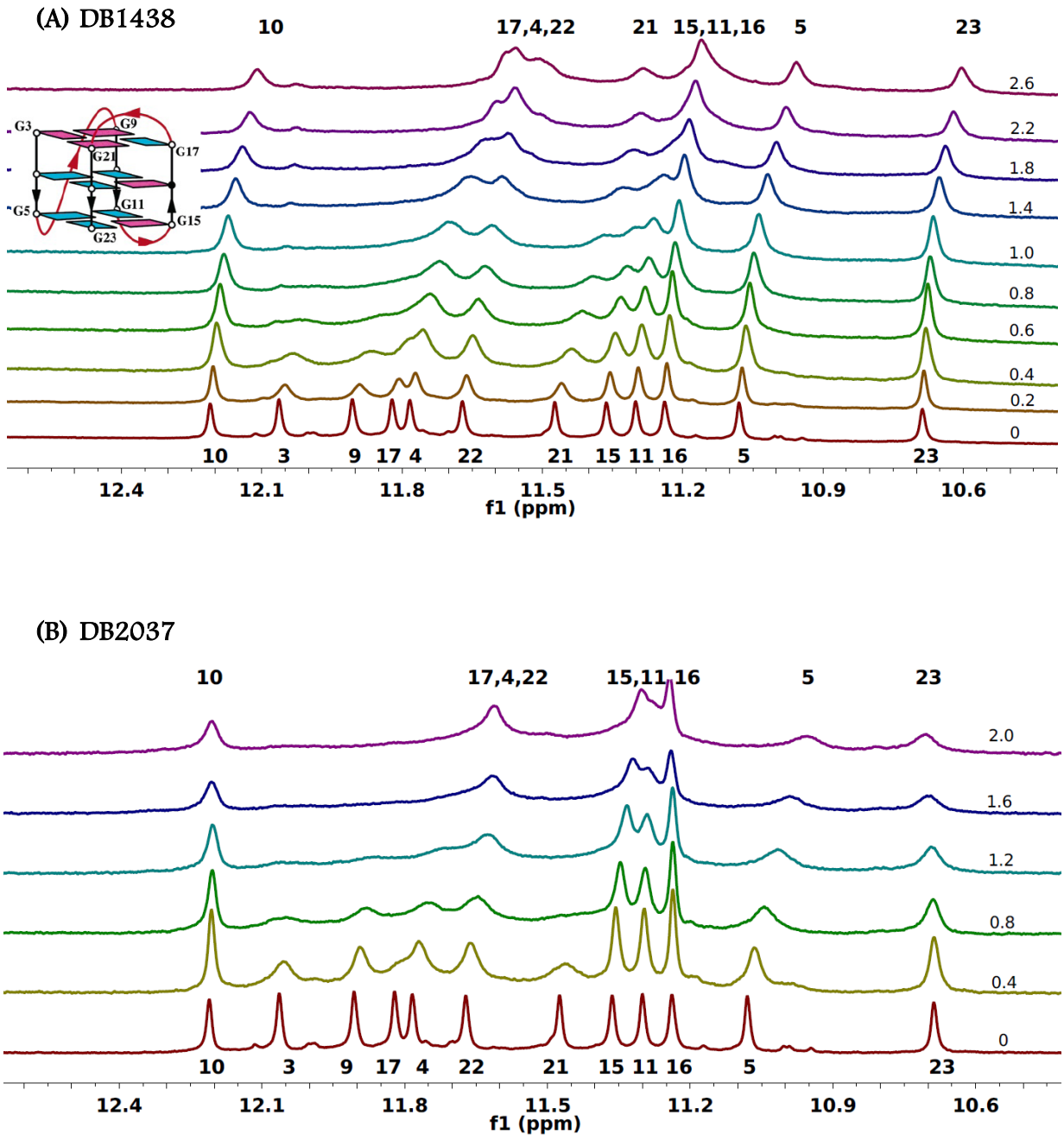
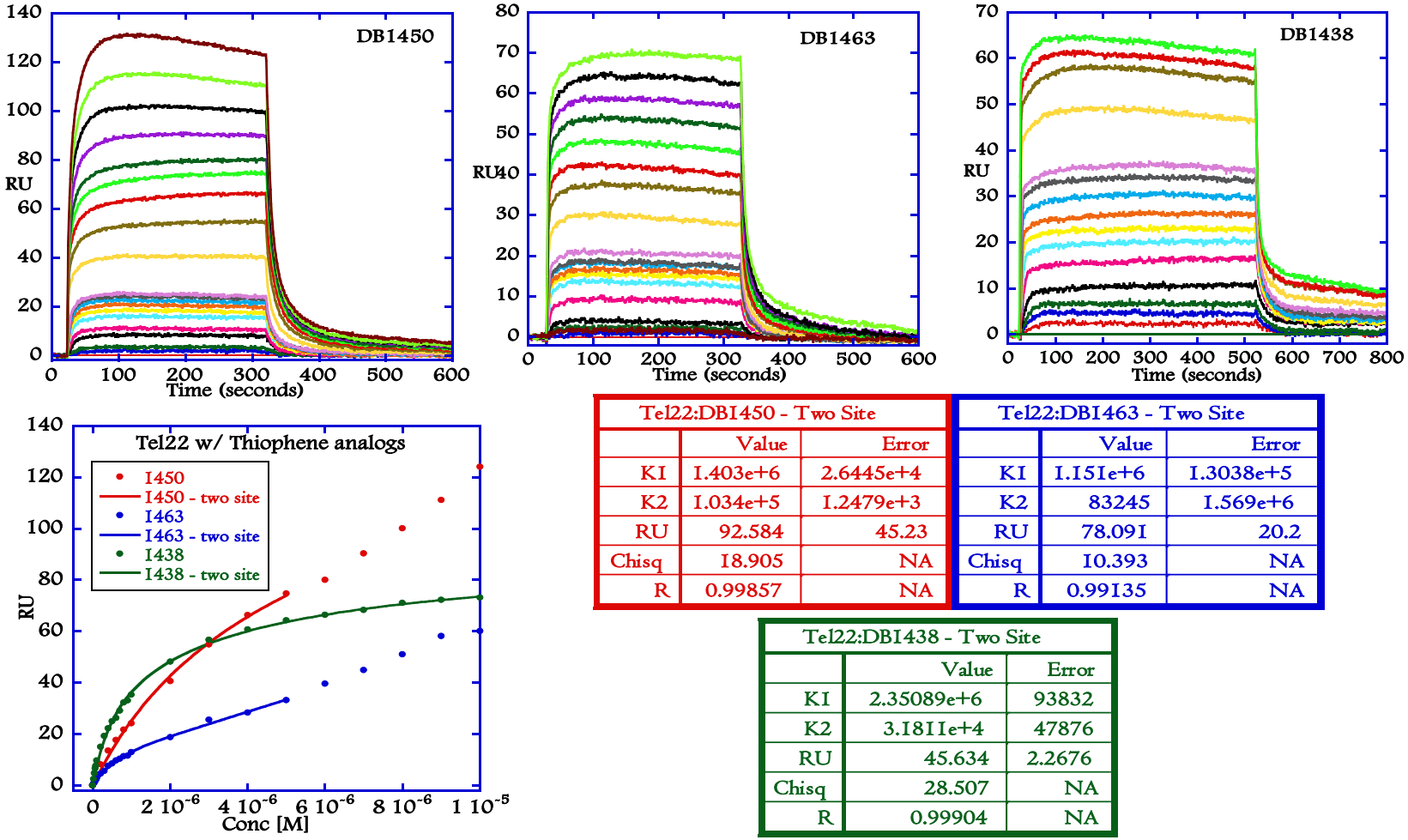


Figure 4.8: Imino proton spectra of Tel24 with (A) DB1438 and (B) DB2037 at 25 °C in 10 mM K_2HPO_4/80 mM KCl, pH 7.0.

Inset shows the hybrid-1 fold of Tel24 reported in Luu et al [34]. Proton assignments of Tel24 were provided by Dr. Anh T. Phan



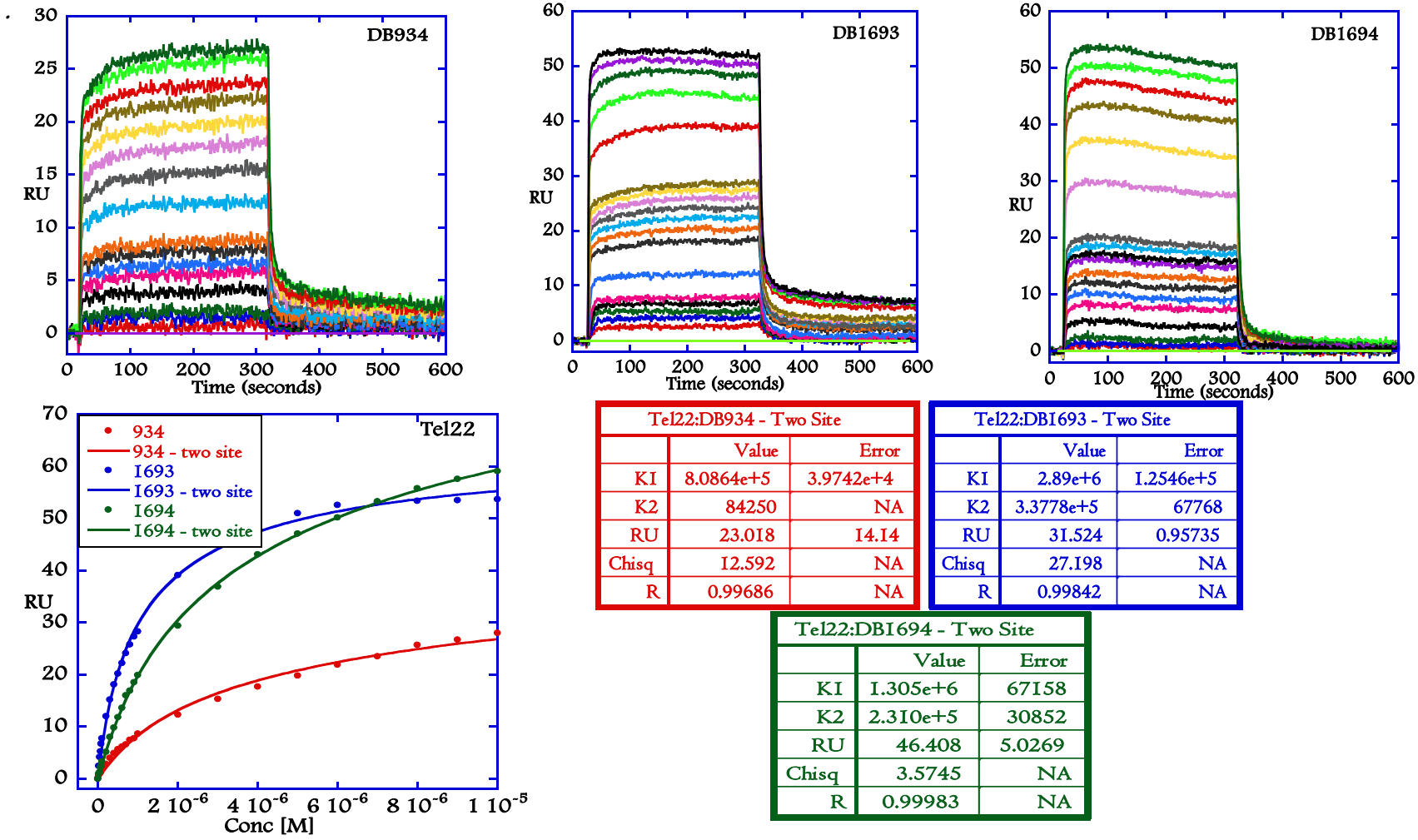
Tel22:DBI450 - Two Site		
	Value	Error
KI	I.403e+6	2.6445e+4
K2	I.034e+5	I.2479e+3
RU	92.584	45.23
Chisq	I8.905	NA
R	0.99857	NA

Tel22:DBI463 - Two Site		
	Value	Error
KI	I.I5Ie+6	I.3038e+5
K2	83245	I.569e+6
RU	78.09I	20.2
Chisq	I0.393	NA
R	0.99I35	NA

Tel22:DBI438 - Two Site		
	Value	Error
KI	2.35089e+6	93832
K2	3.I8I Ie+4	47876
RU	45.634	2.2676
Chisq	28.507	NA
R	0.99904	NA

Figure 4.9: SPR sensorgrams for binding of representative thiophene analogs with Tel22 quadruplex sequence. DB1450 (top left), DB1463 (top middle) and DB1438 (top right) to Tel22 in 10 mM TRIS buffer containing 100 mM K^+ at 25 °C.

The quadruplex curves range in ligand concentration from 10 nM for the bottom curve to 10 μM for the top curve. Steady-state binding plots (bottom left) fit to a two-site model (Materials and Methods). The steady-state curves were fitted only for low compound concentrations, since at higher concentrations, non-specific binding is occurring. The concentration values are for unbound compound concentration in the flow solution. Binding parameters (bottom right) obtained by fitting to a two-site model.



Tel22:DB934 - Two Site		
	Value	Error
K1	8.0864e+5	3.9742e+4
K2	84250	NA
RU	23.018	14.14
Chisq	12.592	NA
R	0.99686	NA

Tel22:DB1693 - Two Site		
	Value	Error
K1	2.89e+6	1.2546e+5
K2	3.3778e+5	67768
RU	31.524	0.95735
Chisq	27.198	NA
R	0.99842	NA

Tel22:DB1694 - Two Site		
	Value	Error
K1	1.305e+6	67158
K2	2.310e+5	30852
RU	46.408	5.0269
Chisq	3.5745	NA
R	0.99983	NA

Figure 4.10: SPR sensorgrams for binding of DB832 analogs with Tel22 quadruplex sequence. DB934 (top left), DB1693 (top middle) and DB1694 (top right) to Tel22 in 10 mM TRIS buffer containing 100 mM K^+ at 25 °C.

The quadruplex curves range in ligand concentration from 10 nM for the bottom curve to 10 μM for the top curve. Steady-state binding plots (bottom left) fit to a two-site model (Materials and Methods). The concentration values are for unbound compound concentration in the flow solution. Binding parameters (bottom right) obtained by fitting to a two-site model.

Table 4.1: Thermal melting results[a] at different ratios (listed in bracket) for Tel22[b] and two duplex sequences (Dickerson and GC20) with 5-5-6 and related ring systems at 100 mM K^+.

Compound	Tel22 (2:1)	Tel22 (4:1)	AATT (2:1)	(GC)20 (2:1)
DB832	19.1	26.1	3.0	3.7
DB1093	0.00	5.00	5.0	2.5
DB1438	18.2	25.5	2.1	2.5
DB1450	17.6	24.2	8.1	9.1
DB1463	17.3	23.1	8.5	3.4
DB934	16.8	22.6	2.0	2.3
DB1693	22.5	27.2	5.1	2.1
DB1694	16.5	23.8	1.0	2.3
DB1949	ND	11.5	2.3	ND
DB1972[c]	ND	13.0	4.0	ND
DB2037[c]	ND	20.0	7.6	ND
DB1999[c]	ND	9.50	2.5	4.8
Distamycin[c]	ND	0.00	7.0	ND
RHPS4[c]	ND	9.30	3.0	ND

[a] ΔTm values (complex – free DNA, in °C) are estimated as the mid-points of the melting transition.

[b] Ligand free single-strand concentration in this study is 2-3 μM.

[c] Performed by Caterina Musetti in our laboratory. The ΔTm values listed are at 50 mM K^+ concentration.

ND: Not determined

Table 4.2: Equilibrium binding constants[a] determined by SPR for representative DB compounds with Tel22.

Compound	K_1	K_2
DB1450	1.40	0.10
DB1463	1.15	0.08
DB1438	2.35	0.03
DB1693	2.89	0.34
DB1694	1.30	0.23
DB934	0.81	0.08
DB832[b]	3.79	0.23

[a] Binding Constants ($K \times 10^6\ M^{-1}$) determined with a two-site model (See Figure 4.9 and 4.10 for fitting)

[b] Binding Constant with Tel24 sequence (Chapter 2, Figure 2.11)

4.5 REFERENCES

[1] a)J. Adams, P. Elliott, *Oncogene* **2000**, *19*, 6687 ; b)P. G. Baraldi, A. Bovero, F. Fruttarolo, D. Preti, M. A. Tabrizi, M. G. Pavani, R. Romagnoli, *Medicinal Research Reviews* **2004**, *24*, 475; c)W. Denny, *Eur J Med Chem* **2001**, *36*, 577 ; d)W. A. Denny, *Current Medicinal Chemistry* **2001**, *8*, 533; e)S. Nelson, L. Ferguson, W. Denny, *Cell & Chromosome* **2004**, *3*, 2; f)J. Woynarowski, *Biochim Biophys Acta* **2002**, *1587*, 300

[2] a)S. Burge, G. N. Parkinson, P. Hazel, A. K. Todd, S. Neidle, *Nucleic Acids Res* **2006**, *34*, 5402; b)T. R. Cech, *Cell* **2004**, *116*, 273; c)J. Dai, M. Carver, D. Yang, *Biochimie* **2008**, *90*, 1172; d)A. T. Phan, V. Kuryavyi, D. J. Patel, *Curr Opin Struct Biol* **2006**, *16*, 288.

[3] D. Monchaud, *Organic & biomolecular chemistry* **2008**, *6*, 627.

[4] a)S. M. Kerwin, *Current Pharmaceutical Design* **2000**, *6*, 441; b)S. M. Kerwin, *Bioorganic & medicinal chemistry letters* **2001**, *11*, 2411; c)J. L. Mergny, C. Helene, *Nat Med* **1998**, *4*, 1366; d)S. Neidle, G. N. Parkinson, *Biochimie* **2008**, *90*, 1184; e)S. Neidle, D. E. Thurston, *Nat Rev Cancer* **2005**, *5*, 285.

[5] a)W. D. Wilson, F. A. Tanious, A. Mathis, D. Tevis, J. E. Hall, D. W. Boykin, *Biochimie* **2008**, *90*, 999; b)B. M. Watkins, *Trends in Parasitology* **2003**, *19*, 477; c)J. V. Walker, J. L. Nitiss, *Cancer Investigation* **2002**, *20*, 570; d)P. Trouiller, P. L. Olliaro, *International Journal of Infectious Diseases* **1999**, *3*, 61; e)M. N. Soeiro, E. M. De Souza, C. E. Stephens, D. W. Boykin, *Expert Opinion on Investigational Drugs* **2005**, *14*, 957; f)R. Palchaudhuri, P. J. Hergenrother, *Current Opinion in Biotechnology* **2007**, *18*, 497.

[6] a)A. R. Urbach, P. B. Dervan, *Proceedings of the National Academy of Sciences of the United States of America* **2001**, *98*, 4343; b)M. d. N. Soeiro, E. M. d. Souza, D. W. Boykin, *Expert Opinion on Therapeutic Patents* **2007**, *17*, 927; c)M. L. Kopka, C. Yoon, D. Goodsell, P. Pjura, R. E. Dickerson, *Proceedings of the National Academy of Sciences of the United States of America* **1985**, *82*, 1376; d)P. B. Dervan, R. W. Bürli, *Current Opinion in Chemical Biology* **1999**, *3*, 688; e)P. Dervan, B. Edelson, *Curr Opin Struct Biol* **2003**, *13*, 284

[7] a)J. M. Woynarowski, R. D. Sigmund, T. A. Beerman, *Biochemistry* **1989**, *28*, 3850; b)U. H. Mortensen, T. Stevnsner, S. Krogh, K. Olesen, O. Westergaard, B. J. Bonven, *Nucl. Acids Res.* **1990**, *18*, 1983; c)T. A. Beerman, J. M. Woynarowski, R. D. Sigmund, L. S. Gawron, K. E. Rao, J. W. Lown, *Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression* **1991**, *1090*, 52; d)D. Bartulewicz, K. Bielawski, A. Bielawska, *Archiv der Pharmazie* **2002**, *335*, 422; e)J. W. George, S. Ghate, S. W. Matson, J. M. Besterman, *Journal of Biological Chemistry* **1992**, *267:*, 10683; f)R. M. Brosh, Jr, J. K. Karow, E. J. White, N. D. Shaw, I. D. Hickson, V. A. Bohr, *Nucl. Acids Res.* **2000**, *28*, 2420.

[8] a)M. Sands, M. A. Kron, R. B. Brown, *Reviews of Infectious Diseases* **1985**, *7*, 625; b)R. D. PEARSON, E. L. HEWLETT, *Annals of Internal Medicine* **1985**, *103*, 782; c)A. J. Nok, *Parasitology Research* **2003**, *90*, 71; d)M. P. Hutchinson, H. J. C. Watson, *Transactions of the Royal Society of Tropical Medicine and Hygiene* **1962**, *56*, 227; e)E. A. ELAMIN, A. M. HOMEIDA, S. E. I. ADAM, M. M. MAHMOUD, *Journal of Veterinary*

Pharmacology and Therapeutics **1982**, *5*, 259; f)F. Doua, T. W. Miezan, J. R. S. Singaro, F. B. Yapo, T. Baltz, *Am J Trop Med Hyg* **1996**, *55*, 586.

[9] G. D. Larsen TA, Cascio D, Grzeskowiak K, Dickerson RE., *J Biomol Struct Dyn.* **1989**, *7*, 14.

[10] E. Trotta, E. D'Ambrosio, G. Ravagnan, M. Paci, *Nucl. Acids Res.* **1995**, *23*, 1333.

[11] a)G. Riou, J. Benard, *Biochemical and Biophysical Research Communications* **1980**, *96*, 350; b)C. Brack, E. Delain, G. Riou, B. Festy, *Journal of Ultrastructure Research* **1972**, *39*, 568.

[12] D. G. Brown, M. R. Sanderson, E. Garman, S. Neidle, *Journal of Molecular Biology* **1992**, *226*, 481.

[13] a)S. F. Queener, *journal of medicinal chemistry* **1995**, *38*, 4739; b)V. Pyrgos, S. Shoham, E. Roilides, T. J. Walsh, *Paediatric Respiratory Reviews* **2009**, *10*, 192; c)J. E. Conte, D. Chernoff, D. W. Feigal, P. Joseph, C. McDonald, J. A. Golden, *Annals of Internal Medicine* **1990**, *113*, 203.

[14] a)C. M. Nunn, S. Neidle, *journal of medicinal chemistry* **1995**, *38*, 2317; b)K. J. Edwards, T. C. Jenkins, S. Neidle, *Biochemistry* **1992**, *31*, 7104; c)K. R. Fox, C. E. Sansom, M. F. G. Stevens, *FEBS Letters* **1990**, *266*, 150.

[15] T. Sun, Y. Zhang, *Nucl. Acids Res.* **2008**, *36*, 1654.

[16] A. Feddersen, K. Sack, *J. Antimicrob. Chemother.* **1991**, *28*, 437.

[17] a)D. W. Boykin, A. Kumar, J. Spychala, M. Zhou, R. J. Lombardy, W. D. Wilson, C. C. Dykstra, S. K. Jones, J. E. Hall, *journal of medicinal chemistry* **1995**, *38*, 912; b)S. M. Rahmathullah, J. E. Hall, B. C. Bender, D. R. McCurdy, R. R. Tidwell, D. W. Boykin, *journal of medicinal chemistry* **1999**, *42*, 3994.

[18] a)A. M. W. Stead, P. G. Bray, I. G. Edwards, H. P. DeKoning, B. C. Elford, P. A. Stocks, S. A. Ward, *Molecular Pharmacology* **2001**, *59*, 1298; b)A. Purfield, R. Tidwell, S. Meshnick, *Malaria Journal* **2009**, *8*, 104; c)I. Francesconi, W. D. Wilson, F. A. Tanious, J. E. Hall, B. C. Bender, R. R. Tidwell, D. McCurdy, D. W. Boykin, *journal of medicinal chemistry* **1999**, *42*, 2260; d)E. A. Steck, K. E. Kinnamon, D. E. Davidson Jr, R. E. Duxbury, A. J. Johnson, R. E. Masters, *Experimental Parasitology* **1982**, *53*, 133.

[19] a)L. Wang, C. Bailly, A. Kumar, D. Ding, M. Bajic, D. W. Boykin, W. D. Wilson, *Proceedings of the National Academy of Sciences of the United States of America* **2000**, *97*, 12; b)L. Wang, C. Carrasco, A. Kumar, C. E. Stephens, C. Bailly, D. W. Boykin, W. D. Wilson, *Biochemistry* **2001**, *40*, 2511.

[20] C. Bailly, C. Tardy, L. Wang, B. Armitage, K. Hopkins, A. Kumar, G. B. Schuster, D. W. Boykin, W. D. Wilson, *Biochemistry* **2001**, *40*, 9770.

[21] M. J. Cocco, L. A. Hanakahi, M. D. Huber, N. Maizels, *Nucl. Acids Res.* **2003**, *31*, 2944.

[22] L. Martino, A. Virno, B. Pagano, A. Virgilio, S. Di Micco, A. Galeone, C. Giancola, G. Bifulco, L. Mayol, A. Randazzo, *Journal of the American Chemical Society* **2007**, *129*, 16048.

[23] E. W. White, F. Tanious, M. A. Ismail, A. P. Reszka, S. Neidle, D. W. Boykin, W. D. Wilson, *Biophysical Chemistry* **2007**, *126*, 140.

[24] a)S. Mallena, M. P. H. Lee, C. Bailly, S. Neidle, A. Kumar, D. W. Boykin, W. D. Wilson, *Journal of the American Chemical Society* **2004**, *126*, 13659; b)P. Bilik, F. Tanious, A. Kumar, W. D. Wilson, D. W. Boykin, P. Colson, C. Houssier, M. Facompré, C. Tardy, C. Bailly, *ChemBioChem* **2001**, *2*, 559.

[25] E. Gavathiotis, R. A. Heald, M. F. G. Stevens, M. S. Searle, *Journal of Molecular Biology* **2003**, *334*, 25.

[26] a)F. A. Tanious, B. Nguyen, W. D. Wilson, in *Methods in Cell Biology, Vol. Volume 84* (Eds.: J. C. Dr. John, Dr. H. William Detrich, III), Academic Press, **2008**, pp. 53; b)B. Nguyen, F. A. Tanious, W. D. Wilson, *Methods* **2007**, *42*, 150.

[27] T. M. Davis, W. D. Wilson, *Analytical Biochemistry* **2000**, *284*, 348.

[28] a)G. S. H. Lee, M. A. Wilson, B. R. Young, *Organic Geochemistry* **1998**, *28*, 549; b)M. Liu, X.-a. Mao, C. Ye, H. Huang, J. K. Nicholson, J. C. Lindon, *Journal of Magnetic Resonance* **1998**, *132*, 125.

[29] J. Dash, P. S. Shirude, S.-T. D. Hsu, S. Balasubramanian, *Journal of the American Chemical Society* **2008**, *130*, 15950.

[30] R. M. Pagni, *Journal of Chemical Education* **1998**, *75*, 1095.

[31] a)B. Nordén, *Applied Spectroscopy Reviews* **1978**, *14*, 157 ; b)B. Norden, M. Kubista, T. Kurucsev, *Quarterly Reviews of Biophysics* **1992**, *25*, 51.

[32] a)W. D. Wilson, B. Nguyen, F. A. Tanious, A. Mathis, J. E. Hall, C. E. Stephens, D. W. Boykin, *Current Medicinal Chemistry - Anti-Cancer Agents* **2005**, *5*, 389; b)T. Nishimura, T. Okobira, A. M. Kelly, N. Shimada, Y. Takeda, K. Sakurai, *Biochemistry* **2007**, *46*, 8156; c)B. Nguyen, C. Tardy, C. Bailly, P. Colson, C. Houssier, A. Kumar, D. W. Boykin, W. D. Wilson, *Biopolymers* **2002**, *63*, 281; d)M. Munde, M. A. Ismail, R. Arafa, P. Peixoto, C. J. Collar, Y. Liu, L. Hu, M.-H. David-Cordonnier, A. Lansiaux, C. Bailly, D. W. Boykin, W. D. Wilson, *Journal of the American Chemical Society* **2007**, *129*, 13732; e)Y. Miao, M. P. H. Lee, G. N. Parkinson, A. Batista-Parra, M. A. Ismail, S. Neidle, D. W. Boykin, W. D. Wilson, *Biochemistry* **2005**, *44*, 14701; f)T. Banerjee, R. Mukhopadhyay, *Biochemical and Biophysical Research Communications* **2008**, *374*, 264.

[33] Q. Chen, I. D. Kuntz, R. H. Shafer, *Proceedings of the National Academy of Sciences of the United States of America* **1996**, *93*, 2635.

[34] K. N. Luu, A. T. Phan, V. Kuryavyi, L. Lacroix, D. J. Patel, *J Am Chem Soc* **2006**, *128*, 9963.

[35] L. D. Han H, Rangan A, Hurley LH., *Journal of the American Chemical Society* **2001**, *123*, 11.

[36] a)J. Li, J. J. Correia, L. Wang, J. O. Trent, J. B. Chaires, *Nucleic Acids Res* **2005**, *33*, 4649; b)Y. Xu, Y. Noguchi, H. Sugiyama, *Bioorg Med Chem* **2006**, *14*, 5584; c)G. N. Parkinson, M. P. Lee, S. Neidle, *Nature* **2002**, *417*, 876.

[37] A. Ambrus, D. Chen, J. Dai, T. Bialis, R. A. Jones, D. Yang, *Nucleic Acids Res* **2006**, *34*, 2723.

[38] A. T. Phan, V. Kuryavyi, H. Y. Gaw, D. J. Patel, *Nat Chem Biol* **2005**, *1*, 167.

[39] N. Maizels, *Nat Struct Mol Biol* **2006**, *13*, 1055.

[40] a)X. Yang, H. Robinson, Y. Gao, A. Wang, *Biochemistry* **2000**, *39*, 10950 ; b)R. L. Rich, D. G. Myszka, *Journal of Molecular Recognition* **2010**, *23*, 1; c)R. L. Rich, D. G. Myszka, *Journal of Molecular Recognition* **2008**, *21*, 355.

5 EVALUATION OF AZACYANINES, NAPHTHALENE DIIMIDES AND DIARYL UREAS AS NEW SCAFFOLDS FOR G-QUADRUPLEX RECOGNITION: SURFACE PLASMON RESONANCE STUDIES

5.1 INTRODUCTION

Stabilization of quadruplex architecture by small molecules is emerging as a potential anticancer approach since it is thought to interfere with oncogenic expression and telomeric maintenance in cancer cells. Interestingly, several classes of small molecules have been developed that efficiently target G-quadruplex DNA. Most of the quadruplex-interactive small molecules discovered to this time are planar, aromatic, heterocyclic scaffolds that are generally derived from duplex DNA intercalators. The planar units of these molecules are conformationally constrained and, as a result, primarily interact with quadruplex units by stacking on either one or both ends of the terminal G-tetrads. Because of the structural similarity to duplex DNA intercalators, many of these quadruplex-binding molecules exhibit limited selectivity for quadruplex over duplex structures. Binding to non-targeted duplex sequences can result in significant loss of valuable compound with potential non-selective cytotoxicity. Since potential quadruplex-forming sequences are common throughout the genome, selectivity for a particular quadruplex structure over other quadruplex motifs is also a concern in the design of quadruplex-binding compounds. Binding to non-targeted quadruplex sequences might also have deleterious effect on regulation of non-targeted genes. Increasing selectivity of small molecules for their targets is therefore an important focus of research.

Several classes of small molecules that interact with selective quadruplex forming motifs via end-stacking have been extensively investigated [1]. Porphyrins and acridines are two important classes of small molecules that have been thoroughly investigated and have shown

potential to lead to clinical agents [1b, 2]. Porphyrins are a class of heterocyclic macrocycles composed of pyrrole subunits linked through methine bridges. Porphyrins are known to bind to and stabilize different types of G-quadruplexes and, in some cases, to facilitate G-quadruplex formation [2b, 3]. TMPyP4 (Figure 5.1) is a commercially-available cationic porphyrin that has been widely studied and shown to bind to quadruplex DNA sequences [4]. TMPyP4 binds to human telomeric DNA by stacking externally on each of the terminal quartets, with the positively charged groups oriented toward the sugar-phosphate DNA backbone [2b]. However, several other secondary binding modes have also been reported for TMPyP4 [5], but the external end-stacking is established as the most common interaction mode. The stabilization of quadruplex conformations of telomeric DNA by TMPyP4 has shown to prevent recognition of the telomere by telomerase and resulting in effective inhibition of the enzyme [3b, 6]. TMPyP4 also binds to and stabilizes the G-rich strand of the intramolecular *c-myc* sequence, with each molecule stacked externally on the G-tetrad surface [3d]. Moreover, binding of TMPyP4 to *c-myc* also induces a switch from a parallel conformation of *c-myc* to an antiparallel conformation; this conformational switch has been shown to directly repress the transcriptional activity and effectively inhibit tumor growth *in vivo* [2a, 7]. Therefore, TMPyP4 has a potential to be developed as an effective chemotherapeutic agent, but its associated toxicity is also a major concern [8].

In spite of being a great quadruplex-stabilizer and an effective telomerase inhibitor, TMPyP4 has several drawbacks that would prevent it from being used therapeutically [9]. TMPyP4 has an enhanced selectivity for duplex sequences and can efficiently stack in the grooves of AT-rich sequences [10]. An exquisite specificity is imperative to target the relatively few quadruplex structures in the presence of an excess of double-stranded DNA in the genome to attain the maximum effectiveness. Also, the broad spectra of porphyrin molecules are known to produce photo-induced cytotoxicity due to their unique chemical structure, resulting

in indiscriminant cell damage [11]. TMPyP4 also binds preferentially to intermolecular quadruplexes over the biologically relevant intramolecular quadruplexes. This intermolecular binding has been demonstrated through the TMPyP4-induced formation of anaphase bridges in sea urchins [3b]. The porphyrin core of TMPyP4 is smaller (~10 Å) than the size of the G-quartet (~13 Å) which results in the molecule binding with an offset from the center of the terminal G-quartet [2b, 4]. TMPyP4 only overlaps with two out of four of the guanines in the terminal G-quartets. The binding of TMPyP4 to the quadruplex could potentially be improved if it were modified to increase the contact surface between the compound and the G-quartets, in order to maximize stacking interactions.

The acridine scaffold is of great interest since its DNA affinity and intercalative properties make it an important pharmacophore for the design of antitumor drugs targeting DNA [1a, 1b, 2d, 12]. The global substitution on the heterocycle is crucial for a specific biological activity and also selectivity for tumor cells. To increase the strength of DNA binding in an effort to design more potent molecules, different strategies have been used either by adding cationic substituents or by preparing dimeric molecules in which two intercalating structures are linked together [1b, 13]. Tri-substituted acridine, BRACO19 (Figure 5.1), and a pentacyclic acridine, RHPS4 (Figure 5.1), selectively bind to the single-strand overhang of telomeric DNA, induce the formation of G-quadruplex structures and inhibit cell growth and telomerase activity at sub-cytotoxic doses [14]. Both RHPS4 and BRACO19 cause chromosomal end-to-end fusions and anaphase bridges, which are mitotic defects that typically result in nearly immediate cellular senescence or apoptosis [15]. Furthermore, it has been discovered that RHPS4 also has the ability to act synergistically with several common cancer drugs such as Taxol, Doxorubicin, and others [16]. However, some antagonism has also occurred with drugs such as Cisplatin that have targets similar to RHPS4 [17]. *In vivo* studies with BRACO19 also reveal it has positive combinatorial affects with Taxol in promoting cancer-cell senescence [18].

However, pharmacokinetic studies have shown that BRACO-19 and its associated compounds pose stability problems during the preparation of dosage forms, their storage and after application [19]. This limits its therapeutic applications, since any new drug candidate should be submitted to thorough stability investigations prior to *in vitro* and *in vivo* tests. Cell permeability studies with Caco-2 cells, a standard model for intestinal drug absorption, indicate that BRACO-19 might not be suitable for oral administration [20]; this was also supported by tumor xenografts in mice that did not respond to an oral treatment with BRACO19 [21]. The drawbacks associated with BRACO19 are partially attributed to its chemical structure. It contains two basic pyrrolidine rings that are very likely to be protonated under physiological conditions; that is, the molecule is positively charged. This results in good water solubility but strongly decreases the interaction with hydrophobic structures like cell membranes and significantly reduces the cellular uptake potential of BRACO19 [20].

Telomestatin (Figure 5.1), a natural product macrocycle with a larger ring system than TMPyP4, has been shown to bind strongly to and stabilize intramolecular basket-type G-quadruplex structures [3c, 22]. Studies have shown that Telomestatin binds to the ends of the basket conformation of the intramolecular human telomeric sequence [2c]. Telomestatin is also able to induce and stabilize G-quadruplexes in the absence of added monovalent cations, which is a unique characteristic among small molecules [22]. Telomestatin is the strongest telomerase inhibitor ever reported of any G-quadruplex interactive small molecule, with an IC_{50} of ca. 0.005 μM [3c] as compared to TmPyP4, with an IC_{50} of ca 0.6 μM [3b]. Telomestatin is also highly selective for quadruplex DNA over duplex DNA [2c], likely due to the close match between its size, shape and macrocycle structure to that of the G-quartet. Nevertheless, being a natural product, Telomestatin is very difficult to obtain, especially in quantities needed for widespread therapeutic use. The complex ring system is difficult to synthesize, and at present would be cost-prohibitive to produce commercially. Also, since Telomestatin is an uncharged

molecule, it has low solubility in water, which is not only undesirable from a therapeutic standpoint, but also makes it difficult to study [23].

Approaches to platinum-based anti-cancer targeting have included conjugation of known minor-groove DNA binding ligands and intercalators into discrete platinum complexes as well as noncovalent dinuclear and trinuclear complexes [24]. Platinum (II) complexes have also been shown to interact with the human telomeric G-quadruplex [25]. Terpyridine-platinum (Pt-TPY) complexes selectively trap the adenine residues in the loops of the antiparallel quadruplex conformation of the wild-type human telomeric DNA, and their interaction being driven by the aromatic surface of the ligand [26]. Platinum-quinacridine hybrid (Pt-MPQ) (Figure 5.1) interacts with G-quadruplex DNA via a dual noncovalent/covalent binding mode, targeting preferentially the guanines constitutive of external G-tetrads [26-27]. However, the lack of specificity for a particular DNA sequence, an inherent drawback observed in platinum chemistry, makes it very difficult to design platinum-based quadruplex-selective drugs [28].

The unique physicochemical properties of these ligands and its relation to the observed cytotoxicity render important challenges in modifying these small molecules in an effective way to circumvent any deleterious effects. Alternatively, new inroads can be made towards the design and development of new classes of small molecules that can interact with target sequences with enhanced selectivity and exhibit very little side effects. The compounds mentioned above have the potential to be less harmful than current chemotherapies which often affect all rapidly proliferating cells rather than just tumor cells. Even if none of these small molecules prove to be directly applicable to clinical medicine, the insight these compounds have provided into the function of telomerase and telomere structure will be indispensable to the future of cancer therapeutics.

The success of the agents described above to selectively target G-quadruplexes has led to the design of different classes of potential quadruplex-interacting small molecules. The interaction of several classes has been evaluated with several quadruplex forming motifs (Figure 5.2) from the genome using surface plasmon resonance (SPR)-biosensor studies. The selectivity of these different classes of ligands has also been evaluated with various duplex forming sequences.

On the basis of modeling and comparative analysis of the binding of Coralyne, a crescent-shaped molecule known to intercalate duplex and triplex DNA [29], and slightly larger than typical DNA intercalators, a series of azacyanines (Figure 5.3) were designed and synthesized by Dr. Nick Hud's group at Georgia Institute of Technology. The fused ring systems of azacyanines were hypothesized to be marginally too large to intercalate a Watson-Crick duplex but effectively stack at the terminal G-quartets. The synthesis of azacyanines has been reported [30], and the relatively ease of synthesis of these molecules makes them ideal candidates to study as quadruplex-interactive agents.

Naphthalene imides and diimides (ND) are a different class of compounds that are shown to bind to duplex DNA [31]. Several show *in vivo* anti-cancer activity and two (Amonafide and Elinafide) have been evaluated in anti-cancer clinical trials in humans [32]. A series of disubstituted NDs have been previously reported as quadruplex-interactive ligands but showed only low affinity [1g]. Nevertheless, the planarity of the ND moiety would be an interesting starting point for the development of more complex ligands and possibly possess higher affinities for quadruplex motifs. The introduction of different side chains along the ND core would produce a greater ligand diversity that may discriminate between different types of G-quadruplexes and also with much higher selectivity over non-targeted sequences. Using this design principle, Dr. Stephen Neidle's group at University of London developed a series of tri

and tetra-substituted analogs of NDs (Figure 5.4), in hopes of improving binding and selectivity. The recent advancements in efficient synthesis routes for this class of ligands also makes them ideal candidates to study as quadruplex-interactive agents [33].

Substituted urea analogs with the diaryl functionality are another class of compounds that is present in numerous clinically-approved therapeutic agents [34]. The diaryl urea skeleton is represented in many kinase inhibitors, for example, multi-targeted tyrosine kinase inhibitors and inhibitors of insulin-like growth factor I receptor signaling [35]. The diaryl scaffold represents planar, non-polycyclic systems that deviate from the normal quadruplex-interacting ligands, which are exemplified by planar but polycyclic and fused-ring systems. Structure-based design indicated that incorporation of additional phenyl-carbamoyl groups to the 1, 3-diphenyl scaffold allows all four guanine bases of a G-tetrad to be targeted, with addition of cationic side chains to enhance G-quadruplex potency, selectivity and aid solubility of these ligands [36]. Using the simple and efficient "click-chemistry" approach [37], Dr. Stephen Neidle's group designed a series of 1, 3-diphenyl derivatives with different ring systems and varying side-chain lengths. These compounds were also evaluated for their quadruplex recognition potential by using the very powerful biosensor-SPR methods.

5.2 MATERIALS AND METHODS

5.2.1 Sample Preparation

The 5′-biotin labeled oligonucleotide sequences Tel22, d[AGGG(TTAGGG)$_3$]; Tel24, d[TTGGG(TTAGGG)$_3$A]; Tel26, d[AAAGGG(TTAGGG)$_3$AA]; *c-kit1*, d[(AGGG)$_2$CGCTGGGAGGAGGG]; *bcl-2*, d[GGGCGCGGGAGGAATTGGGCGGG]; *c-kit2*, d[(CGGG)$_2$CGCGAGGGAGGGG]; *c-myc*, d[(AGGGTGGGG)$_2$A]; Dickerson, d[CGAATTCGTTTTCGAATTCG]; GC20, [CGCGCGCGTTTTCGCGCGCG]; (AT)$_7$, d[CC(AT)$_7$AGCCCCCGC(TA)$_7$TGG] were purchased with HPLC purification and mass

spectrometry characterization from Integrated DNA Technologies (Coralville, IA). Azacyanines were synthesized by Dr. Nick Hud's group at Georgia Institute of Technology and are reported in [30]. Naphthalene Diimides (NDs) and Diaryl Ureas were synthesized by Dr. Stephen Neidle's group at University of London and are reported [38]. Stock solutions containing 1 mM of each compound were prepared in double distilled water and diluted to working concentrations immediately before use with buffer.

5.2.2 Immobilization of DNA and Biosensor SPR Experiments

Biosensor SPR experiments were performed with a four-channel BIAcore 2000 optical biosensor system (BIAcore, Inc.) and streptavidin-coated sensor chips. All DNA samples, for either duplex- or quadruplex-binding experiments, were used as single strands to prevent dissociation in the SPR flow system. The chips were prepared for use by conditioning with a series of 1 min injections of 1 M NaCl in 50 mM NaOH followed by extensive washing with buffer. 5'-Biotinylated DNA samples (25-50 nM) in HBS buffer were immobilized on the flow cell surface by non-covalent capture as previously described [39]. Three flow cells were used to immobilize DNA samples, and any one of the flow cell was left blank as a control. Interaction analysis was performed by using steady-state methods with multiple injections of increasing compound concentrations over the immobilized DNA surface at 25 °C. Biosensor experiments were conducted in filtered, degassed HEPES buffer (10 mM HEPES, 100 mM KCl, 3 mM EDTA, 0.005 v/v of 10% P20 BIACORE surfactant, pH 7.3) at 25 °C. Flow cell 1 was left blank as a reference, while flow cells 2-4 were immobilized with DNA on a streptavidin-derivatized gold chip (SA chip from BIAcore) by manual injection of DNA stock solutions (flow rate of 1 μL/min) until the desired value of DNA response was obtained (350-400 RU). Compound solutions were prepared in with the running buffer by serial dilutions from stock solution. Typically, a series of different ligand concentrations (1 nM to 10 μM from 20 mM H_2O stock) were injected onto the chip (flow rate of 50 μL/min, 5-10 min) until a constant steady-state

response was obtained followed by a dissociation period (buffer, 10 min). After every cycle, the chip surface was regenerated (20 s injection of 10 mM glycine solution, pH 2.0) followed by multiple buffer injections.

The instrument response (RU) in the steady-state region is proportional to the amount of bound drug and was typically determined by linear averaging over a 10-20 s or longer time span, depending on the length of the steady-state plateau. The predicted maximum response per bound compound in the steady-state region (RUmax) was determined from the DNA molecular weight, the amount of DNA on the flow cell, the compound molecular weight, and the refractive index gradient ratio of the compound and DNA, as previously described [40]. In most of the cases, the observed RU values at high concentrations were greater than RUmax, pointing to more than one binding site in these DNA sequences. The number of binding sites was estimated by fitting plots of RU versus C_{free}. These methods can also be used to determine an empirical RUmax value. The RUmax value is required to convert the observed response (RU) to the standard binding parameter *r* (moles of drug bound per moles of DNA hairpin)

$$r = RU/RUmax$$

which is useful for comparison of a compound binding to different DNAs. To obtain the binding constants, the data were evaluated with different interaction models to obtain an optimal fit using BIAevaluation (BIAcore Inc.) and Kaleidagraph (Synergy Software) software for nonlinear least-squares optimization of the binding parameters:

One site: $r = (K_1C_{free})/(1 + K_1C_{free})$

Two site: $r = (K_1C_{free} + 2K_1K_2C_{free}^2)/(1 + K_1C_{free} + K_1K_2C_{free}^2)$

Three site: $r = (K_1C_{free} + 2K_1K_2C_{free}^2 + 3K_1K_2K_3C_{free}^3)/(1 + K_1C_{free} + K_1K_2C_{free}^2 + K_1K_2K_3C_{free}^3)$

where K_1, K_2 and K_3 are equilibrium constants for three types of binding sites and C_{free} is the concentration of the compound in equilibrium with the complex and is fixed by the concentration in the flow solution.

5.3 RESULTS AND DISCUSSION

5.3.1 Evaluation of Azacyanines

SPR was used to quantitatively evaluate the interaction between azacyanines (Figure 5.3) and a series of quadruplex-forming human telomeric DNA sequences (Tel24 and Tel26, Materials and Methods) in order to gain insight into the selectivity of this class of ligands. Tel24 (Figure 5.2A) and Tel26 (Figure 5.2A) have been shown to predominantly exist as a hybrid-1 quadruplex conformation based on NMR structural characterization by independent laboratories [41]. The three –TTA– loops between the guanine stretches of the hybrid-1 scaffold are primarily oriented in 5'-diagonal/lateral/lateral-3' manner, and the bases involved in the tetrad formation retain the same glycosidic conformation and loop orientation in both the sequences. However, the capping bases at the 5' and the 3'-ends of Tel24 and Tel26 are very different. Therefore compounds that can interact at the ends of either Tel24 or Tel26 conformation can have different binding affinities due to the structural variability at the ends of these two sequences. These two oligonucleotides were immobilized in different flow cells on the same sensor chip, and a range of compound concentrations were injected to monitor the interactions with DNA. Suitable blank control injections with running buffer (HEPES buffer containing 80 mM KCl at pH 7.4) were also performed, and the resulting sensorgrams were subtracted from the compound sensorgrams to obtain the final concentration-dependent graphs for azacyanines. The binding ratio arises from the RU as saturation of binding sites is approached in SPR experiments. During SPR titration, the increase of RU values is directly proportional to the amount of drug bound to DNA molecules immobilized on the sensor chip.

The plot of RU versus the unbound concentration of each of the compounds was fitted to a two-site binding model (Materials and Methods) and the binding constants were determined by averaging the SPR response in the steady-state region. The compounds all showed rapid association and dissociation kinetics with a steady-state plateau that is expected from their planar structure with small substituents. The SPR response versus free compound concentration curves were fit well for both DNA structures with one tight binding site and one or two approximately degenerate, much weaker, secondary sites for both sequences. The sensorgrams and binding plots for the concentration-dependent binding (two-site model, Materials and Methods) of the azacyanines with Tel24, Tel26 and duplex DNA are shown in Figures 5.8-5.10. All ligands tested exhibited a similar, strong primary association constant (Table 5.1) with both DNAs, suggesting that G-quadruplex binding is general to this class of compounds. However, **Aza3** (Tel24, 1.34 x 10^6 M^{-1}; Tel26, 1.74 x 10^6 M^{-1}) and **Aza4** (Tel24, 1.00 x 10^6 M^{-1}; Tel26, 1.22 x 10^6 M^{-1}) showed slightly higher affinities for Tel26 conformer than Tel24 (Table 5.1); whereas, **Aza5** (Tel24, 2.68 x 10^6 M^{-1}; Tel26, 2.11 x 10^6 M^{-1}) exhibited a slightly higher affinity for Tel24 conformer than Tel26. The thermal melting studies (data not shown) of Tel24 and Tel26 have shown that Tel24 is a much more stable conformer than Tel26. The hybrid-fold of Tel26 as determined by NMR reveals the presence of a novel A3·A9·A21 capping structure and covers the top G-quartet [41a]. NMR structure of Tel24 conformation shows the formation of an A·T basepair by T1 and A20 and a reverse A·T basepair by T13 and A24 and the 5' and the 3' ends of this conformation, respectively [41b]. This difference in the end structures between Tel24 and Tel26 might have an indirect effect on the ligand stacking and the observed differences in the binding affinities. **Aza5** also exhibited 2 times stronger binding to both the sequences than **Aza3** or **Aza4** (Table 5.1). This can be due to the presence of additional methoxy groups on both ends of the **Aza5,** potentially contributing towards more favorable hydrophobic interactions with the terminal G-quartet. Interestingly,

Aza4 showed slightly weaker affinities for both the sequences compared to **Aza3**. This can be due to a change in the stacking orientation of the molecule on the terminal G-quartet due to the replacement of nitrogen by sulfur in **Aza4.** However, the small difference in the affinities of **Aza3** and **Aza4** is not significant enough to make any valid arguments.

NMR studies of **Aza3** with Tel24 sequence (conducted by Aaron Engelhart in Dr. Nick Hud's lab at Georgia Institute of Technology) revealed that the ligand was in intermediate exchange on the NMR timescale [30]. A number of aromatic resonances broaden in a site-specific fashion, consistent with the local chemical environment of various residues being perturbed differentially by ligand binding. Interestingly, broadening occurs for both the aromatic residues of the external G-tetrads (G3, G9, G17 and G21 of 5'-tetrad and G4, G15 and G23 of 3'-tetrad) and both A·T base pairs, providing support for a mixed intercalation-exterior stacking mode of binding (i.e., between the exterior tetrads and the capping A·T base pairs of the loops). The primary strong binding site obtained from SPR studies is in excellent agreement with the NMR results and the observation of a weak secondary site also agrees with the moderate NOE effects at the 5' end of the Tel24 conformation [30]. A preliminary model constructed based on the SPR and NMR data of **Aza3** with Tel24 shows that compound preferentially stacking on the terminal tetrad at the 3'-end (Figure 5.6B).

SPR experiments were also performed with intramolecular duplex sequences (Dickerson, GC20, $(AT)_7$, Materials and Methods) with different groove widths to determine the selectivity of these compounds. The duplex sequences with varying groove-widths were chosen to eliminate any bias towards a particular groove dimension and, also, these strands contain a variety of sites known to favor various modes of ligand binding in duplex DNA, including A-tracts and both pyrimidine–purine steps. The azacyanine ligands bound poorly to

all duplex DNA strands investigated and the SPR data could not be fit to an exponential binding curve (Figure 5.8-5.10). The upper limit for the K_A in each **Aza3**– and **Aza4**–dsDNA (double stranded DNA) pair is in the range of 10^4 M^{-1}. Remarkably, no **Aza5**–duplex interaction was detected under the SPR conditions used (Figure 5.10, top right). Fluorescence binding data from Dr. Nick Hud's group also confirm that these ligands exhibit marked selectivity for quadruplex over duplex-DNA and over 100-fold in the case of **Aza5** [30]. This group of compounds is clearly quite promising for development as selective quadruplex targeting agents.

5.3.2 Evaluation of Naphthalene Diimides

A series of di- (2ND), tri- (3ND) and tetra-substituted (4ND) analogs of naphthalene diimides (Figure 5.4) with varying lengths of side-chains and end groups were synthesized in Dr. Stephen Neidle's laboratory. SPR was used to quantitatively evaluate the interaction between naphthalene diimides with a wide range of oncogenic and telomeric DNA sequences (Materials and Methods) in order to gain insight into the selectivity of this class of ligands. The conformational diversity exhibited by the various quadruplex-forming sequences (Figure 5.2) makes them very interesting to characterize their ligand binding properties. Table 5.2 lists the equilibrium binding constants of a series of naphthalene diimides with several sequences. The binding constants were obtained using the procedure described in 5.2.2 and 5.3.1. From the Table 5.2, it is readily apparent that all the NDs exhibit very little selectivity for duplex DNA - the first requirement in developing ligands - to have virtually zero affinity for non-target sequences. The response exhibited in SPR with increasing concentrations of the ligands were very low (<5 RU) to obtain reliable binding data for interaction with duplex DNA (sensorgrams for duplex not shown).

SPR results of ND analogs with several quadruplex-forming sequences show that they can selectively distinguish between quadruplex conformations and, particularly, with high affinity for Tel22, *c-kit1* and *c-kit2* sequences (Figure 5.11-5.16, sensorgrams for 2ND01 not included). Due to poor solubility of 3ND07 under SPR conditions, reliable sensorgrams could not be obtained. The wild-type human telomeric sequence Tel22 is known to adopt multiple conformations under physiological conditions of K^+, with hybrid-1 and hybrid-2 (Figure 5.2-A, B) as the major conformers [41-42]. Exposed terminal quartets and the three accessible grooves of the hybrid conformation should make it a relatively easy target for ligand recognition. The oncogenic promoter quadruplex sequences of *c-kit1* and *c-kit2* folds into a parallel topology also with exposed terminal quartets and accessible grooves [43] (Figure 5.2-C). This scaffold can also be exploited for selective targeting with small molecules. SPR results show all the compounds binding to the aforementioned quadruplex motifs, but significant differences can be seen in the sensorgrams (Figure 5.11-5.16). The most striking is in the dissociation rate, which is much slower for 4-ND and 3-ND analogs than for 2-ND (Table 5.2, Tel22, k_d). Because of surface absorption of the compounds in the initial period of injection, it is not possible to quantitatively determine the association kinetics constants. For the dissociation reaction, however, it is clear that the disubstituted ND dissociates in the first few seconds of buffer flow while the apparent half-life for dissociation ($1/k_d$) of the tetra-substituted ND is approximately 35 s (Table 5.2, Tel22, k_d). The **3ND03** (Figure 5.16, Table 5.2, k_d = 0.046 s^{-1}) analog has dissociation rates that are more similar to the 4-NDs. The sensorgrams for all compounds reach a steady-state plateau between 100-200 s after initiation of compound injection (Figure 5.11-5.16). The steady-state RU values were determined by averaging in the steady state region and were plotted versus the free compound concentration in the flow solution for determination of the equilibrium binding constants (Figure 5.11-5.16: bottom left, bottom right panels). Analysis of the binding of ND analogs shows a single very strong binding site and a weaker secondary site with a K ca. 50x weaker. 4ND analogs show

much higher affinity than 3ND and 2ND counterparts. The binding of **2ND01** is almost 10x lower with K= 4.5 x 10^6 M^{-1} and a still weaker second binding site. Interestingly, all the compounds exhibited higher affinity for Tel22 conformation than the Tel24 sequence (Table 5.2). Tel24 is shown to predominantly fold into a hybrid-1 conformation with capping base-pairs at the terminal quartets. The capping units might have a negative effect on the ligand stacking at the terminal quartet resulting in lower binding affinities. It is also readily apparent from equilibrium binding constants (Table 5.2) that all the ND analogs have very little selectivity towards *bcl-2* sequence. NMR structural studies have shown that *bcl-2* sequence folds into a very stable hybrid-2 type quadruplex conformation [44] (Figure 5.2, D). The hybrid-2 core is usually comprised of three accessible grooves of varying geometries, but the terminal tetrads are capped by loop bases that can further stabilize the conformation. End-stacking ligands typically require the terminal tetrads to be completely exposed in order to have a maximum stacking effect, and any "blockage" would adversely affect their interaction at the ends. In this *bcl-2* sequence the terminal quartets are flanked by very stable capping structures (A10·T15 base-pair at 5'-end, and a stable G·C cap at the 3'-end) that would effectively prevent any ligand binding at the ends. Moreover, even with accessible grooves, ND analogs exhibited very weak to zero binding (sensorgrams not shown with *bcl-2* sequence) for this conformation suggesting that the sidechains might be too bulky to effectively recognize the grooves. All the ND analogs except **4NDO1** (Table 5.2, K=8.9 x 10^6 M^{-1}) exhibited very weak response (< 15 RU) even for some of the highest ligand concentrations (1 μM - 10 μM) to accurately determine the binding constants and, in most cases, the binding constants were less than 10^5 M^{-1}.

SPR results of ND compounds show that the binding constants observed for the *c-kit1* and *c-kit2* are among the strongest observed for this quadruplex architecture [38b] (Table 5.2).

The proto-oncogene promoter quadruplex forming sequences of *c-kit1* and *c-kit2* are shown to fold into a parallel conformation [43] with accessible grooves and exposed terminal tetrads (Figure 5.2-C). The tetra-substituted analogs displayed higher affinities for *c-kit1* and *c-kit2* than the tri-substituted and di-substituted counterparts (Table 5.2). The binding of **2ND01** is almost 10x lower for *c-kit2* in comparison with 4ND analogs, and exhibited virtually no binding for *c-kit1* sequence (sensorgrams not shown). Among all the tetra-substituted analogs, **4ND02** displayed the highest binding (*c-kit1*: K=34.3 x 10^6 M^{-1}, *c-kit2*: K=47.3 x 10^6 M^{-1}). However, other tetra-substituted analogs also displayed high binding affinities for these two sequences. Therefore, a clear correlation between the length of the alkylamino side-chain and the binding affinities could not be established. However, tetra-substituted analogs generally displayed an enhanced affinity relative to that of the di- or tri-substituted NDs. Interestingly, SPR results also show the ND analogs with very little selectivity for *c-myc* quadruplex sequence (Table 5.2).The 19mer *c-myc* also folds into a very stable parallel quadruplex topology as shown from NMR studies [7, 45]. In this parallel topology (Figure 5.2, E), three of the grooves are effectively blocked by the diagonally running loop bases, but the terminal tetrads are completely exposed due to the absence of any capping structures. End-stacking ligands should effectively recognize the terminal tetrads due to the lack of any steric hindrance from the capping bases. Surprisingly, this was not observed with the ND analogs with the *c-myc* sequence. A plausible explanation for this can be the long alkylamino sidechains tethered with ring systems in the end being sterically hindered by the loop bases running along the quadruplex grooves.

FRET thermal melting studies of naphthalene diimides were performed on selective quadruplex systems and a duplex sequence by Dr. Stephen Neidle's group. The low affinity of NDs for duplex DNA is clearly evident from the melting data (Table 5.3). All the NDs exhibited very small increase in Tm upon complex formation with duplex DNA except for **3ND03** (ΔTm

~ 11.5°C) and **4ND02** (ΔTm ~ 7.5°C). Systematic study of several other ND analogs has shown that tri-substituted analogs showed slightly higher ΔTm for duplex DNA than the tetra-substituted or di-substituted analogs, probably because of steric reasons [38b]. Thermal melting analyses with various quadruplex systems show the ligands possess the highest stabilization potential for the quadruplex systems. The tetra-substituted analogs showed highest ΔTm for both oncogenic and telomeric quadruplex conformations than the tri- or di-substituted counterparts. **2ND01** showed very small change in ΔTm for the different quadruplex systems (Tel22-5.25; *c-kit1*-2.5; *c-kit2*-7.75), further validating the importance of increasing number of sidechains in NDs contributing to maximize the potential interaction with the loops or the grooves. These results are also in excellent correlation with the SPR studies for the high selectivity exhibited by tetra-substituted analogs for the quadruplex systems. Interestingly, **4ND09** showed high selectivity for Tel22 (ΔTm ~27.8°C) than *c-kit1* (ΔTm ~0°C) and *c-kit2* (ΔTm ~4.25°C) sequences. It is important to design highly-tuned ligands to discriminate between almost identical micro-environments of different target sequences, such as the ones exhibited by terminal tetrads of Tel22 and *c-kit1/2* conformations, and these compounds illustrate one promising route for such design.

Figure 5.7-A shows a preliminary model of **4ND08** bound at the 3'-end of the human telomeric quadruplex model system after 5 ns of molecular dynamics simulation performed at Dr. Stephen Neidle's lab. The naphthalene core of 4ND08 is optimally stacked on top of the terminal tetrad with the four tethered alkyl-amino side chains making favorable hydrophobic interactions with the loops or the grooves of the quadruplex system. These results are highly encouraging because the ND platform can be further exploited to develop improved class of ligands that can exhibit virtually zero affinity for the plethora of duplexes in the genome and with greater selectivity for specific quadruplex systems.

5.3.3 Evaluation of Diaryl Ureas

Several modifications were performed along the phenyl moieties of the 1,2-diphenyl scaffold (Figure 5.5) by Dr. Stephen Neidle's group to evaluate the effect of the side chains and ring systems on quadruplex recognition. SPR was used to quantitatively characterize the interaction of these ligands against quadruplex sequences originating from the human telomere (Tel22), as well as from the *c-kit1, c-kit2* and *c-myc* proto-oncogenes. G-quadruplex selectivity was further assessed using the Dickerson duplex sequence. Several of the compounds with a triazole linker (Figure 5.2: WD419, WD422, WD423 and WD442) at the phenyl rings exhibited poor solubility even at modest concentrations required for SPR, as a result, these compounds could not be evaluated. In terms of G-quadruplex affinity, SPR produced equilibrium binding constants which fit a single strong binding-site model with a significantly weaker secondary binding observed in most cases (Figures 5.17-5.19, Table 5.4). The SPR results for quadruplex sequences indicate the strongest binding for each compound is to *c-kit2,* with *c-myc, c-kit1* and Tel22 having weaker and more similar affinities. The SPR results indicate a distinct order of selectivity with *c-kit2* > *c-myc* > *c-kit1* > Tel22, with the interaction of ligand **WD313** being particularly enhanced for the *c-myc* quadruplex (K=14.1 x 10^6 M^{-1}), which also has slower dissociation kinetics. Interestingly, none of the compounds exhibited any binding to Tel24 or *bcl-2* quadruplex-forming sequences (sensorgrams not shown), also observed in the case of naphthalene diimides (Section 5.3.2). The side-chain length dependence of G-quadruplex interaction was also assessed by SPR measurements. SPR demonstrated that the n = 1 side chain of **WD313** was optimal for general G-quadruplex DNA interaction over **WD308** (n = 2) and **WD263** (n = 3) (Table 5.4). SPR results further showed the G-quadruplex:duplex DNA selectivity is significantly enhanced, with no duplex DNA interaction detected (K = < 1 x 10^4) under the SPR experimental conditions used. This represents at least 50-100 fold selectivity for G-quadruplex vs. duplex DNA.

FRET based thermal melting studies were performed by Dr. Stephen Neidle's group to assess the G-quadruplex stabilization and selectivity of the ligand library (Table 5.5). The ΔT_m values for the ligands showed that compounds with n = 1 side chain are less effective quadruplex stabilizers than the n = 2/3 analogues, which show approximately equivalent behavior. The FRET melting data are directly proportional to side-chain length, a trend which is not consistent with the results from the SPR assay. The results are for different temperatures and it may be that in this instance affinity and thermal stability cannot be directly compared. For the Tel22, *c-kit-1*, *c-kit-2* and duplex DNAs where FRET and SPR results can be compared, however, there is good agreement (Table 5.4 and 5.5). Also, FRET-based assessment of the interactions of these ligands with the *c-myc* G-quadruplex was not possible, due to the exceptional stability of the *c-myc* quadruplex fold. The best agreement observed between the FRET and SPR assays is for G-quadruplex vs. duplex DNA selectivity.

5.4 CONCLUSION

Azacyanines, naphthalene diimides and diaryl urea analogs have been well characterized in this study using the biosensor technique. The high selectivity exhibited by these classes of ligands over duplex-DNA, and also the selectivity exhibited between quadruplex conformations makes them an interesting set of ligands to further develop as potential therapeutic agents for a quadruplex-mediated mechanism of controlling biochemical processes. The planar portion of DNA-directed drugs subtly modulates their ability to recognize the terminal quadruplex structural arrangements. Quadruplex end-stacking ligands, in most cases, have displayed better binding affinities when tethered with long side-chains. The side chains have potential to make additional interactions with the grooves, loops or the backbone of the quadruplex. Aza5, with a methoxy substituent, displayed slightly better affinity than Aza3 or Aza4. Therefore, potential opportunities exist where the tethered side-

chain moieties of azacyanines could be manipulated to increase their affinity for quadruplexes. The extended, fused aromatic ring systems of naphthalene diimides have shown to have affinities very similar to well known G-quadruplex binders structurally related to acridines or TMPyP4. The tetra-substituted analogs displayed better quadruplex affinities than the tri- or di-substituted counterparts. The additional side-chains of the 4NDs potentially engage in additional interactions with various structural components of quadruplex systems. The synthetic ease of naphthalene diimides can facilitate studying more systematic substitutions along the planar ring systems, coupled with the opportunities to modulate side-chains for better quadruplex affinities. Diarylurea-based scaffolds, which constitute non-polycyclic, non-fused ring systems, have been shown to have significant selectivity for parallel-type quadruplex systems, with an almost 20-fold weaker selectivity for hybrid-type quadruplex systems. The non-polycyclic ring systems are perhaps more optimized to recognize terminal quartets of parallel quadruplex architecture. Therefore, diarylurea-based compounds can be finely-tuned to discriminate between very similar quadruplex architectures with subtle structural differences.

The number of quadruplex-interactive ligands has increased dramatically in the past few years; however, there has been very limited success in effectively converting these ligands into potential therapeutic agents. This is partially due to the lack of a better understanding of quadruplex-ligand interactions at the molecular level. In addition, the "off-target" effects of these compounds could also contribute to the variations of cellular effects. Systematic modification of ligands coupled with better methodologies to characterize quadruplex-ligand complexes will significantly augment the current knowledge towards developing more potent small molecules.

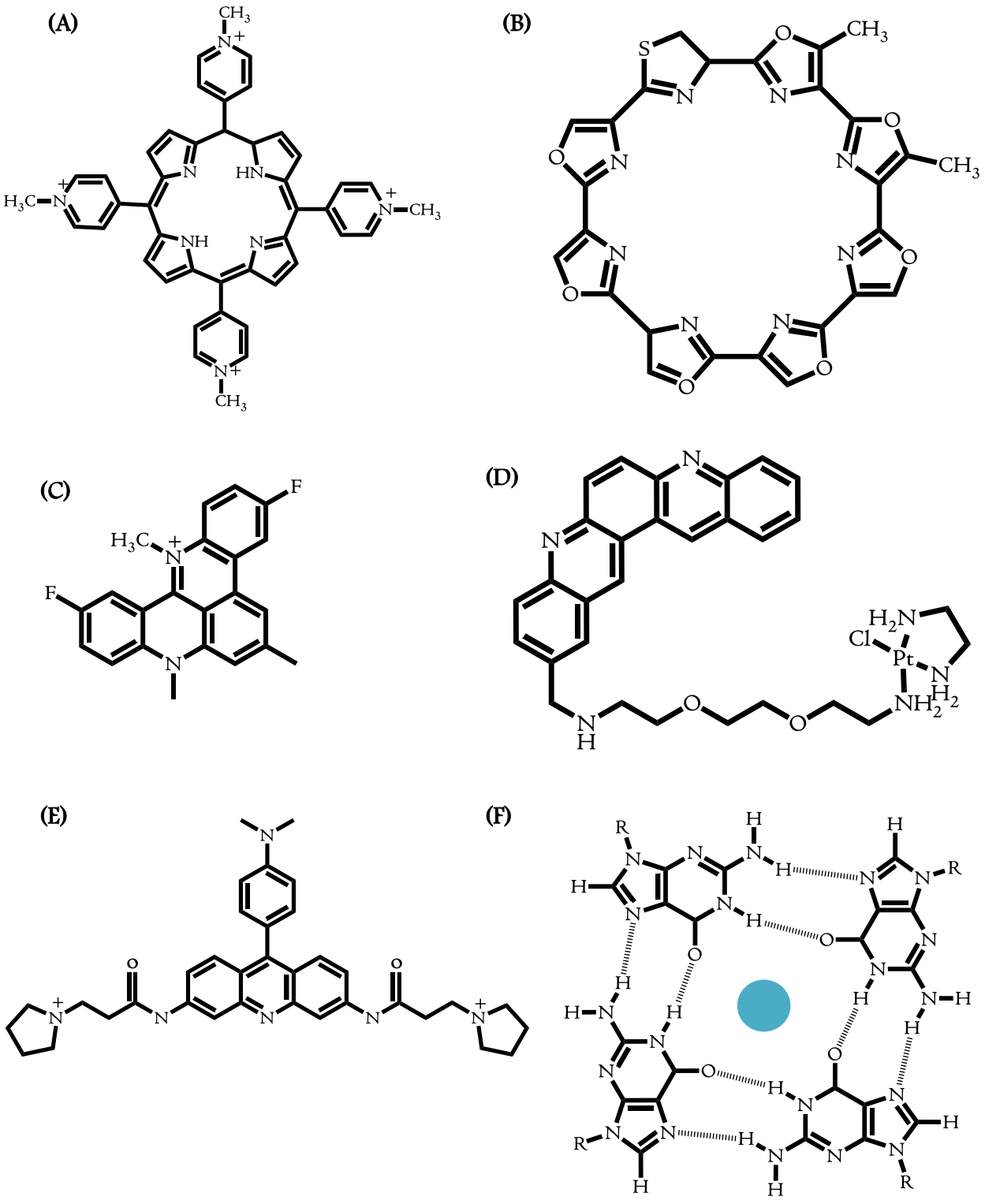


Figure 5.1: Structures of (A) TMPyP4, (B) Telomestatin, (C) RHPS4, (D) Pt-MPQ, (E) BRACO-19, and (F) G-quartet with bound cation shown as a blue circle.

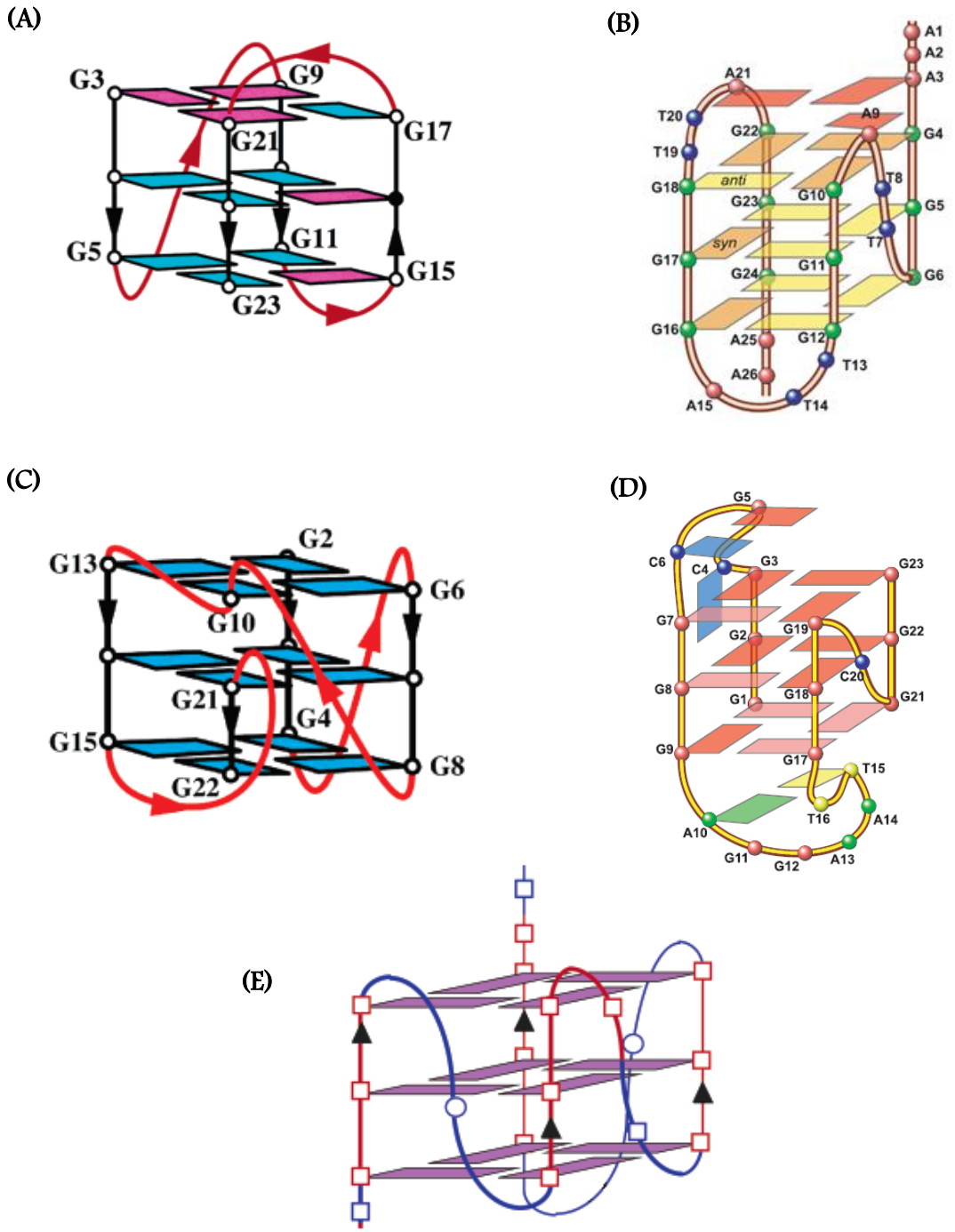
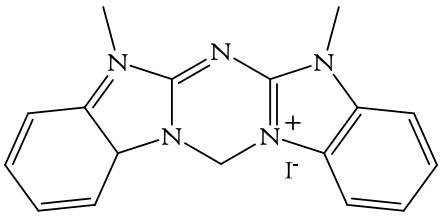


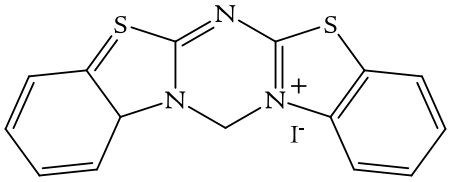
Figure 5.2: Folding patterns of different quadruplex-forming motifs used in SPR-Biosensor studies.

(A) Tel24: Hybrid1 [41b], (B) Tel26: Hybrid1 [41a], (C) c-kit1/c-kit2: Parallel [43], (D) bcl-2: Hybrid2 [44b], and (E) c-myc: Parallel [5b, 7, 45]. All the images are directly obtained from their published journals without further permission.

Aza3



Aza4



Aza5

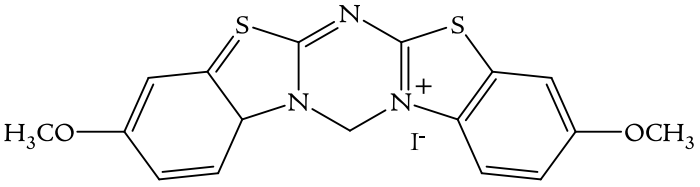


Figure 5.3: Chemical Structures of Azacyanines.

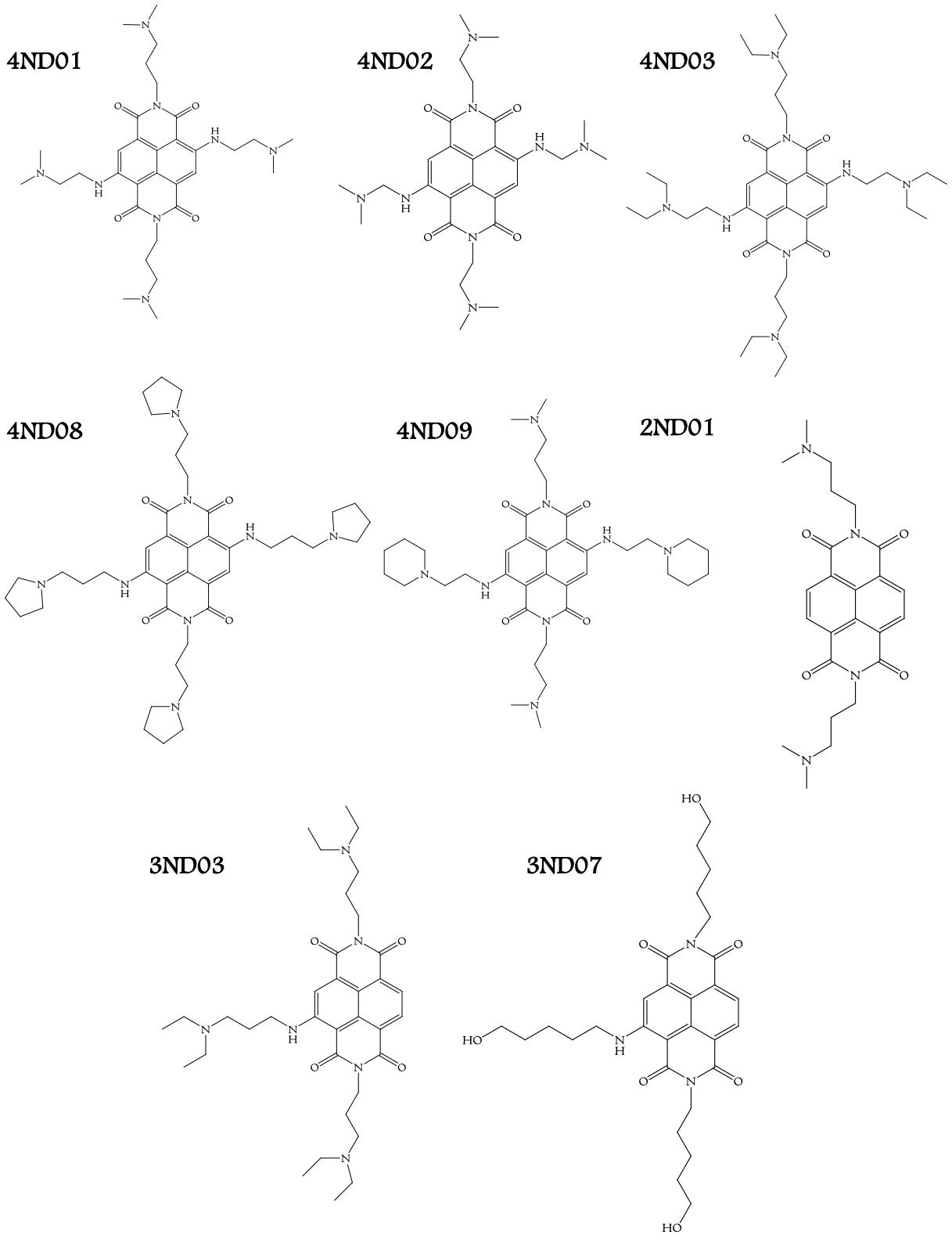


Figure 5.4: Chemical Structures of Naphthalene Diimides.

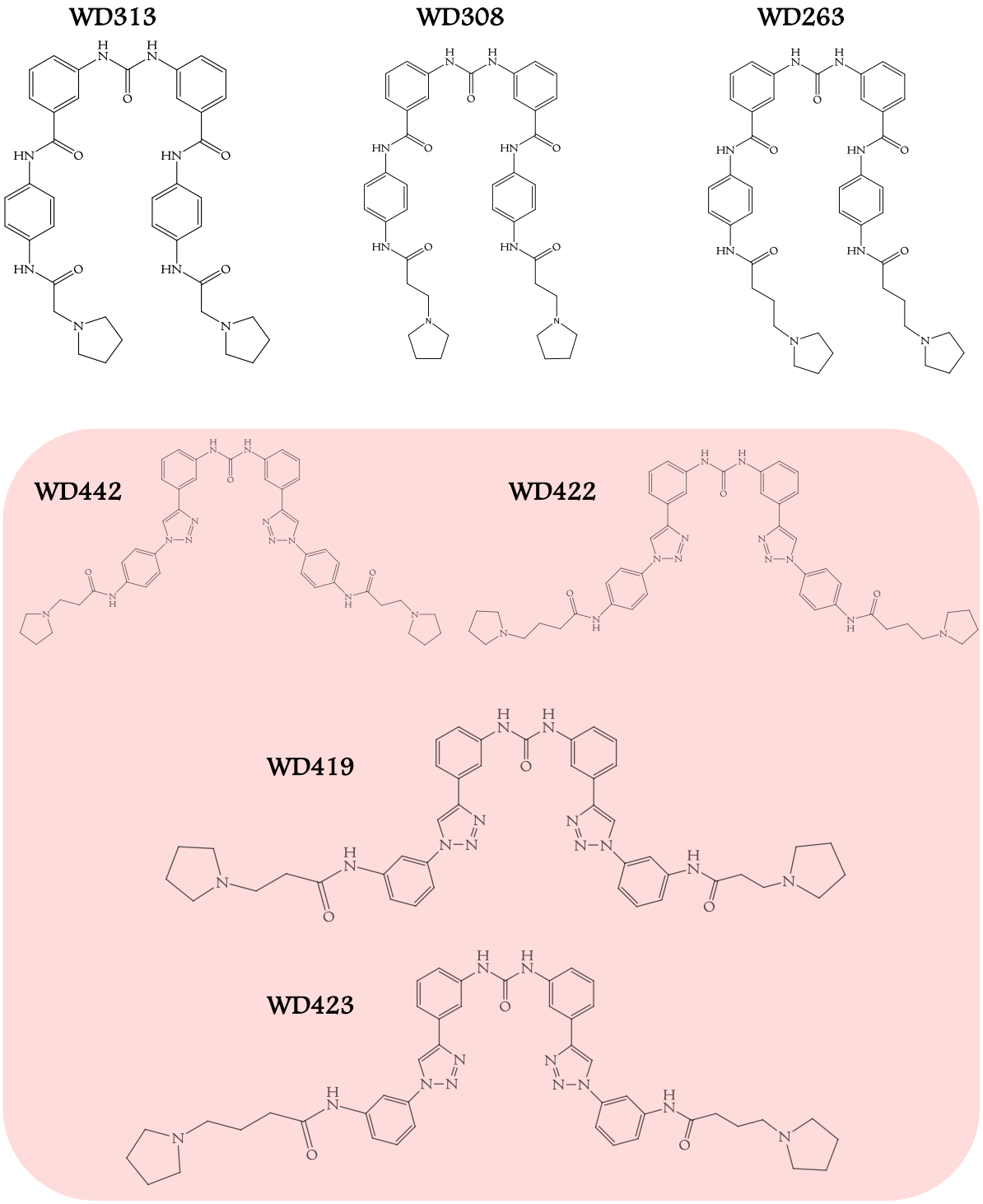


Figure 5.5: Chemical Structures of Diaryl Ureas.

Compounds highlighted could not be evaluated with SPR due to solubility issues.

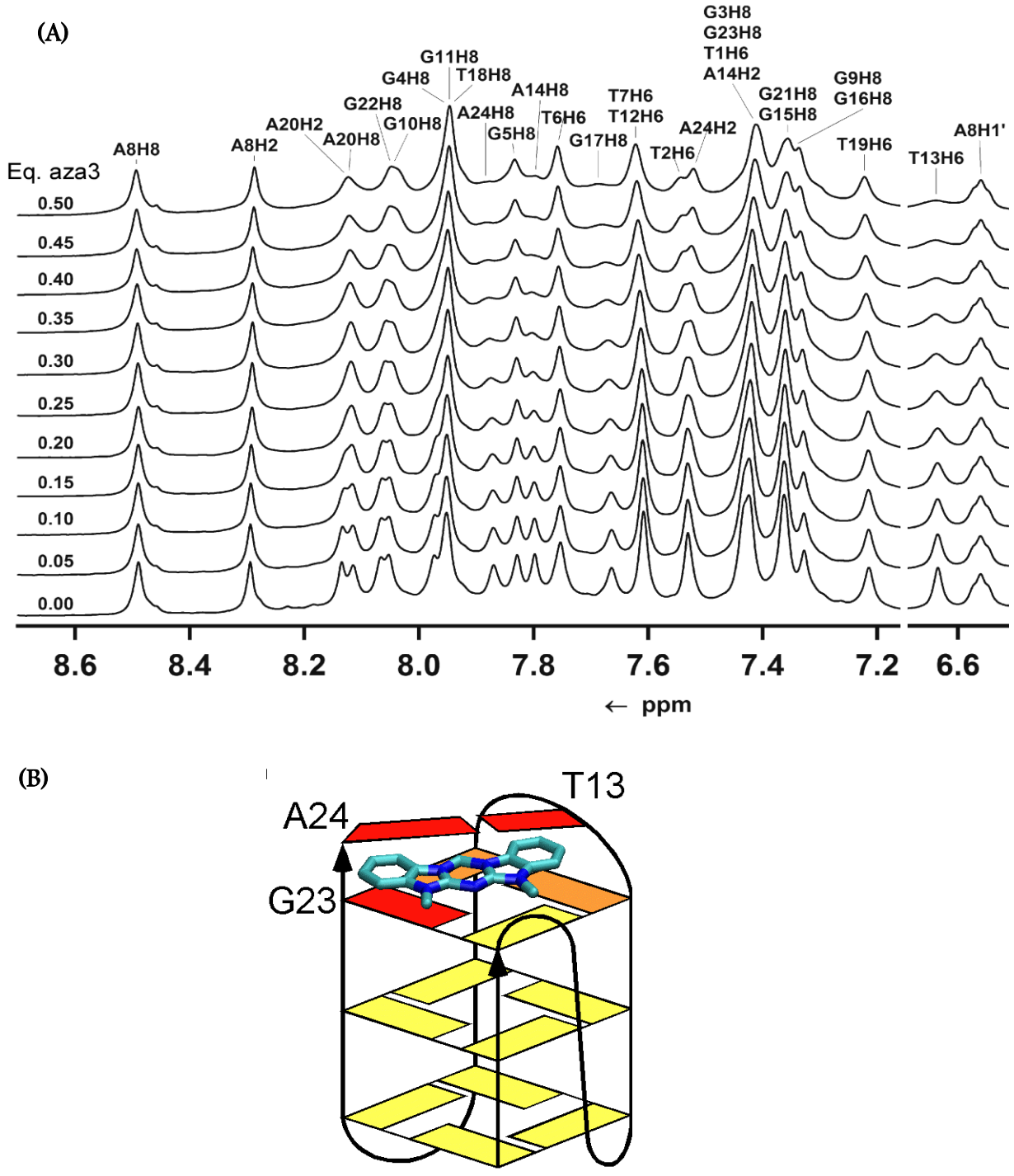


Figure 5.6: Interaction of Azacyanine-3 with Tel24 quadruplex motif.

(A) Aromatic region of ^{1}H-NMR spectra of Tel24 in the presence of Aza3. (B) A schematic representation of Aza3 bound Tel24 based on preliminary ^{1}H-NMR studies [30].

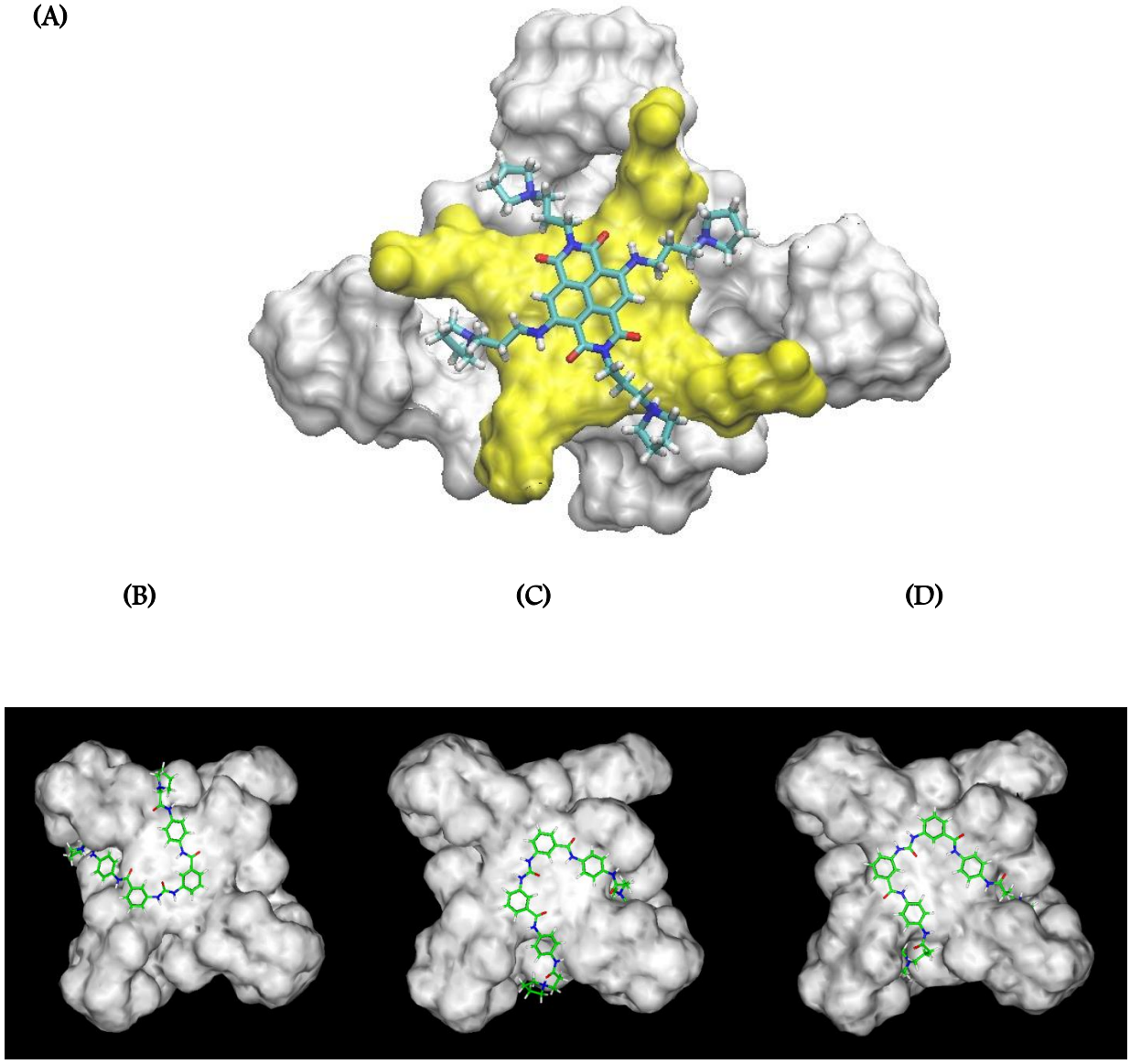
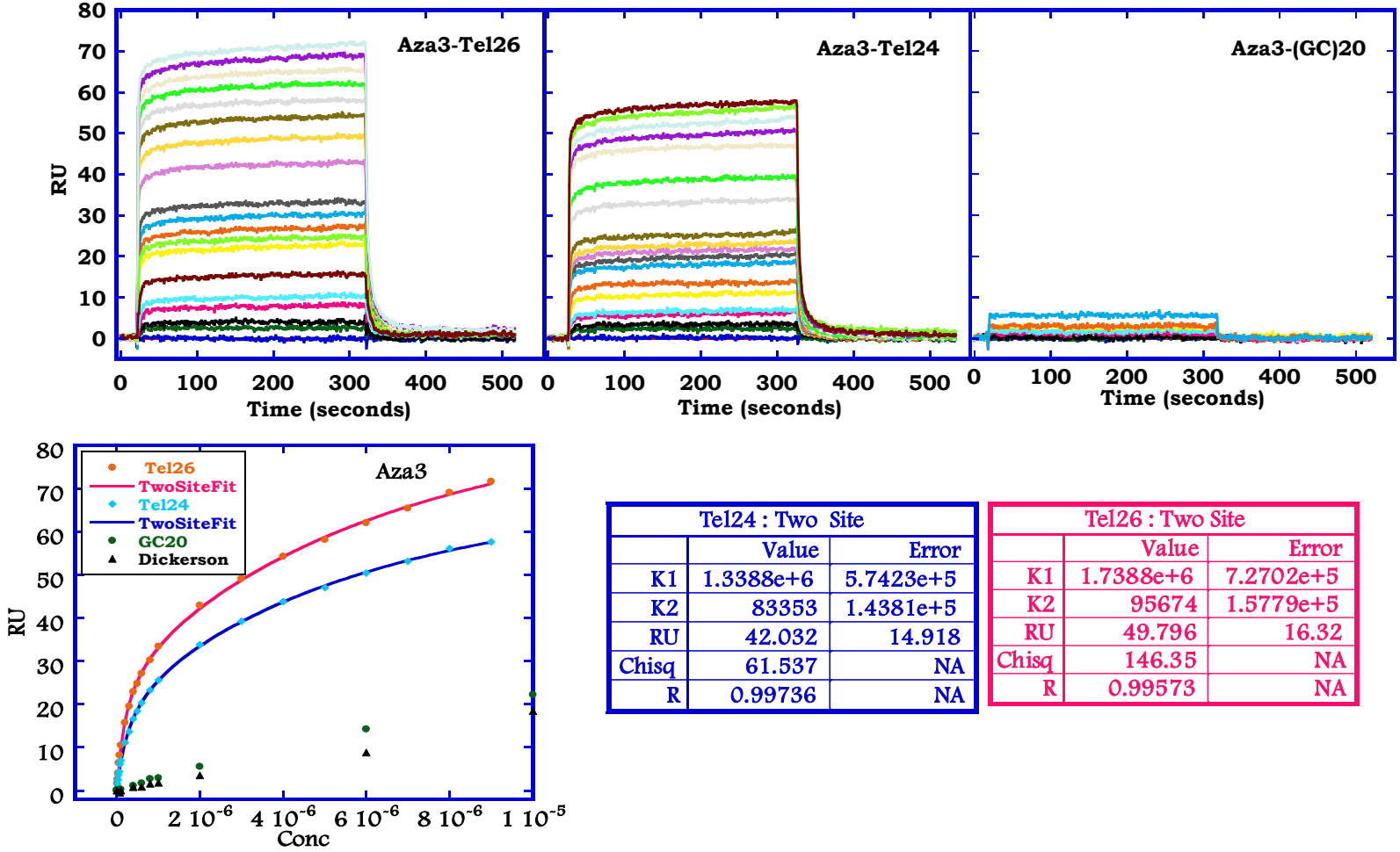


Figure 5.7: Molecular models of ligands complexed with a parallel topology of human telomeric quadruplex conformation. (A) 4ND08 (B) WD313 (C) WD308 (D) WD263.

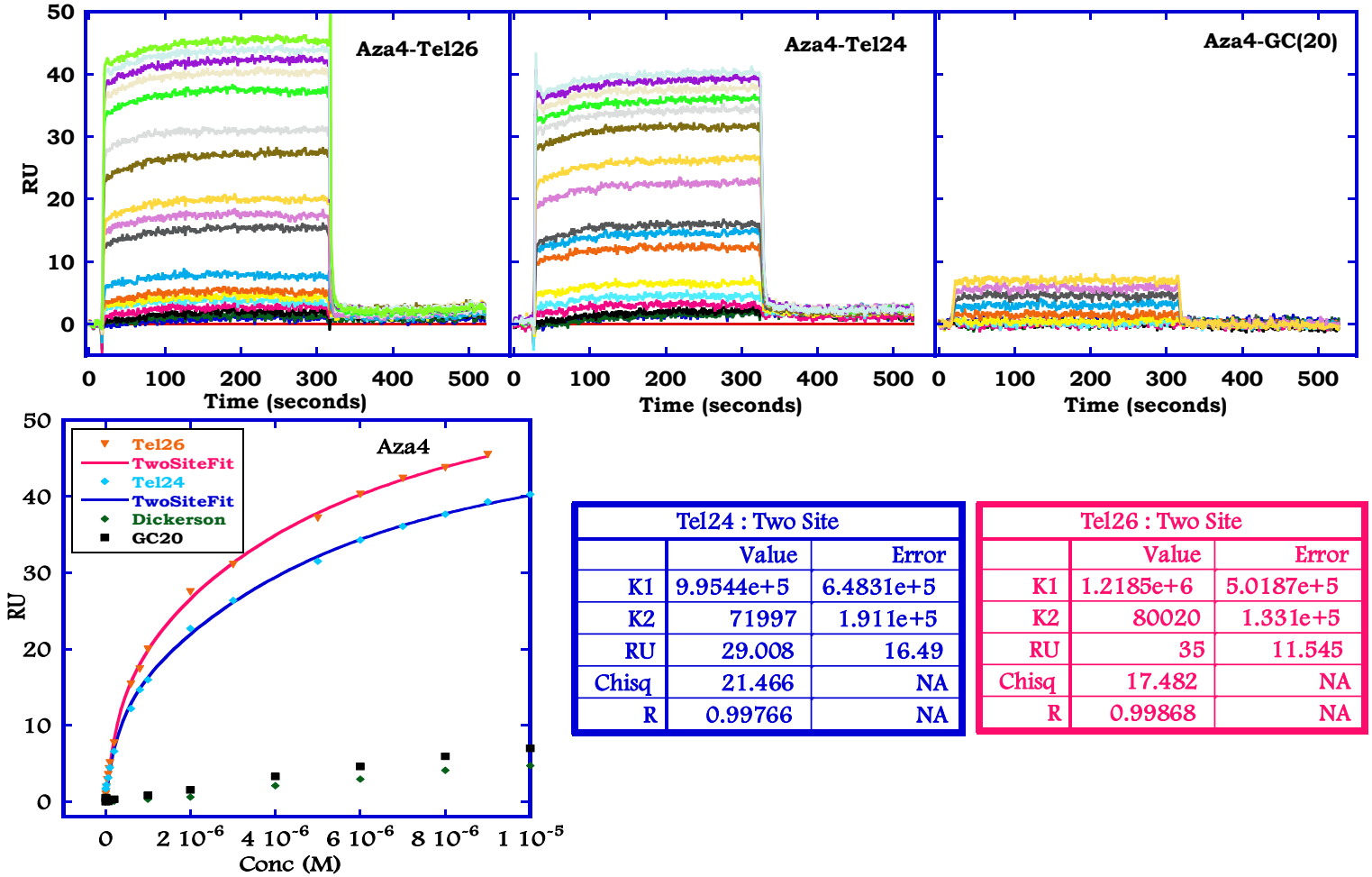


Tel24 : Two Site		
	Value	Error
K1	1.3388e+6	5.7423e+5
K2	83353	1.4381e+5
RU	42.032	14.918
Chisq	61.537	NA
R	0.99736	NA

Tel26 : Two Site		
	Value	Error
K1	1.7388e+6	7.2702e+5
K2	95674	1.5779e+5
RU	49.796	16.32
Chisq	146.35	NA
R	0.99573	NA

Figure 5.8: SPR sensorgrams for binding of Aza3 to the immobilized G-quadruplexes formed by Tel26 (top left), Tel24 (top middle) and GC(20) duplex (top right) in HEPES buffer containing 80 mM KCl at 25 °C.

The quadruplex curves range in Aza3 concentration from 100 nM for the bottom curve to 10 μM for the top curve. Steady-state binding plots (bottom left) fit to a two-site model (Materials and Methods). The concentration values are for unbound compound concentration in the flow solution. Binding parameters (bottom right) obtained by fitting to a two-site model.

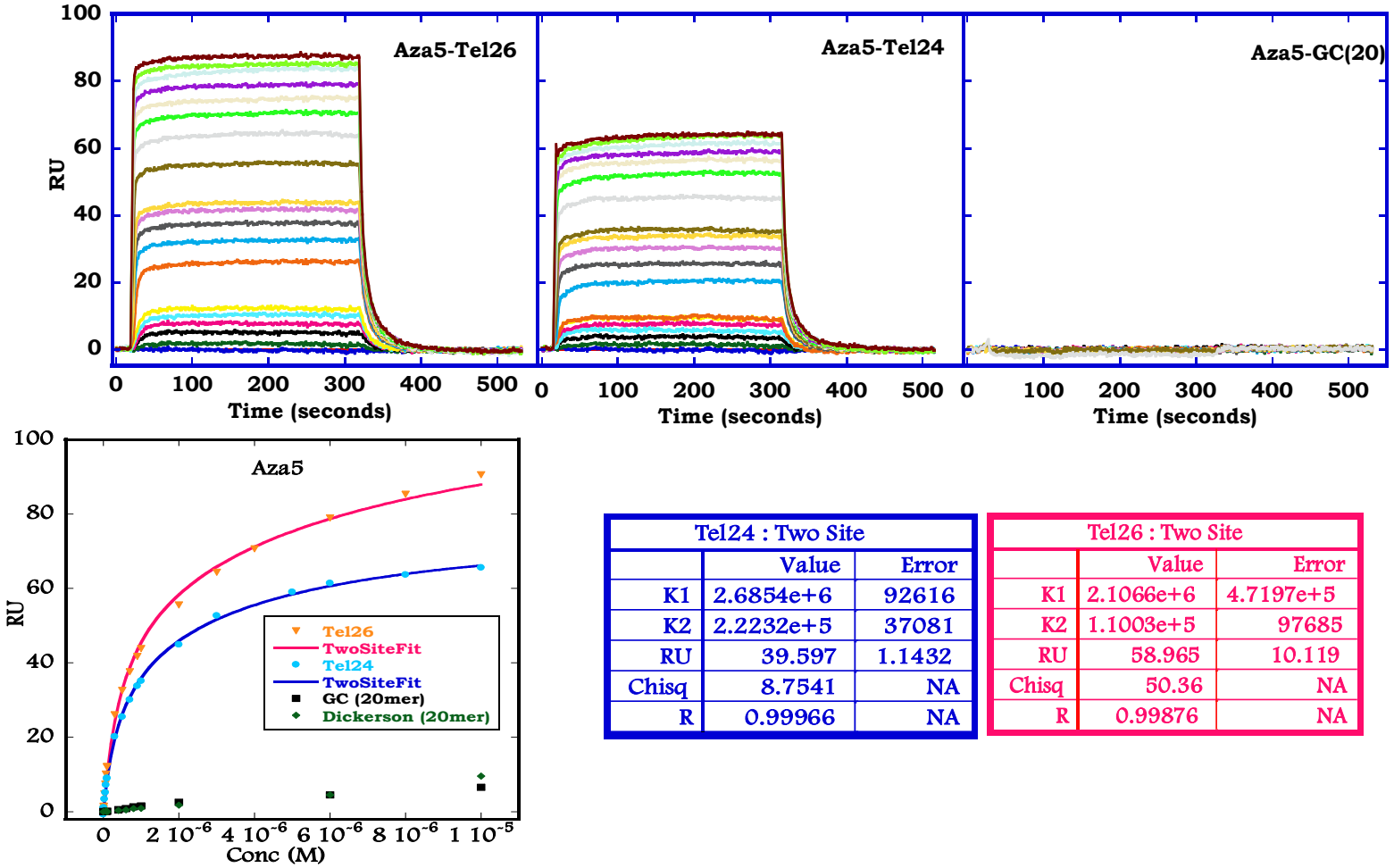


Tel24 : Two Site		
	Value	Error
K1	9.9544e+5	6.4831e+5
K2	71997	1.911e+5
RU	29.008	16.49
Chisq	21.466	NA
R	0.99766	NA

Tel26 : Two Site		
	Value	Error
K1	1.2185e+6	5.0187e+5
K2	80020	1.331e+5
RU	35	11.545
Chisq	17.482	NA
R	0.99868	NA

Figure 5.9: SPR sensorgrams for binding of Aza4 to the immobilized G-quadruplexes formed by Tel26 (top left), Tel24 (top middle) and GC(20) duplex (top right) in HEPES buffer containing 80 mM KCl at 25 °C.

The quadruplex curves range in Aza4 concentration from 100 nM for the bottom curve to 10 μM for the top curve. Steady-state binding plots (bottom left) fit to a two-site model (Materials and Methods). The concentration values are for unbound compound concentration in the flow solution. Binding parameters (bottom right) obtained by fitting to a two-site model.

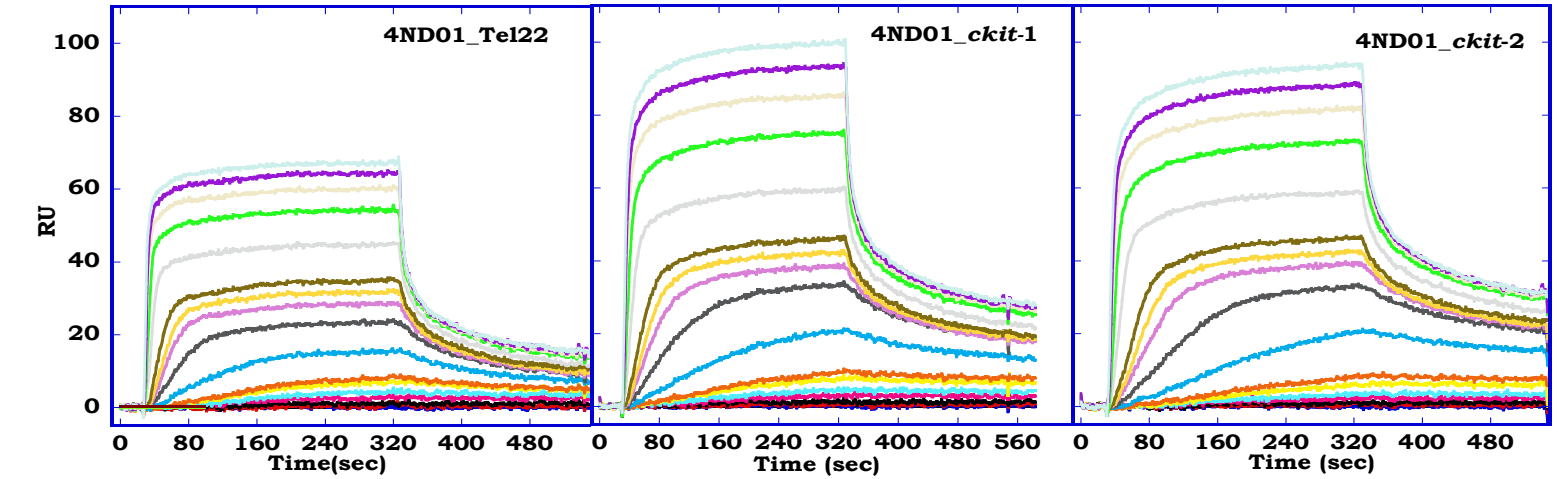


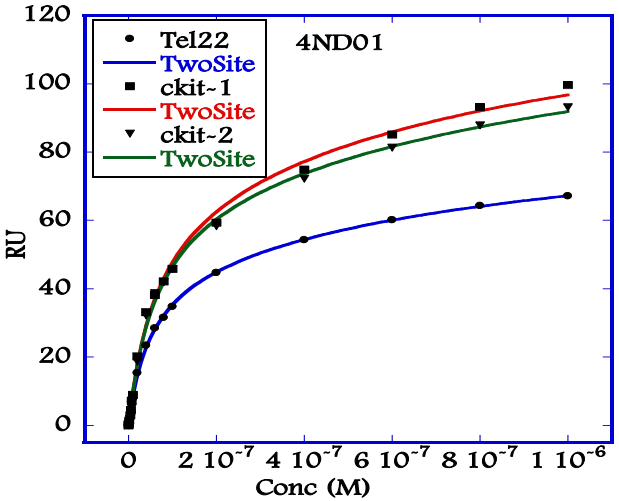
Tel24 : Two Site		
	Value	Error
K1	2.6854e+6	92616
K2	2.2232e+5	37081
RU	39.597	1.1432
Chisq	8.7541	NA
R	0.99966	NA

Tel26 : Two Site		
	Value	Error
K1	2.1066e+6	4.7197e+5
K2	1.1003e+5	97685
RU	58.965	10.119
Chisq	50.36	NA
R	0.99876	NA

Figure 5.10: SPR sensorgrams for binding of Aza5 to the immobilized G-quadruplexes formed by Tel26 (top left), Tel24 (top middle) and GC(20) duplex (top right) in HEPES buffer containing 80 mM KCl at 25 °C.

The quadruplex curves range in Aza5 concentration from 100 nM for the bottom curve to 10 μM for the top curve. Steady-state binding plots (bottom left) fit to a two-site model (Materials and Methods). The concentration values are for unbound compound concentration in the flow solution. Binding parameters (bottom right) obtained by fitting to a two-site model.





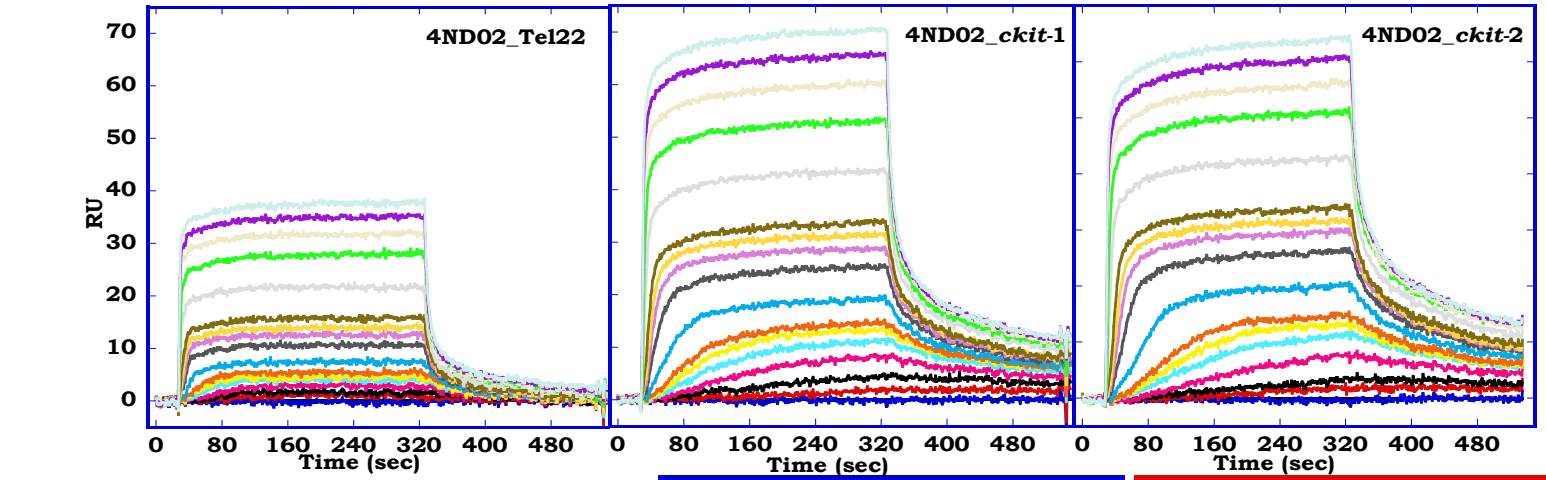
Tel22 : Two Site		
	Value	Error
K1	2.1185e+7	2.9235e+6
K2	9.1917e+5	4.6177e+5
RU	46.542	4.4852
Chisq	6.0163	NA
R	0.9997	NA

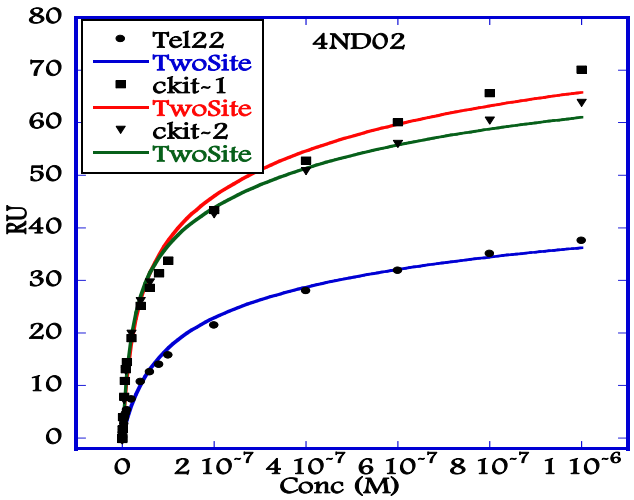
ckit-1 : Two Site		
	Value	Error
K1	1.6763e+7	6.4251e+6
K2	8.4834e+5	1.2339e+6
RU	68.445	19.686
Chisq	63.457	NA
R	0.99847	NA

ckit-2 : Two Site		
	Value	Error
K1	1.6612e+7	4.9839e+6
K2	6.3455e+5	6.5972e+5
RU	68.607	13.648
Chisq	66.375	NA
R	0.99827	NA

Figure 5.11: SPR sensorgrams for binding of 4ND01 to the immobilized G-quadruplexes formed by Tel22 (top left), *ckit-1* (top middle) and *ckit-2* (top right) in HEPES buffer containing 80 mM KCl at 25 °C.

The quadruplex curves range in 4ND01 concentration from 10 nM for the bottom curve to 1 μM for the top curve. Steady-state binding plots (bottom left) fit to a two-site model (Materials and Methods). The concentration values are for unbound compound concentration in the flow solution. Binding parameters (bottom right) obtained by fitting to a two-site model.





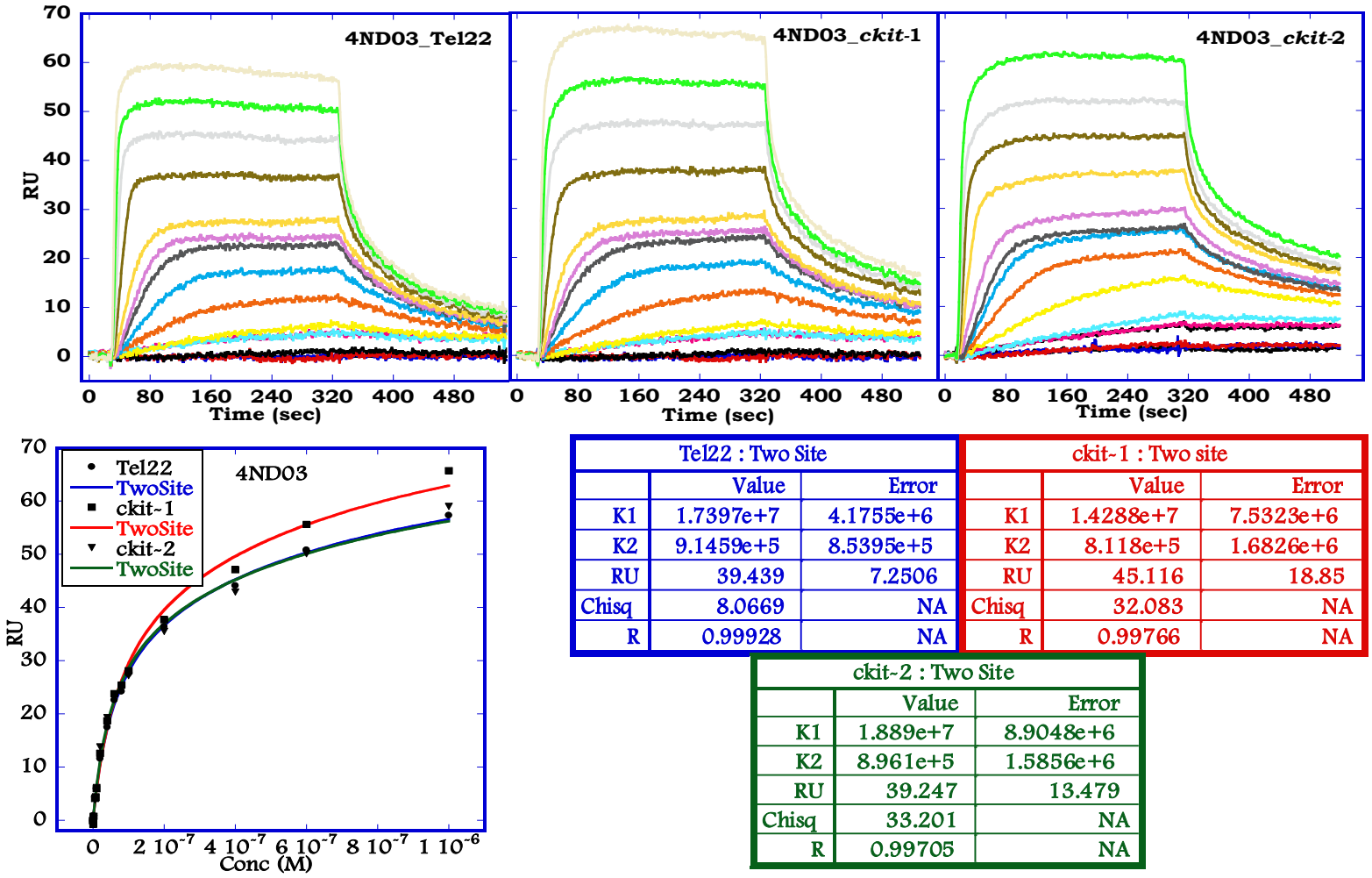
Tel22 : Two Site		
	Value	Error
K1	1.5283e+7	7.9779e+6
K2	9.6731e+5	2.1145e+6
RU	25.087	11.035
Chisq	23.723	NA
R	0.99565	NA

ckit-1 : Two Site		
	Value	Error
K1	3.4322e+7	1.3581e+7
K2	1.4188e+6	2.2114e+6
RU	41.966	11.686
Chisq	121.62	NA
R	0.99331	NA

ckit-2 : Two Site		
	Value	Error
K1	4.7277e+7	1.0841e+7
K2	1.6365e+6	1.4435e+6
RU	37.951	5.7095
Chisq	47.471	NA
R	0.99701	NA

Figure 5.12: SPR sensorgrams for binding of 4ND02 to the immobilized G-quadruplexes formed by Tel22 (top left), *ckit-1* (top middle) and *ckit-2* (top right) in HEPES buffer containing 80 mM KCl at 25 °C.

The quadruplex curves range in 4ND02 concentration from 10 nM for the bottom curve to 1 μM for the top curve. Steady-state binding plots (bottom left) fit to a two-site model (Materials and Methods). The concentration values are for unbound compound concentration in the flow solution. Binding parameters (bottom right) obtained by fitting to a two-site model.



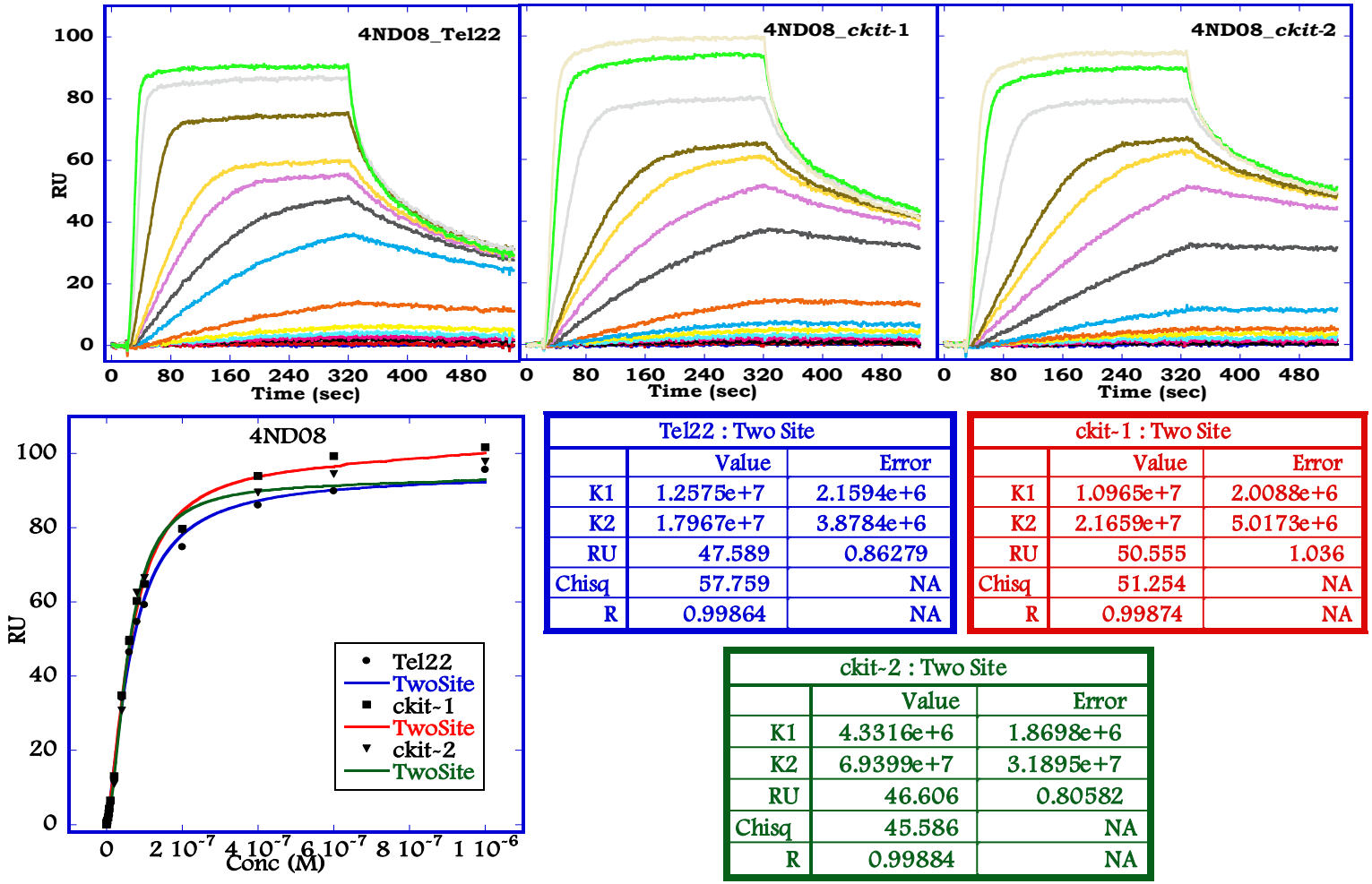
Tel22 : Two Site		
	Value	Error
K1	1.7397e+7	4.1755e+6
K2	9.1459e+5	8.5395e+5
RU	39.439	7.2506
Chisq	8.0669	NA
R	0.99928	NA

ckit-1 : Two site		
	Value	Error
K1	1.4288e+7	7.5323e+6
K2	8.118e+5	1.6826e+6
RU	45.116	18.85
Chisq	32.083	NA
R	0.99766	NA

ckit-2 : Two Site		
	Value	Error
K1	1.889e+7	8.9048e+6
K2	8.961e+5	1.5856e+6
RU	39.247	13.479
Chisq	33.201	NA
R	0.99705	NA

Figure 5.13: SPR sensorgrams for binding of 4ND03 to the immobilized G-quadruplexes formed by Tel22 (top left), *ckit-1* (top middle) and *ckit-2* (top right) in HEPES buffer containing 80 mM KCl at 25 °C.

The quadruplex curves range in 4ND03 concentration from 10 nM for the bottom curve to 1 μM for the top curve. Steady-state binding plots (bottom left) fit to a two-site model (Materials and Methods). The concentration values are for unbound compound concentration in the flow solution. Binding parameters (bottom right) obtained by fitting to a two-site model.



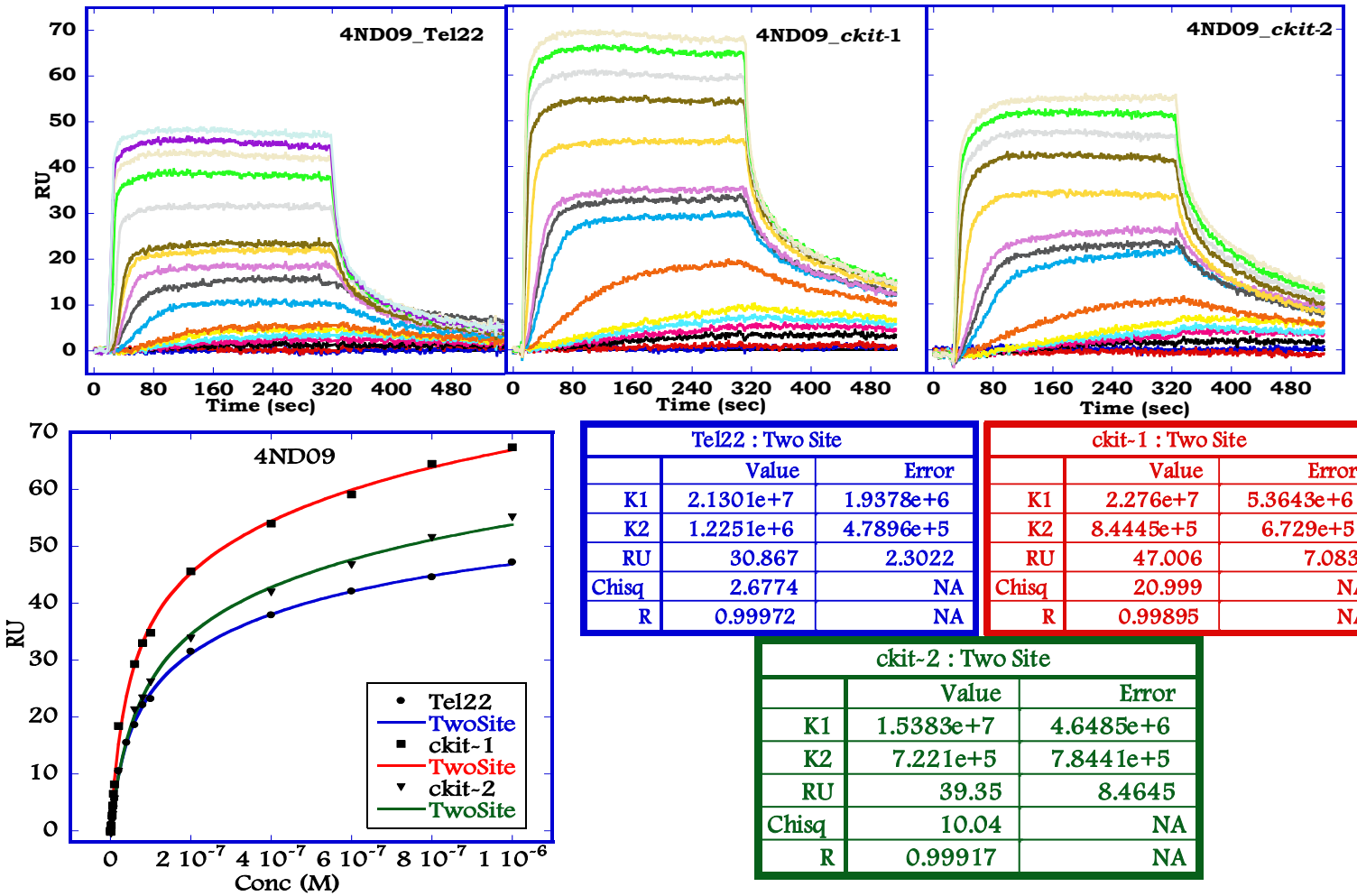
Tel22 : Two Site		
	Value	Error
K1	1.2575e+7	2.1594e+6
K2	1.7967e+7	3.8784e+6
RU	47.589	0.86279
Chisq	57.759	NA
R	0.99864	NA

ckit-1 : Two Site		
	Value	Error
K1	1.0965e+7	2.0088e+6
K2	2.1659e+7	5.0173e+6
RU	50.555	1.036
Chisq	51.254	NA
R	0.99874	NA

ckit-2 : Two Site		
	Value	Error
K1	4.3316e+6	1.8698e+6
K2	6.9399e+7	3.1895e+7
RU	46.606	0.80582
Chisq	45.586	NA
R	0.99884	NA

Figure 5.14: SPR sensorgrams for binding of 4ND08 to the immobilized G-quadruplexes formed by Tel22 (top left), *ckit-1* (top middle) and *ckit-2* (top right) in HEPES buffer containing 80 mM KCl at 25 °C.

The quadruplex curves range in 4ND08 concentration from 10 nM for the bottom curve to 1 μM for the top curve. Steady-state binding plots (bottom left) fit to a two-site model (Materials and Methods). The concentration values are for unbound compound concentration in the flow solution. Binding parameters (bottom right) obtained by fitting to a two-site model.



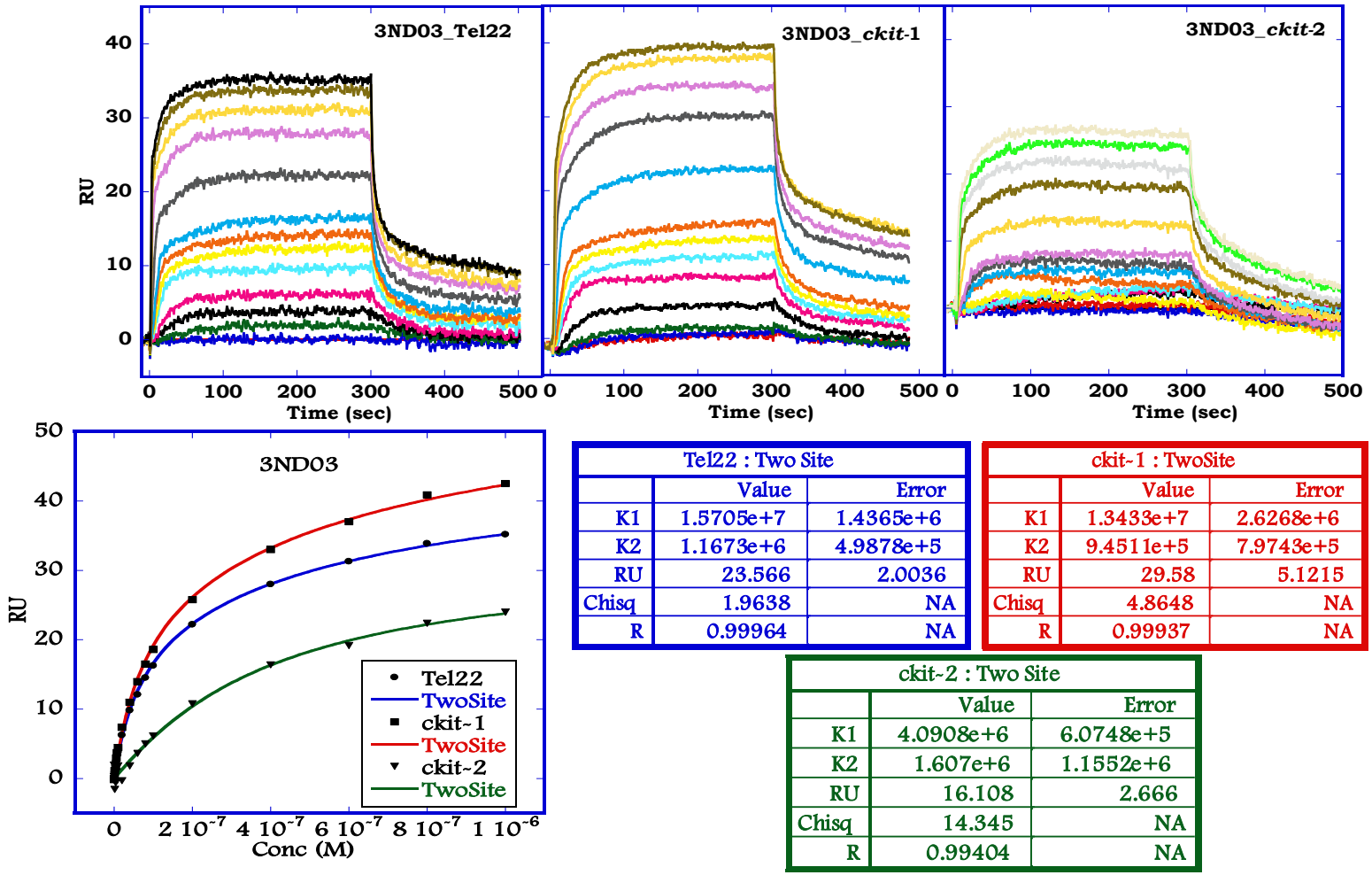
Tel22 : Two Site		
	Value	Error
K1	2.1301e+7	1.9378e+6
K2	1.2251e+6	4.7896e+5
RU	30.867	2.3022
Chisq	2.6774	NA
R	0.99972	NA

ckit-1 : Two Site		
	Value	Error
K1	2.276e+7	5.3643e+6
K2	8.4445e+5	6.729e+5
RU	47.006	7.083
Chisq	20.999	NA
R	0.99895	NA

ckit-2 : Two Site		
	Value	Error
K1	1.5383e+7	4.6485e+6
K2	7.221e+5	7.8441e+5
RU	39.35	8.4645
Chisq	10.04	NA
R	0.99917	NA

Figure 5.15: SPR sensorgrams for binding of 4ND09 to the immobilized G-quadruplexes formed by Tel22 (top left), *ckit-1* (top middle) and *ckit-2* (top right) in HEPES buffer containing 80 mM KCl at 25 °C.

The quadruplex curves range in 4ND09 concentration from 10 nM for the bottom curve to 1 μM for the top curve. Steady-state binding plots (bottom left) fit to a two-site model (Materials and Methods). The concentration values are for unbound compound concentration in the flow solution. Binding parameters (bottom right) obtained by fitting to a two-site model.



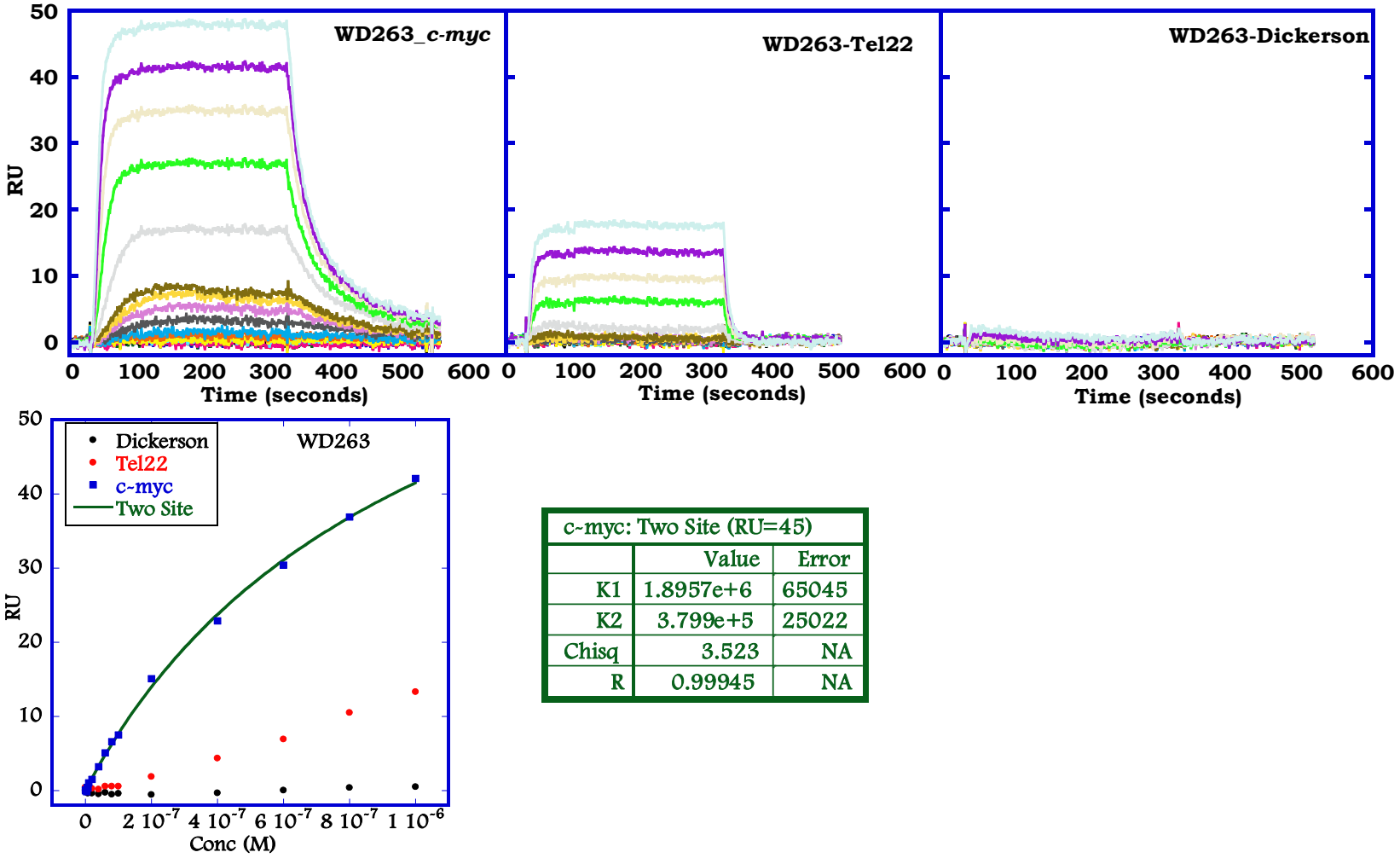
Tel22 : Two Site		
	Value	Error
K1	1.5705e+7	1.4365e+6
K2	1.1673e+6	4.9878e+5
RU	23.566	2.0036
Chisq	1.9638	NA
R	0.99964	NA

ckit-1 : TwoSite		
	Value	Error
K1	1.3433e+7	2.6268e+6
K2	9.4511e+5	7.9743e+5
RU	29.58	5.1215
Chisq	4.8648	NA
R	0.99937	NA

ckit-2 : Two Site		
	Value	Error
K1	4.0908e+6	6.0748e+5
K2	1.607e+6	1.1552e+6
RU	16.108	2.666
Chisq	14.345	NA
R	0.99404	NA

Figure 5.16: SPR sensorgrams for binding of 3ND03 to the immobilized G-quadruplexes formed by Tel22 (top left), *ckit-1* (top middle) and *ckit-2* (top right) in HEPES buffer containing 80 mM KCl at 25 °C.

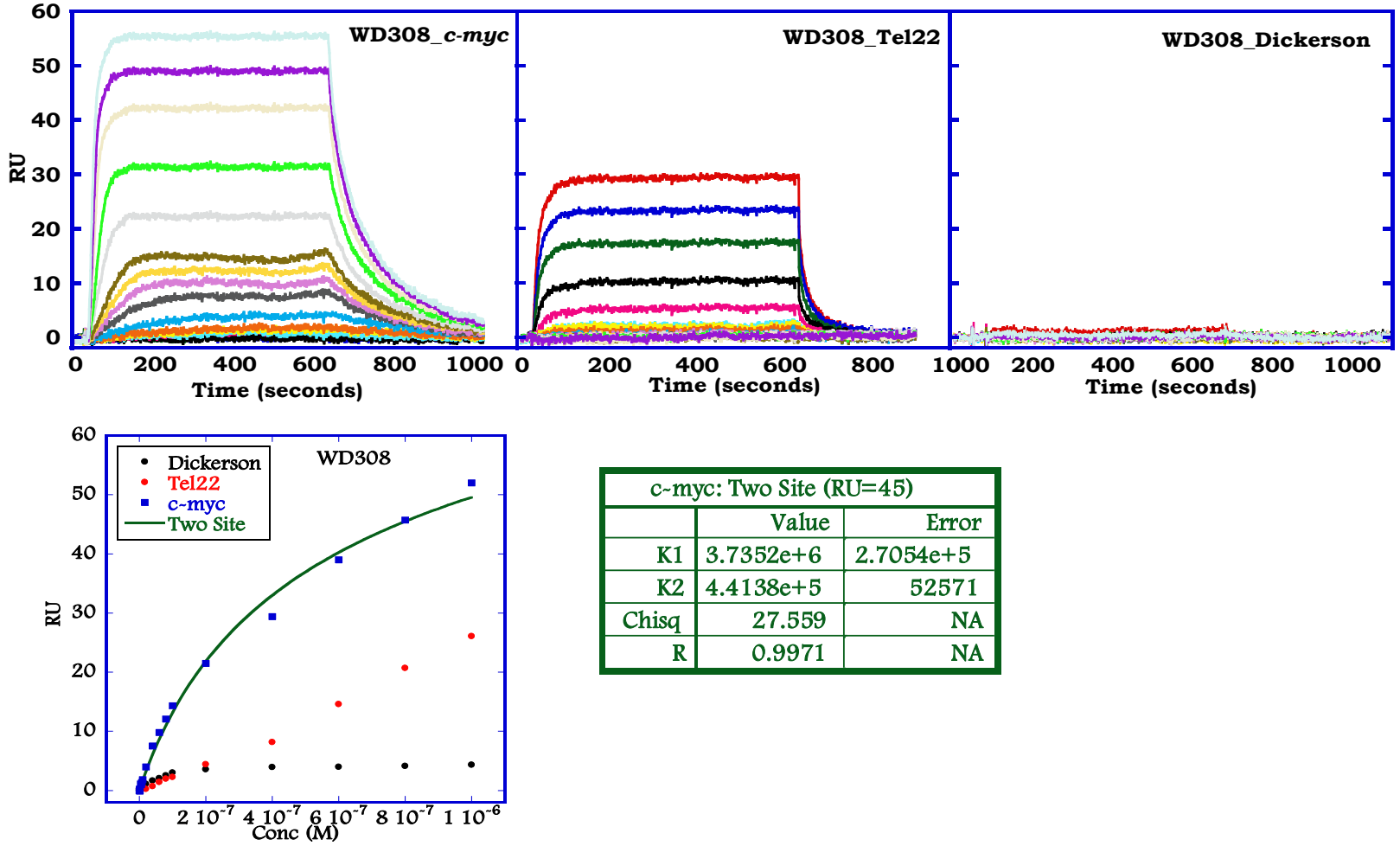
The quadruplex curves range in 3ND03 concentration from 10 nM for the bottom curve to 1 μM for the top curve. Steady-state binding plots (bottom left) fit to a two-site model (Materials and Methods). The concentration values are for unbound compound concentration in the flow solution. Binding parameters (bottom right) obtained by fitting to a two-site model.



c-myc: Two Site (RU=45)		
	Value	Error
K1	1.8957e+6	65045
K2	3.799e+5	25022
Chisq	3.523	NA
R	0.99945	NA

Figure 5.17: SPR sensorgrams for binding of WD263 to the immobilized G-quadruplexes formed by *c-myc* (top left), Tel22 (top middle) and Dickerson (top right) in HEPES buffer containing 80 mM KCl at 25 °C.

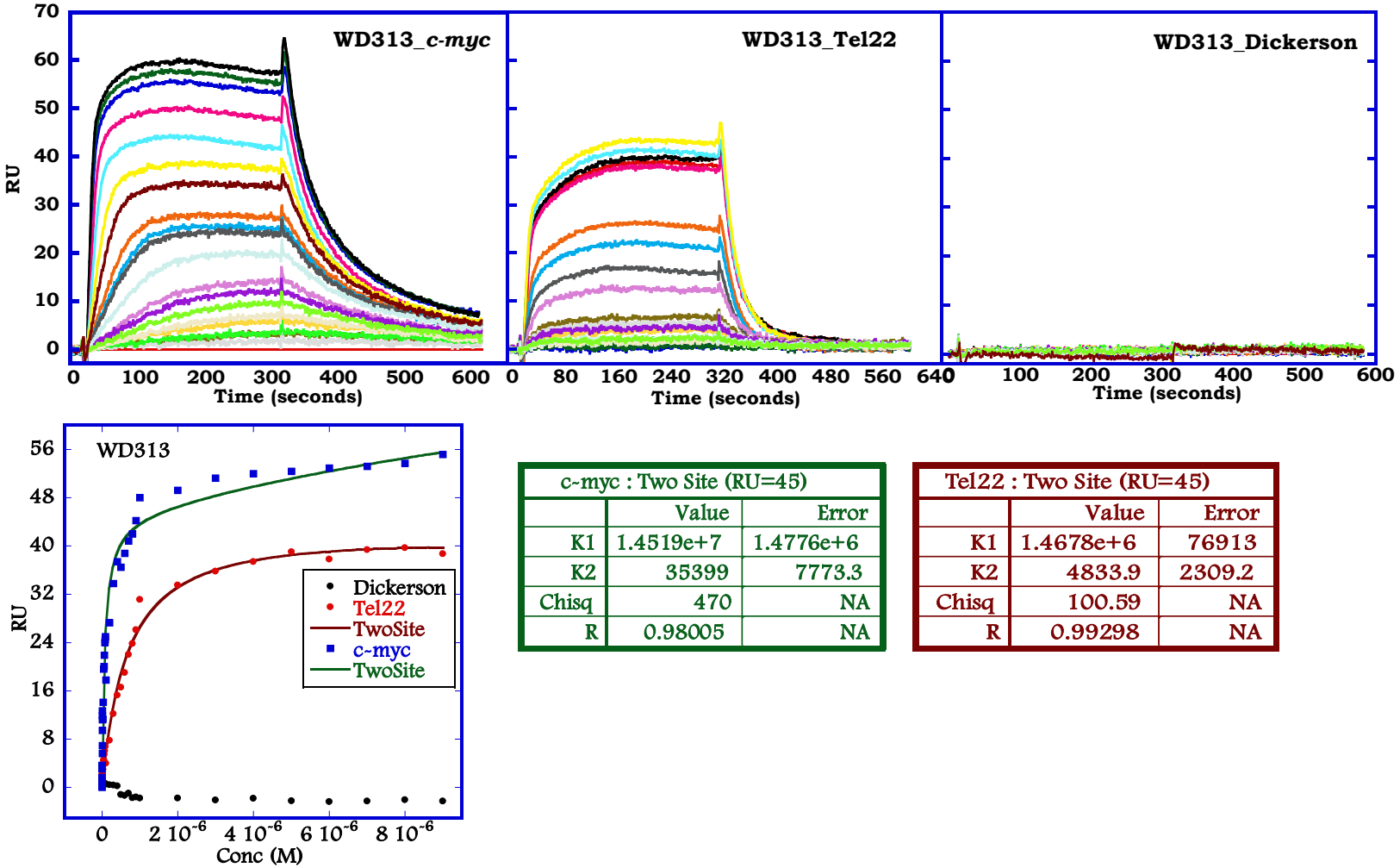
The quadruplex curves range in WD263 concentration from 10 nM for the bottom curve to 1 μM for the top curve. Steady-state binding plots (bottom left) fit to a two-site model (Materials and Methods). The concentration values are for unbound compound concentration in the flow solution. Binding parameters (bottom right) obtained by fitting to a two-site model.



c-myc: Two Site (RU=45)		
	Value	Error
K1	3.7352e+6	2.7054e+5
K2	4.4138e+5	52571
Chisq	27.559	NA
R	0.9971	NA

Figure 5.18: SPR sensorgrams for binding of WD308 to the immobilized G-quadruplexes formed by *c-myc* (top left), Tel22 (top middle) and Dickerson (top right) in HEPES buffer containing 80 mM KCl at 25 °C.

The quadruplex curves range in WD308 concentration from 10 nM for the bottom curve to 1 μM for the top curve. Steady-state binding plots (bottom left) fit to a two-site model (Materials and Methods). The concentration values are for unbound compound concentration in the flow solution. Binding parameters (bottom right) obtained by fitting to a two-site model.



c-myc : Two Site (RU=45)		
	Value	Error
K1	1.4519e+7	1.4776e+6
K2	35399	7773.3
Chisq	470	NA
R	0.98005	NA

Tel22 : Two Site (RU=45)		
	Value	Error
K1	1.4678e+6	76913
K2	4833.9	2309.2
Chisq	100.59	NA
R	0.99298	NA

Figure 5.19: SPR sensorgrams for binding of WD313 to the immobilized G-quadruplexes formed by *c-myc* (top left), Tel22 (top middle) and Dickerson (top right) in HEPES buffer containing 80 mM KCl at 25 °C.

The quadruplex curves range in WD313 concentration from 10 nM for the bottom curve to 10 μM for the top curve. Steady-state binding plots (bottom left) fit to a two-site model (Materials and Methods). The concentration values are for unbound compound concentration in the flow solution. Binding parameters (bottom right) obtained by fitting to a two-site model.

Table 5.1: Equilibrium binding constants determined by SPR for azacyanines with Tel26 and Tel24.

Compound	Tel26	Tel24
Aza3	1.74	1.34
Aza4	1.22	1.00
Aza5	2.11	2.68

[a] Binding Constants ($K \times 10^6\ M^{-1}$) are for the single strong binding site. In addition to the K_1 value shown for the single strong binding site, all compounds have one or two much weaker binding sites with $K = 2 \times 10^5 \pm 1 \times 10^5\ M^{-1}$

[b] Binding constants for the duplex sequences were not determined because of their very low steady-state response even for high compound concentrations.

Table 5.2: Equilibrium binding constants[a] determined by SPR for Naphthalene Diimides with several quadruplex-forming motifs.

Compound	AATT[b]	Tel22	Tel22 k_d (s^{-1})	Tel24-H1	*bcl2*-H2	*c-myc*	*c-kit1*	*c-kit2*
4ND01	---	21.1	0.023	11.9	8.90	Weak	16.8	16.6
4ND02	---	15.3	0.099	1.13	Weak	Weak	34.3	47.3
4ND03	---	17.4	0.019	12.7	Weak	Weak	14.3	18.9
4ND08	---	12.6	0.036	Weak	Weak	Weak	11.0	4.33
4ND09	---	21.3	0.077	0.69	Weak	Weak	22.7	15.4
3ND03	---	15.7	0.046	3.32	Weak	Weak	13.4	4.09
3ND07[c]	---	---	---	---	---	---	---	---
2ND01	---	4.00	---[d]	Weak	Weak	Weak	Weak	1.94

[a] *Binding Constants ($K \times 10^6$ M^{-1}) are for the single strong binding site. In addition to the K_1 value shown for the single strong binding site, all compounds have one or two much weaker binding sites with $K_2 \sim 10^5$ M^{-1}*

[b] *Binding Constants for duplex sequence were too low to be determined using the same experimental procedures due to very low response units.*

[c] *Reliable sensorgrams could not be obtained due to the poor solubility of the compound even at SPR concentrations.*

[d] *Too fast to determine dissociation constant.*

Weak: Response units were too low (< 15 units) even at 1 μM ligand concentration to accurately determine the binding constants and in most cases the binding constants were in the range of low 10^5 M^{-1}

Table 5.3: FRET melting studies[a,b] of selective intramolecular quadruplex-forming sequences[c] and a duplex sequence with naphthalene diimides.

Compound	Duplex[d]	Tel22	*c-kit1*	*c-kit2*
ND01	4.50	33.2	29.8	36.5
4ND02	7.50	28.0	25.3	29.8
4ND03	4.25	35.2	33.8	36.3
4ND08	2.00	34.5	31.5	39.3
4ND09	0.25	27.8	0.00	4.25
3ND03	11.5	26.2	24.5	28.5
3ND07	-0.50	10.5	0.75	4.75
2ND01	2.75	5.25	2.50	7.75

[a] *Performed by members from Dr. Stephen Neidle's group at University of London.*

[b] *ΔTm values (complex – free DNA, in °C) are estimated as the mid-points of the melting transition and the values listed are the mean of two determinations.*

[c, d] *Ligand free single-strand concentration in this study is 0.5 μM.*

Table 5.4: Equilibrium binding constants[a] determined by SPR for Diaryl Ureas with several quadruplex-forming motifs.

Compound	AATT[b]	Tel22	Tel24-H1	bcl2-H2	*c-myc*	*c-kit1*	*c-kit2*
WD263	---	0.40	NB	NB	1.90	1.22	3.87
WD308	---	1.10	Weak	Weak	3.73	2.36	7.61
WD313	---	1.47	Weak	Weak	14.5	2.10	9.66
WD419[c]	---	---	---	---	---	---	---
WD422[c]	---	---	---	---	---	---	---
WD423[c]	---	---	---	---	---	---	---
WD442[c]	---	---	---	---	---	---	---

[a] Binding constants (K x 10^6 M^{-1}) are for the single strong binding site. In addition to the K_1 value shown for the single strong binding site, all compounds have one or two much weaker binding sites with K_2 ~ 10^5 M^{-1} or lower in some cases.

[b] Binding Constants for duplex sequence were too low to be determined using the same experimental procedures due to very low response units.

[c] Reliable sensorgrams could not be obtained due to the poor solubility of the compound.

Weak: Response units were too low (< 15 units) even at 1 μM ligand concentration to accurately determine the binding constants and in most cases the binding constants were less than 10^5

NB: No binding was detected due to a very low response (< 5 units)

Table 5.5: FRET melting studies[a,b] of selective intramolecular quadruplex-forming sequences[c] and a duplex sequence with diaryl ureas.

Compound	Duplex[d]	Quadruplex[e]	*c-kit1*	*c-kit2*
WD263	0.00	12.0	10.1	17.1
WD308	0.40	13.5	10.2	18.7
WD313	0.00	7.90	3.60	11.3

[a] Performed by members from Dr. Stephen Neidle's group at University of London.

[b] ΔTm values (complex – free DNA, in °C) are estimated as the mid-points of the melting transition and the values listed are the mean of two determinations.

[c, d] Ligand free single-strand concentration in this study is 0.5 μM.

[e] F21T, d[$G_3(T_2AG_3)_3$]

5.5 REFERENCES

[1] a)S. M. Haider, G. N. Parkinson, S. Neidle, *Journal of Molecular Biology* **2003**, *326*, 117; b)M. J. B. Moore, *journal of medicinal chemistry* **2006**, *49*, 582; c)P. J. Perry, S. M. Gowan, A. P. Reszka, P. Polucci, T. C. Jenkins, L. R. Kelland, S. Neidle, *journal of medicinal chemistry* **1998**, *41*, 3253; d)P. J. Perry, M. A. Read, R. T. Davies, S. M. Gowan, A. P. Reszka, A. A. Wood, L. R. Kelland, S. Neidle, *journal of medicinal chemistry* **1999**, *42*, 2679; e)P. J. Perry, A. P. Reszka, A. A. Wood, M. A. Read, S. M. Gowan, H. S. Dosanjh, J. O. Trent, T. C. Jenkins, L. R. Kelland, S. Neidle, *journal of medicinal chemistry* **1998**, *41*, 4873; f)L. Rossetti, M. Franceschin, S. Schirripa, A. Bianco, G. Ortaggi, M. Savino, *Bioorganic & medicinal chemistry letters* **2005**, *15*, 413; g)C. Sissi, L. Lucatello, A. Paul Krapcho, D. J. Maloney, M. B. Boxer, M. V. Camarasa, G. Pezzoni, E. Menta, M. Palumbo, *Bioorganic & Medicinal Chemistry* **2007**, *15*, 555; h)F. Koeppel, J.-F. Riou, A. Laoui, P. Mailliet, P. B. Arimondo, D. Labit, O. Petitgenet, C. Helene, J.-L. Mergny, *Nucl. Acids Res.* **2001**, *29*, 1087; i)W. M. David, J. Brodbelt, S. M. Kerwin, P. W. Thomas, *Anal Chem* **2002**, *74*, 2029; j)M. Franceschin, E. Pascucci, A. Alvino, D. D'Ambrosio, A. Bianco, G. Ortaggi, M. Savino, *Bioorganic & medicinal chemistry letters* **2007**, *17*, 2515.

[2] a)C. L. Grand, H. Han, R. M. Muñoz, S. Weitman, D. D. Von Hoff, L. H. Hurley, D. J. Bearss, *Molecular Cancer Therapeutics* **2002**, *1*, 565; b)L. D. Han H, Rangan A, Hurley LH., *Journal of the American Chemical Society* **2001**, *123*, 11; c)M.-Y. Kim, M. Gleason-Guzman, E. Izbicka, D. Nishioka, L. H. Hurley, *Cancer Research* **2003**, *63*, 3247; d)C. Martins, M. Gunaratnam, J. Stuart, V. Makwana, O. Greciano, A. P. Reszka, L. R. Kelland, S. Neidle, *Bioorganic & medicinal chemistry letters* **2007**, *17*, 2293; e)M. J. B. Moore, C. M. Schultes, J. Cuesta, F. Cuenca, M. Gunaratnam, F. A. Tanious, W. D. Wilson, S. Neidle, *journal of medicinal chemistry* **2005**, *49*, 582.

[3] a)H. Han, C.L. Cliff, and L.H. Hurley,, *Biochemistry* **1999**, *38*, 66981; b)W. R. Izbicka E, Raymond E, Davidson KK, Lawrence RA, Sun D, Windle BE, Hurley LH, Von Hoff DD., *Cancer Research* **1999**, *59*, 6; c)K. Shin-ya, K. Wierzba, K.-i. Matsuo, T. Ohtani, Y. Yamada, K. Furihata, Y. Hayakawa, H. Seto, *Journal of the American Chemical Society* **2001**, *123*, 1262; d)G. C. Siddiqui-Jain A, Bearss DJ, Hurley LH., *Proceedings of the National Academy of Sciences of the United States of America* **2002**, *99*, 6.

[4] F. X. Han, R. T. Wheelhouse, L. H. Hurley, *Journal of the American Chemical Society* **1999**, *121*, 3561.

[5] a)A. Arora, S. Maiti, *The Journal of Physical Chemistry B* **2008**, *112*, 8151; b)M. W. Freyer, R. Buscaglia, K. Kaplan, D. Cashman, L. H. Hurley, E. A. Lewis, *Biophysical Journal* **2007**, *92*, 2007; c)N. Nagesh, R. Buscaglia, J. M. Dettler, E. A. Lewis, *Biophysical Journal* **2010**, *98*, 2628.

[6] R. T. Wheelhouse, D. Sun, H. Han, F. X. Han, L. H. Hurley, *Journal of the American Chemical Society* **1998**, *120*, 3261.

[7] J. Seenisamy, E. M. Rezler, T. J. Powell, D. Tye, V. Gokhale, C. S. Joshi, A. Siddiqui-Jain, L. H. Hurley, *Journal of the American Chemical Society* **2004**, *126*, 8702.

[8] a)E. Nyarko, T. Hara, D. J. Grab, A. Habib, Y. Kim, O. Nikolskaia, T. Fukuma, M. Tabata, *Chemico-Biological Interactions* **2004**, *148*, 19; b)S. Y. Rha, E. Izbicka, R. Lawrence, K. Davidson, D. Sun, M. P. Moyer, G. D. Roodman, L. Hurley, D. Von Hoff, *Clinical Cancer Research* **2000**, *6*, 987; c)Y.-A. Shieh, S.-J. Yang, M.-F. Wei, M.-J. Shieh, *ACS Nano* **2010**, *4*, 1433.

[9] a)D. G. Hilmey, M. Abe, M. I. Nelen, C. E. Stilts, G. A. Baker, S. N. Baker, F. V. Bright, S. R. Davies, S. O. Gollnick, A. R. Oseroff, S. L. Gibson, R. Hilf, M. R. Detty, *journal of medicinal chemistry* **2001**, *45*, 449; b)C. E. Stilts, M. I. Nelen, D. G. Hilmey, S. R. Davies, S. O. Gollnick, A. R. Oseroff, S. L. Gibson, R. Hilf, M. R. Detty, *journal of medicinal chemistry* **2000**, *43*, 2403.

[10] a)D. L. Banville, L. G. Marzilli, W. D. Wilson, *Biochemical and Biophysical Research Communications* **1983**, *113*, 148; b)J. A. Strickland, L. G. Marzilli, K. M. Gay, W. D. Wilson, *Biochemistry* **1988**, *27*, 8870; c)T. Uno, K. Aoki, T. Shikimi, Y. Hiranuma, Y. Tomisugi, Y. Ishikawa, *Biochemistry* **2002**, *41*, 13059.

[11] a)B. Armitage, *Chemical Reviews* **1998**, *98*, 1171; b)D. T. Croke, L. Perrouault, M. A. Sari, J. P. Battioni, D. Mansuy, C. Helene, T. Le Doan, *Journal of Photochemistry and Photobiology B: Biology* **1993**, *18*, 41; c)J. M. Nussbaum, M. E. A. Newport, M. Mackie, N. B. Leontis, *Photochemistry and Photobiology* **1994**, *59*, 515.

[12] a)N. H. Campbell, *Journal of the American Chemical Society* **2008**, *130*, 6722; b)C. Hounsou, L. Guittat, D. Monchaud, M. Jourdan, N. Saettel, J.-L. Mergny, M.-P. Teulade-Fichou, *ChemMedChem* **2007**, *2*, 655.

[13] a)M.-K. Cheng, C. Modi, J. C. Cookson, I. Hutchinson, R. A. Heald, A. J. McCarroll, S. Missailidis, F. Tanious, W. D. Wilson, J.-L. Mergny, C. A. Laughton, M. F. G. Stevens, *journal of medicinal chemistry* **2008**, *51*, 963; b)J. Debray, W. Zeghida, M. Jourdan, D. Monchaud, M.-L. Dheu-Andries, P. Dumy, M.-P. Teulade-Fichou, M. Demeunynck, *Organic & biomolecular chemistry* **2009**, *7*; c)Y.-T. Fu, B. R. Keppler, J. Soares, M. B. Jarstfer, *Bioorganic & Medicinal Chemistry* **2009**, *17*, 2030; d)M. Laronze-Cochard, Y.-M. Kim, B. Brassart, J.-F. Riou, J.-Y. Laronze, J. Sapi, *European Journal of Medicinal Chemistry* **2009**, *44*, 3880.

[14] a)E. Gavathiotis, R. A. Heald, M. F. G. Stevens, M. S. Searle, *Journal of Molecular Biology* **2003**, *334*, 25; b)S. M. Gowan, R. Heald, M. F. G. Stevens, L. R. Kelland, *Molecular Pharmacology* **2001**, *60*, 981; c)R. A. Heald, C. Modi, J. C. Cookson, I. Hutchinson, C. A. Laughton, S. M. Gowan, L. R. Kelland, M. F. G. Stevens, *journal of medicinal chemistry* **2001**, *45*, 590; d)C. Leonetti, M. Scarsella, G. Riggio, A. Rizzo, E. Salvati, M. D'Incalci, L. Staszewsky, R. Frapolli, M. F. Stevens, A. Stoppacciaro, M. Mottolese, B. Antoniani, E. Gilson, G. Zupi, A. Biroccio, *Clinical Cancer Research* **2008**, *14*, 7284.

[15] a)C. Leonetti, S. Amodei, C. D'Angelo, A. Rizzo, B. Benassi, A. Antonelli, R. Elli, M. F. G. Stevens, M. D'Incalci, G. Zupi, A. Biroccio, *Molecular Pharmacology* **2004**, *66*, 1138; b)E. Salvati, C. Leonetti, A. Rizzo, M. Scarsella, M. Mottolese, R. Galati, I. Sperduti, M. F. G. Stevens, M. D'Incalci, M. Blasco, G. Chiorino, S. Bauwens, B. Horard, E. Gilson, A. Stoppacciaro, G. Zupi, A. Biroccio, *The Journal of Clinical Investigation* **2007**, *117*, 3236.

[16] P. Phatak, J. C. Cookson, F. Dai, V. Smith, R. B. Gartenhaus, M. F. G. Stevens, A. M. Burger, *Br J Cancer* **2007**, *96*, 1223.

[17] J. C. Cookson, F. Dai, V. Smith, R. A. Heald, C. A. Laughton, M. F. G. Stevens, A. M. Burger, *Molecular Pharmacology* **2005**, *68*, 1551.

[18] M. Gunaratnam, O. Greciano, C. Martins, A. P. Reszka, C. M. Schultes, H. Morjani, J.-F. Riou, S. Neidle, *Biochemical Pharmacology* **2007**, *74*, 679.

[19] a)S. Taetz, T. E. Mürdter, J. Zapp, S. Boettcher, C. Baldes, E. Kleideiter, K. Piotrowska, U. F. Schaefer, U. Klotz, C. M. Lehr, *International Journal of Pharmaceutics* **2008**, *357*, 6; b)R. J. Ward, C. Autexier, *Molecular Pharmacology* **2005**, *68*, 779.

[20] S. Taetz, C. Baldes, T. Mürdter, E. Kleideiter, K. Piotrowska, U. Bock, E. Haltner-Ukomadu, J. Mueller, H. Huwer, U. Schaefer, U. Klotz, C. M. Lehr, *Pharmaceutical Research* **2006**, *23*, 1031.

[21] A. M. Burger, F. Dai, C. M. Schultes, A. P. Reszka, M. J. Moore, J. A. Double, S. Neidle, *Cancer Research* **2005**, *65*, 1489.

[22] M.-Y. Kim, H. Vankayalapati, K. Shin-ya, K. Wierzba, L. H. Hurley, *Journal of the American Chemical Society* **2002**, *124*, 2098.

[23] T. Shalaby, A. O. von Bueren, M.-L. Hürlimann, G. Fiaschetti, D. Castelletti, T. Masayuki, K. Nagasawa, A. Arcaro, I. Jelesarov, K. Shin-ya, M. Grotzer, *Molecular Cancer Therapeutics* **2010**, *9*, 167.

[24] N. Wheate, L. Webster, C. Brodie, J. Grant Collins, *Anti-Cancer Drug Design* **2000**, *15*, 313.

[25] a)R. Kieltyka, P. Englebienne, N. Moitessier, H. Sleiman, **2010**, pp. 223; b)S. Redon, S. Bombard, M.-A. Elizondo-Riojas, J.-C. Chottard, *Nucl. Acids Res.* **2003**, *31*, 1605.

[26] H. Bertrand, S. Bombard, D. Monchaud, M.-P. Teulade-Fichou, *NUCLEIC ACIDS SYMP SER (OXF)* **2008**, *52*, 163.

[27] I. Ourliac-Garnier, M.-A. Elizondo-Riojas, S. Redon, N. P. Farrell, S. Bombard, *Biochemistry* **2005**, *44*, 10620.

[28] a)A. L. Harris, X. Yang, A. Hegmans, L. Povirk, J. J. Ryan, L. Kelland, N. P. Farrell, *Inorganic Chemistry* **2005**, *44*, 9598; b)M. J. Hannon, *Pure and Applied Chemistry* **2007**, *79*, 19.

[29] a)S. S. Jain, M. Polak, N. V. Hud, *Nucl. Acids Res.* **2003**, *31*, 4608; b)Ö. Persil, C. T. Santai, S. S. Jain, N. V. Hud, *Journal of the American Chemical Society* **2004**, *126*, 8644; c)M. Polak, N. V. Hud, *Nucl. Acids Res.* **2002**, *30*, 983.

[30] Ö. P. Çetinkol, A. E. Engelhart, R. K. Nanjunda, W. D. Wilson, N. V. Hud, *ChemBioChem* **2008**, *9*, 1889.

[31] a)F. A. Tanious, S. F. Yen, W. D. Wilson, *Biochemistry* **1991**, *30*, 1813; b)S. F. Yen, E. J. Gabbay, W. D. Wilson, *Biochemistry* **1982**, *21*, 2070.

[32] a)M. F. Brana, A. Ramos, *Current Medicinal Chemistry -Anti-Cancer Agents* **2001**, *1*, 237; b)E. Van Quaquebeke, T. Mahieu, P. Dumont, J. Dewelle, F. Ribaucour, G. Simon, S. Sauvage, J.-F. Gaussin, J. Tuti, M. El Yazidi, F. Van Vynckt, T. Mijatovic, F. Lefranc, F. Darro, R. Kiss, *journal of medicinal chemistry* **2007**, *50*, 4122.

[33] C. Thalacker, C. Röger, F. Würthner, *The Journal of Organic Chemistry* **2006**, *71*, 8098.

[34] a)V. P. Cokic, B. B. Beleslin-Cokic, M. Tomic, S. S. Stojilkovic, C. T. Noguchi, A. N. Schechter, *Blood* **2006**, *108*, 184; b)E. Fibach, P. Prasanna, G. Rodgers, D. Samid, *Blood* **1993**, *82*, 2203.

[35] a)A. Garofalo, L. Goossens, A. Lemoine, A. Farce, Y. Arlot, P. Depreux, *Journal of Enzyme Inhibition and Medicinal Chemistry* **2010**, *25*, 158; b)U. R. Khire, D. Bankston, J. Barbosa, D. R. Brittelli, Y. Caringal, R. Carlson, J. Dumas, T. Gane, S. L. Heald, B. Hibner, J. S. Johnson, M. E. Katz, N. Kennure, J. Kingery-Wood, W. Lee, X.-G. Liu, T. B. Lowinger, I. McAlexander, M.-K. Monahan, R. Natero, J. Renick, B. Riedl, H. Rong, R. N. Sibley, R. A. Smith, D. Wolanin, *Bioorganic & medicinal chemistry letters* **2004**, *14*, 783; c)J. Lyons, S. Wilhelm, B. Hibner, G. Bollag, *Endocr Relat Cancer* **2001**, *8*, 219.

[36] D. Monchaud, *Organic & biomolecular chemistry* **2008**, *6*, 627.

[37] a)W. C. Drewe, S. Neidle, *Chemical Communications* **2008**, 5295; b)H. C. Kolb, M. G. Finn, K. B. Sharpless, *Angewandte Chemie International Edition* **2001**, *40*, 2004.

[38] a)W. C. Drewe, R. Nanjunda, M. Gunaratnam, M. Beltran, G. N. Parkinson, A. P. Reszka, W. D. Wilson, S. Neidle, *journal of medicinal chemistry* **2008**, *51*, 7751; b)F. Cuenca, O. Greciano, M. Gunaratnam, S. Haider, D. Munnur, R. Nanjunda, W. D. Wilson, S. Neidle, *Bioorganic & medicinal chemistry letters* **2008**, *18*, 1668.

[39] a)E. R. Lacy, N. M. Le, C. A. Price, M. Lee, W. D. Wilson, *Journal of the American Chemical Society* **2002**, *124*, 2153; b)S. Mazur, F. A. Tanious, D. Ding, A. Kumar, D. W. Boykin, I. J. Simpson, S. Neidle, W. D. Wilson, *Journal of Molecular Biology* **2000**, *300*, 321; c)B. Nguyen, C. Tardy, C. Bailly, P. Colson, C. Houssier, A. Kumar, D. W. Boykin, W. D. Wilson, *Biopolymers* **2002**, *63*, 281; d)L. Wang, C. Bailly, A. Kumar, D. Ding, M. Bajic, D. W. Boykin, W. D. Wilson, *Proceedings of the National Academy of Sciences of the United States of America* **2000**, *97*, 12.

[40] T. M. Davis, W. D. Wilson, *Analytical Biochemistry* **2000**, *284*, 348.

[41] a)A. Ambrus, D. Chen, J. Dai, T. Bialis, R. A. Jones, D. Yang, *Nucl. Acids Res.* **2006**, *34*, 2723; b)K. N. Luu, A. T. Phan, V. Kuryavyi, L. Lacroix, D. J. Patel, *Journal of the American Chemical Society* **2006**, *128*, 9963.

[42] a)J. Li, J. J. Correia, L. Wang, J. O. Trent, J. B. Chaires, *Nucl. Acids Res.* **2005**, *33*, 4649; b)G. N. Parkinson, M. P. H. Lee, S. Neidle, *Nature* **2002**, *417*, 876.

[43] a)H. Fernando, A. P. Reszka, J. Huppert, S. Ladame, S. Rankin, A. R. Venkitaraman, S. Neidle, S. Balasubramanian, *Biochemistry* **2006**, *45*, 7854; b)A. T. Phan, V. Kuryavyi, S. Burge, S. Neidle, D. J. Patel, *Journal of the American Chemical Society* **2007**, *129*, 4386; c)S. Rankin, A. P. Reszka, J. Huppert, M. Zloh, G. N. Parkinson, A. K. Todd, S. Ladame, S. Balasubramanian, S. Neidle, *Journal of the American Chemical Society* **2005**, *127*, 10584.

[44] a)J. Dai, D. Chen, R. A. Jones, L. H. Hurley, D. Yang, *Nucl. Acids Res.* **2006**, *34*, 5133; b)J. Dai, T. S. Dexheimer, D. Chen, M. Carver, A. Ambrus, R. A. Jones, D. Yang, *Journal of the American Chemical Society* **2006**, *128*, 1096.

[45] a)A. Ambrus, D. Chen, J. Dai, R. A. Jones, D. Yang, *Biochemistry* **2005**, *44*, 2048; b)A. T. Phan, Y. S. Modi, D. J. Patel, *Journal of the American Chemical Society* **2004**, *126*, 8710.

6 NMR STRUCTURAL CHARACTERIZATION OF A NOVEL HETEROCYCLIC DIAMIDINE THAT RECOGNIZES A UNIQUE GC/AT MOTIF AS AN ANTIPARALLEL STACKED DIMER

6.1 INTRODUCTION

Small molecules have proven to be extremely important to researchers to explore biochemical processes at the cellular and molecular levels, both *in vitro* and *in vivo*. Such molecules have also been proven to be valuable for treating diseases, and most medicines marketed today are from this class. Over the past two decades, research on small molecules acting on nucleic acids has not only led to therapeutically useful drugs but also provided an invaluable source of structural and biological information on nucleic acids. DNA is an intracellular target of many anticancer drugs, and the interaction between small molecules and DNA has been shown to cause DNA damage in cancer cells by inhibiting their growth and ultimately resulting in apoptosis. DNA binding anticancer drugs that can interact with DNA can be categorized into compounds that bind non-covalently and those that bind covalently, and it is interesting that within both categories some compounds can further bind to DNA sequences selectively. The effectiveness of these compounds in potential anti-parasitic and anti-cancer therapy is closely related to the mechanisms by which they induce cell death. In some cases the interaction between drugs and DNA results in DNA structural distortion or damage that inhibits replication or transcription[1]. Another possible mechanism is the selective binding to abundant stretches of –AT– or –GC– repeat sequences that are essential for normal chromosome metabolism or gene transcription [2]. In particular, some compounds could regulate nucleic acid function by targeting specific sequences such as particular genes or their RNA products, or promoters for gene transcription or translation [3], and can be used to target specific DNA sequences acting as important regulation factor in gene expression. Characterization of the sequence specificity of drug–DNA interactions is essential for understanding the drug mechanism of action. By elucidating the molecular determinants of

that specificity, it will be possible to develop principles for the design of new drugs with enhanced anticancer potency.

A number of natural and synthetic compounds which bind to the minor groove of DNA are becoming an important class of tools with which to study ligand–DNA interactions, and increasingly are also of interest as potential anticancer drugs[4]. Polyamides containing N-methylpyrrole (Py) and N-methylimidazole (Im) amino acids bind to predetermined sequences in the minor groove of DNA with affinities and specificities comparable to naturally occurring DNA binding proteins[5]. Distamycin-A and netropsin, generally referred to as "shape-selective" binders, preferentially bind to the narrower minor groove of AT-rich sequences, and the absence of the protruding 2-amino group of guanine based on the shape-fit for maximizing stabilizing van der Waal's contacts[6]. Hydrogen bonding between the groove floor base pairs and the linking amides, and electrostatic stabilizing interactions with the protonated amines are primary contributors to the overall stability of the drug/DNA complex[7]. Hydrogen bonds, van der Waal contacts, and electrostatic interactions between the oligopeptidic pyrrole-carbamoyl frame that ends with an amidine moiety and the DNA double helix result in secure DNA minor groove binding with AT-rich preference. TATA box binding protein (TBP), a component of the basal transcriptional machinery for RNA polymerase II transcribed genes, binds to AT-rich target DNA sequence to promote the transcription, which could be inhibited by some compounds that bind preferentially to AT sequences such as distamycin-A, netropsin, Hoechst 33258, and Py-Im polyamides[8].

Aromatic diamidines, such as DAPI, berenil, pentamidine, and related compounds, were discovered some years ago to have excellent activity against an array of infectious diseases from *Pneumocystis carinii* pneumonia to trypanosomiasis[9]. All of the biologically active aromatic diamidines studied to date have been found to bind strongly to AT-rich sequences in

the minor groove of DNA, and it is clear that there is a correlation between DNA binding for many compounds and the antiparasitic biological activities[10]. These compounds have been discussed in great detail in Chapter 4.

In recent years compounds preferentially binding to GC sequences have piqued the interest of scientific community to further understand the DNA recognition rules. Polyamines, such as spermine, bind in the major groove of GC-rich regions and in the minor groove of AT-rich regions[11]. It has been demonstrated that the binding of polyamines causes macroscopic curvature and bending of DNA and, therefore, the action of polyamines could bring sequence-specific transcription factors and basal transcription factors in close proximity through DNA structure modifications[12]. Trabectedin, an anticancer drug, preferentially binds to GC-rich sequences in DNA, and further binds more strongly to CGG sites compared to GCG, showing the importance of sequence selectivity[13]. The guanine amino group protrudes into the minor groove and generally obstructs the access of drugs to the floor of the groove. The fact that the amino group constitutes a critical negative recognition element for binding of many small molecules in the minor groove of DNA has now been unambiguously demonstrated using DNA molecules in which that group has been either deleted from guanines and/or added to adenines[14]. Given both the strategic position of the amino group in the minor groove and its potential to participate in hydrogen bonding, it was proposed that the introduction of an H-bond acceptor heteroatom in the pyrrole rings of netropsin might permit the drug to bind to GC sequences[15]. Dervan's and Lown's group have extensively exploited this concept and synthesized a series of lexitropsins that has increased selectivity for GC sequences[1, 16]. Among the numerous lexitropsins synthesized so far, imidazole lexitropsins display the most pronounced capacity for binding to GC-containing sequences[17]. Thiazole lexitropsins exhibit a mixed GC recognition influenced by the position of the sulfur on the ring[18]. Despite the elegant design strategy for lexitropsin/DNA complexes, the biological activity of this class of

minor groove binders has not been much improved[4, 17]. So far, the lexitropsin approach has not led to clinically useful drugs, although certain monomeric and dimeric lexitropsins exhibit interesting antiviral or anticancer activities *in vitro* and sometimes *in vivo* [19].

Heterocyclic diamidines designed and developed by Boykin and coworkers have gained significant momentum as a new class of DNA minor groove targeting agents with the possibility of developing into potential antiparastic and antitrypanosomal agents [4, 20]. DB75, a phenyl-furan-phenyl diamidine, is a therapeutically useful heterocyclic dication that has been crystallized as an AATT complex[21]. A neutral, orally available prodrug of DB75 reached phase III clinical trials against human African trypanosomiasis [22]. DB293, a phenyl-furan-benzimidazole dication of DB75, was surprisingly shown to bind to specific sequences of DNA containing mixed GC/AT motif[23]. DNA recognition by DB293 was shown to be strongly cooperative with the formation of a stacked dimer in the minor groove [23-24]. This was a surprising finding considering dications such as netropsin have not been discovered to form stacked dimer with any DNA sequences. The importance of dications as potential antiparasitic agents further necessitates the development of new dimer motifs.

DB832, a phenyl-furan-furan diamidine, was shown to recognize G-rich sequences that can fold into a variety of quadruplex conformations (Chapters 2-4). DB832 was shown to bind in the multiple grooves of human telomeric quadruplex conformation as stacked species; however, a strong, initial end-stacking binding mode was also observed. The unique recognition mode of DB832 coupled with a high binding stoichiometry deemed it difficult to obtain detailed structural information by NMR methodology. The rational design of new derivatives of DB832 to better recognize quadruplex grooves would require detailed structural models for the drug/DNA complex. Structural information would provide the molecular basis for specific recognition and can suggest possible modifications to extend and modify the

sequence-specificity. The goal of understanding duplex groove-binding was probed with a new duplex targeting compound, DB1878. DB1878, a phenyl-furan-indole analog of DB293, was shown to recognize the mixed GC/AT motif as a stacked dimer, but with a significantly higher affinity as seen from surface plasmon resonance studies. Here, we have conducted a detailed 2D NMR analysis of the complex of DB1878 with the mixed GC/AT containing sequence, oligo2-1. The structure reported here is completely different from the traditional DNA recognition exhibited by lexitropsins, and represents an entirely new motif for DNA minor groove recognition. Detailed structural information of the dimer complex of DB1878 with oligo2-1 could potentially aid in designing improved small molecules that can better interact with the guanine-rich grooves of quadruplex structures.

6.2 MATERIALS AND METHODS

6.2.1 Surface Plasmon Resonance (SPR) Studies

Surface plasmon resonance experiments were conducted with a BIAcore T100 instrument as previously described [25]. Biosensor experiments were conducted in filtered, degassed HEPES buffer (10 mM HEPES, 100 mM NaCl, 3 mM EDTA, 0.005 v/v of 10% P20 BIACORE surfactant, pH 7.3) at 25 °C. Flow cell 1 was left blank as a reference, while flow cells 2-4 were immobilized with DNA on a streptavidin-derivatized gold chip (SA chip from BIAcore) by manual injection of DNA stock solutions (flow rate of 1 μL/min) until the desired value of DNA response was obtained (350-400 RU). Typically, a series of different ligand concentrations (1 nM to 10 μM from 20 mM H_2O stock) were injected onto the chip (flow rate of 50 μL/min, 5-10 min) until a constant steady-state response was obtained followed by a dissociation period (buffer, 10 min). After every cycle, the chip surface was regenerated (20 s injection of 10 mM glycine solution, pH 2.0) followed by multiple buffer injections. The RU values in steady-state regions at each concentration were averaged over a 60-sec time zone

and converted to r (moles of compound bound per mole of DNA hairpin) as previously described [25]. The binding affinities were determined by the best fitting plot of r versus free compound concentration with a single site binding model ($K_2 = 0$) or a two-site ($K_1 \neq K_2$) binding model.

One site: $r = (K_1C_{free})/(1 + K_1C_{free})$

Two site: $r = (K_1C_{free} + 2K_1K_2C_{free}^2)/(1 + K_1C_{free} + K_1K_2C_{free}^2)$

where K_1 and K_2 are the macroscopic equilibrium binding constants; C_{free} is the concentration of the compound in equilibrium with the complex and is fixed by the concentration in the flow solution.

6.2.2 Nuclear Magnetic Resonance (NMR) Studies

All NMR experiments were done at 5 °C (unless stated otherwise) on a Bruker Avance™600 using a 5 mm QXI triple resonance z-gradient probehead (Bruker). Oligo2-1, d[CTATGACTCTCGTCATAC], hairpin was purchased from Integrated DNA Technologies with IE purification and mass spectrometry characterization. The DNA was further dialyzed in over long periods of time to further remove any remaining impurities. The final hairpin concentration for all experiments was 1.35 mM. All NMR samples were prepared in 10 mM $NaHPO_4$ buffer containing 20 mM NaCl and 0.1 mM EDTA. The final pH of the NMR samples was adjusted to 7.0 and the sample was annealed prior to collecting data. A 30 mM stock solution of DB1878 was prepared in 99.99% D_2O to perform titrations. Water suppression for samples in 90% H_2O/10% D_2O was achieved with a 1-1 spin-echo or WATERGATE pulse sequences. For experiments in 99.99% D_2O, residual D_2O resonance was suppressed with a low power presaturation pulse. All one-dimensional NMR experiments were collected over 8k or 16k data points with spectral width of 11 ppm (D_2O samples) or 24 ppm (H_2O samples) with a relaxation delay time of 2.5 seconds. Phase-sensitive Double Quantum Filtered

COrrelated SpectroscopY (DQF-COSY)[26], Total Correlated SpectroscopY (TOCSY)[27], NOE SpectroscopY (NOESY)[28], and $^{31}P-^{1}H$ HETCOR experiments were performed by collecting either 1024 or 2048 points in f2, and between 512 and 800 points in f1. NOESY spectra at mixing times ranging from 50 ms to 300 ms were collected with a spectral width of 10 ppm (for D_2O samples) or 24 ppm (for H_2O samples) in each dimension, with a delay time of 3 seconds. Heteronuclear single quantum correlation (HSQC) spectra using echo-antiecho mode for the ^{13}C inversion were recorded using z-gradients (2K × 256, 160 scans, 4 seconds delay) were performed on a Bruker Avance™ 600 equipped with a cryoprobe at Vanderbilt University. Quadrature detection in f1 was achieved using states time-proportional phase increment (states-TPPI). Two-dimensional data were zero-filled to 4k x 4k points prior to Fourier transformation, optimized with a SSB function (2, 2) in both dimensions, and treated with automatic baseline correction. All data were processed on a DELL precision T7400 Workstation with RHEL5 kernel using Bruker TOPSPIN™ processing software and peak assignments were performed using SPARKY (UCSF) and MNova (Mestrelab Research, Spain) software.

6.3 RESULTS AND DISCUSSION

6.3.1 Proton Assignments of Oligo2-1

The NMR investigations of the drug-DNA complexes were undertaken by first characterizing the free oligonucleotide. This was followed by examination of the complex formed upon reaction with DB1878. The hairpin duplex and drug-DNA complex were fully assigned using a combination of *through space* and *through bond* interactions from NOESY, TOCSY and DQF-COSY spectra, using the method formulated by Wuthrich and coworkers[29]. It has previously been shown that H1' protons of B-form DNA typically resonate between 6.5 and 5.0 ppm, while H6/H8 protons resonate in the range 8.5-6.5 ppm[29]. Consequently, the

starting point for the assignment of the oligo2-1 sequence was the NOE interactions in this region of the NOESY spectrum i.e. the H6/H8-H1' assignment pathway (Figure 6.2). The H6/H8 of the 5'-terminal residue will interact through space with the H1' of its own deoxyribose only, whereas subsequent H6/H8 protons will interact with both their own H1' protons and with the sugar protons of the preceding residue. The four bases in the loop of oligo2-1, however, are independent of each other and only have cross peaks to their own H1' resonance. Using this assignment strategy, the backbone connectivity of oligo2-1 has been completely assigned (Figure 6.2-A) at 5 °C. A temperature of 5 °C was employed for all NMR experiments to reduce the effects of chemical exchange and also to reduce the amount of minor species that was constantly observed at room temperature. The connectivity followed the standard patterns as for B-DNA duplexes and complete assignment of base and sugar protons for the oligo2-1 sequence was obtained.

The exchangeable proton signals for the imino protons were also assigned from NOESY spectra of oligo2-1 obtained in 90% H_2O sample at 5 °C (Figure 6.3). Imino protons in a Watson-Crick type base-pairing generally resonate in the region between 11.5-14.5 ppm[29]. The imino-imino connectivities for the six base-pairs in the stem region of oligo2-1 can be followed in Figure 6.3, where the corresponding cross-peaks are well resolved. A clear pathway can be established starting from the imino of T2 until the imino of G12, where the stem of oligo2-1 ends. The imino proton signal of the last base-pair in the stem, G18 is absent in the pathway due to the fraying of helix at the ends and also due to high susceptibility for solvent exchange. However, the imino proton of G18 can be seen in 1D spectrum as a broad peak (~ 12.9 ppm, Figure 6.3-B), and undergoes further broadening as higher temperatures, indicating duplex opening at higher temperatures. The imino proton resonances of all the base-pairs involved in Watson-Crick base-pairing disappear around 45 °C suggesting the significant exchange with the solvent around this temperature.

The sequential assignments of other non-exchangeable sugar protons (H2'/H2", H3' and H4') were also performed for oligo2-1 at 5 °C using the backbone connectivity method, and all the assignments are listed in Table 6.1. The sugar H5'/H5" protons could not be unambiguously assigned due to significant overlap.

6.3.2 Proton Assignments of DB1878

The assignments of non-exchangeable protons of DB1878 were performed using the *through bond* connectivity COSY experiment and the expanded aromatic region is shown in Figure 6.4. The B1 proton of DB1878 does not have any COSY-type connections (three-bond coupling) with any other protons of the molecule and therefore appears as a singlet in the spectrum. The same holds true for indole proton (In) of the molecule and appears as singlet upfield of all the protons. Three bond connectivities for F1/F2 and B2/B3 pairs can be clearly established. The phenyl ring of DB1878 rotates rapidly in solution, and, therefore, the two magnetically equivalent ring protons P1/P3 and P2/P4 are not distinguishable. All the proton assignments of DB1878 are further verified with 1D-^{13}C experiment (spectrum not shown). The complete assignment of free DB1878 is listed in Table 6.2.

6.3.3 DB1878 Binds as a Highly Cooperative Dimer

To quantitatively evaluate the binding of DB1878, SPR studies (by Dr. Yang Liu from Dr. Wilson's laboratory) were conducted with oligo2-1, and the sensorgrams are shown in Figure 6.5-A. Binding of DB1878 was highly cooperative and saturates at two molecules per oligo2-1 hairpin. Fitting the steady-state response values to a two-site model gave an initial weak binding with K_1=9.4 x 10^5 M^{-1}, followed by a strong binding with K_2=5.1 x 10^8 M^{-1}. The binding of DB1878 to oligo2-1 was strongly cooperative with a cooperativity factor of greater 500 (K_2/K_1). The concave shape of the Scatchard plot further confirms the strong

positive cooperativity for ligand binding. Cooperative binding of stacked dimers by polyamides and an heterocyclic diamidine, DB293, with GC-containing sequences has been observed and suggests that the relatively wider grooves of GC base-pairs facilitates the formation of stacked dimers.

To further confirm the cooperativity, NMR studies of DB1878 with oligo2-1 were conducted. Figure 6.5-B shows the 2D COSY spectra of the methyl and aromatic proton interactions of the six thymine residues of oligo2-1 at different ratios of DB1878 at 5 °C. In the absence of any ligand (ratio 0:1), the scalar coupling peaks of methyl-to-aromatic protons are clearly observed corresponding to the six thymine residues. At the intermediate ratio (ratio 1:1), two distinct set of cross-peaks are observed suggesting the presence of two different species and the ligand bound in an intermediate exchange. One set of these peaks correspond to the uncomplexed DNA, whereas, the other set is the 2:1 species based on the 2:1 ratio. Further titration of DB1878 (ratio 2:1) resulted in complete disappearance of the peaks from the unbound DNA species, suggesting the complete saturation of the DNA. The clear saturation of the DNA at this ratio suggests that two distinct molecules are bound per mole of DNA. The new set of peaks in the final ratio corresponds to the second set of peaks observed at the intermediate ratio. Therefore, the two sets of peaks in the intermediate ratio correspond to the unbound and completely bound DNA suggesting a strong positive cooperativity as seen from SPR studies.

6.3.4 Characterization of the 2:1 Complex

DB1878 was titrated to oligo2-1 at 5 °C to form the 2:1 complex for further NMR analysis. The aromatic region of the titration spectra of DB1878 to oligo2-1 is shown in Figure 6.6-A at different ratios, along with the spectra of free DB1878 (25 °C). As seen from the

spectra of DB1878, distinct number of protons is seen for the ligand maybe suggesting that the ligand does not form stacked species in solution. Closer inspection of the titration spectra reveals interesting information about the binding of DB1878. The intermediate ratio (ratio 1) spectrum is a mixture of signals for unbound and the completely bound oligo2-1. The unbound to the completely bound oligo2-1 transition can be monitored by some well defined peaks in the spectra. The peak labeled with a "*" belongs to the A17H8 of the unbound oligo2-1. At the intermediate ratio, this peak broadens to some extent and decreases in the intensity. Simultaneously, a new peak is seen in the spectra at the intermediate ratio, labeled with "#". This peak also belongs to the same proton, A17H8, but from the bound oligo2-1, and is not seen in the unbound species (ratio 0). The intensities of the two peaks at the intermediate ratio suggest that an equal population of the bound and unbound oligo2-1 is present and the ligands are in the slow-exchange regime with the DNA. Upon further titration of the ligand (ratio 2), the unbound peak (*) broadens further and its intensity is significantly reduced. Concomitantly, the bound peak (#) sharpens further and is at maximum intensity at the final ratio. A very similar trend is also observed with the A3H8 protons in the bound (†) and the unbound (‡) DNA. This suggests that the DNA is completely saturated upon 2 mole equivalents of DB1878. Thus, the titration spectra of DB1878 with oligo2-1 clearly illustrate that the 2:1 complex and the free DNA are the only species observed in the 1:1 spectrum. The equal population of the completely bound and unbound DNA at the intermediate ratio also confirms the strong cooperative dimerization of the compound.

The aromatic to aromatic region of a COSY spectrum of the 2:1 complex is shown in Figure 6.6-B. This region consists of cross peaks only from the drug protons bound to the DNA. When compared to the COSY spectra of the free DB1878 (Figure 6.4-A), there are clearly twice the number of cross peaks in the new spectrum. This also clearly suggests that there are two distinct DB1878 molecules bound to the DNA. The proton assignments for the two bound

molecules were made using a combination of COSY and NOESY (not shown) and HSQC spectra (Figure 6.7). Every proton resonance of the two molecules is easily distinguishable due to the slow exchange in the bound form. The phenyl ring protons, which are not well resolved in the free DB1878, are clearly distinguishable for both the bound DB1878 in the 2:1 complex due to slow-exchange. The ^{13}C-^{1}H correlation spectra of oligo2-1 with 2:1 ratio of DB1878 (Figure 6.7) shows a total of 20 decoupled peaks. Since each drug molecule has 10 C-H bonds, the 20 peaks that are observed also suggest that two distinct ligands are bound in the complex. Some of the protons resonate very close to each other and therefore have very similar chemical shift values. As a result, data analysis proved difficult to some extent due to the overlapping of the drug proton peaks (discussed later).

6.3.5 Detailed Assignment of DB1878:Oligo2-1 Complex

Non-exchangeable proton peaks of DB1878 complexed with oligo2-1 were assigned using a combination of homonuclear NOESY and TOCSY spectra at different mixing times. The H5-H6 cross peak of C1 provided the entry point for the backbone assignment in the NOESY spectrum (Figure 6.8). The *through space* connection between the base aromatic protons and the sugar H1' proton was clearly traced from C1 to T4 bases. The connection between the G5-H1' and G5-H8 was interrupted since the G5-H1' proton was significantly perturbed and exhibited an extreme upfield shift (~1.2 ppm), indicating a possible ligand induced perturbation. Similarly, C7-H1' also exhibited a significant upfield shift (~1 ppm), possibly suggesting some ring current effect on the H1' proton. The loop bases of oligo2-1 again did not show any inter-base connectivity and had only cross-peaks to their own H1'. On the 3'-strand of oligo2-1, two interruptions were observed. The H1' protons of C14 and T16 again exhibited significant upfield shifts (~1.1 and ~1.3 ppm respectively), suggesting a site-specific, ligand-mediated perturbation of the H1' protons and possible interaction site of the second DB1878

molecule. The combined H1' chemical shift perturbations suggest that the possible interaction site of the two DB1878 molecules is centered between A3•T16 and C7•G12 base pairs. The H1' protons of A6 and A15 also undergo significant upfield shifts (~0.5 and ~0.7 ppm), since they are embedded within the binding site. Interestingly, the H1' protons of G12 and A3 does not exhibit any shifts, considering their base-paired counterparts exhibit tremendous upfield shifts with their H1'. This suggests that, even though the two DB1878 molecules are bound side-by-side in the minor groove, the two ligands are offset from each other and do not completely span the entire proposed binding site: between A3•T16 and C7•G12 base pairs. The offset of the ligands seems rational assuming the positively charged diamidine moieties of the two ligands would avoid each other to reduce charge repulsion, while concurrently maintain a favorable stacking with their overlapped ring systems.

The assignments of the non-exchangeable protons for the 2:1 complex are shown in Table 6.3. To gain better understanding of DB1878 binding site within the minor groove of oligo2-1, the NMR 'footprint' from ligand-induced perturbations to the minor groove proton, H1', is plotted against the sequence position in Figure 6.9-A. Aromatic protons which resides in the major groove were also plotted as a function of sequence position (Figure 6.9-B). The loop regions were not included since they were not likely the binding site for the ligand. The H1' sugar protons reside deeply in the DNA minor groove. Minor groove binding ligands upon complex formation significantly perturbs the local chemical environment of these protons, and therefore, ligand induced changes should be readily observable in the NMR spectra. Several of the sugar H1' protons come into direct contact with the aromatic rings of either of the bound ligands and experience significant upfield ring current perturbations to their chemical shifts. These changes in the chemical shifts are particularly sensitive to the position of the ligand in the minor groove. On the 5'-strand, significant perturbation of the H1' protons are observed starting from the G5 residue and extends until the terminal residue of the stem, C7. On the 3'-

strand, significant perturbations are observed starting from the C14 residue and extends until T16. This strongly suggests the binding site of the two DB1878 molecules is from A3•T16 to C7•G12. Each ligand is able to recognize a minimum of 3 residues on each of the strands based on the chemical shift perturbations. The terminal base-pair does not exhibit any perturbations. In contrast, the aromatic protons, which are located in the DNA major groove, experience smaller effects upon complex formation. The small changes observed might be due to structural perturbations in the DNA or changes in the base stacking interactions, however, a similar pattern of perturbations as observed in the minor groove H1' protons are seen with the major groove aromatic protons.

6.3.6 Extreme Ring Current Effects on H4' Protons

The NOESY spectra of the 2:1 complex were used to comprehensively assign all the non-exchangeable protons of the complex, except for some of the sugar protons that resonate around the water chemical shift that could not be unambiguously assigned. The NOESY spectra of the 2:1 complex exhibited two extremely upfield shifted resonances (~0.85 and 1.09 ppm) with strong NOE cross peaks all along the spectra with various sugar protons and protons from the bases (Figure 6.10-A). The protons that have the most upfield chemical shifts in a normal Watson-Crick B-DNA are the –CH3 protons of thymine residue (in the range of 2.5 - 1.5 ppm). However, all the thymine methyl protons of oligo2-1 in the complex were accounted for and, therefore, the two most upfield proton resonances were considered to be from a minor species of DNA or due to some impurity. A ^{31}P-^{1}H spectrum collected to assign some of the ambiguous sugar protons revealed, surprisingly, two extremely upfield shifted resonances corresponding to the chemical shifts observed in the NOESY spectrum (Figure 6.10-B). ^{31}P spectra is routinely collected to obtain vicinal spin-spin 3-bond (H3', H5'/H5") or 4-bond (H4') coupling. After careful analysis of the spectra supplemented by 3-bond spin-spin

interactions from COSY data of the 2:1 complex, the two upfield shifted resonances were finally assigned to the H4' sugar protons of C7 (0.85 ppm) and T16 (1.09 ppm). H4' sugar protons in B-DNA generally resonate between 4 - 4.5 ppm and in extreme cases have exhibited upfield or downfield shifts within 1.0 – 1.5 ppm of the normal range, as in the case of intercalators with large ring systems [30]. However, here we have identified H4' sugar protons exhibiting an impressive 2.5 – 3.0 ppm upfield shifts. Such a dramatic upfield shift is observed only if there is an extremely shielded ring current effect on protons, and therefore, in this case the shielding of those protons has to be from the ring systems of the bound ligand. The two bases, C7 and T14, are located at the edges of the proposed binding site of DB1878 with their H4' lying along the walls of the DNA minor groove. The stacking of the two DB1878 molecules in an antiparallel manner would put the two indole ring systems on the opposite ends of the binding site and directly on top of the 4' sugar protons (Figure 6.10-C). The orientation of the indole rings in the DNA minor groove reasonably explains the extreme upfield shifts exhibited by the H4' protons due to the large ring current effect. The H1' sugar protons of the same two bases also exhibited significant upfield shifts upon complex formation (Section 6.3.5). This further validates the presence of the ring system of the two molecules in a distinct orientation near the H1' and H4' protons of the two bases to have such a dramatic effect.

6.3.7 Major Ligand-DNA and Ligand-Ligand Interactions

The NOESY spectrum of the 2:1 complex contains several intermolecular DB1878:DNA crosspeaks as well as numerous crosspeaks between the two ligands. Figure 6.11 shows the important drug:DNA and drug:drug interactions in the NOESY spectra of the aromatic and H1' region of the 2:1 complex at 5 °C. All of the NOEs can be used to identify contact points between the DNA and two bound ligands. The terminal diamidines of the two molecules require a novel stacking for binding side-by-side to DNA. From the intermolecular NOEs

between the two DB1878 molecules, the M1B1 resonance signal has a strong crosspeak with the M2P3 resonance. This puts the phenyl ring of M2 molecule in close proximity to the indole ring of M1. Similarly, the M2B1 resonance signal has a strong crosspeak with the M1P2 resonance. These crosspeaks indicate that the two DB1878 molecules are stacked with an offset and make it possible for the terminal diamidines of each molecule to be separated when binding side-by-side into the minor groove. There are two crucial crosspeaks between M1B1 to A6H2 and M2B1 to A15H2 protons. A-H2 protons reside deeply in the minor groove of the DNA and are excellent probes for analyzing small molecules that bind in the DNA minor groove. The high intensity of the two crosspeaks between the B1 of each DB1878 molecule and AH2 protons, place the two B1 protons in direct contact with the two adenine residues. Figure 6.11-B shows the NOESY spectra of the base aromatic to sugar proton region. The three crosspeaks between the G5H1' and M1P1, M1P2 and M1F1 protons place those three protons deep into the minor groove. The crosspeak between C14H1' and M2F1 is of significantly lower intensity than the G5H1' and M1F1 protons. Since the G5 and C14 are involved in a base-pair and the intensities of their H1' protons with the two furan protons from different ligands are not the same, the M2 furan protons are pointing outside the minor groove of the DNA. However, the M2F2 proton exhibits a stronger intensity with the C14H1' proton. This might be due to some spin-diffusion effects. Strong NOE crosspeaks are observed between the drug protons facing outside the minor groove and some of the sugar protons that are located along the walls and the exterior of the minor groove. The M2F1 resonance, which was shown to be facing outside the minor groove, shows NOE crosspeaks with H5'/H5" protons of A15. Similarly, M2In shows NOE contacts with H5'/H5" protons of T16 residue, suggesting that the indole proton of M2 is facing outside the minor groove. Several strong NOE crosspeaks are observed between the two NH protons of the two molecules and DNA protons (spectra not shown). Important NOE crosspeaks are observed between the minor groove protons A6H1' to M1NH and A15H1' to M2NH, suggesting that the NH protons of the two molecules are

oriented into the minor groove of the DNA. Major intermolecular NOEs between the two drug molecules and the drug:DNA are listed in Table 6.4 and Table 6.5 respectively. Important NOE contacts between the two drug molecules and the DNA protons are depicted in Figure 6.12.

6.3.8 Importance of the Central G•C Basepair

The Watson-Crick base-pairing scheme of GC base-pair generally results in the bulky amino group of the guanine residue directly point into the minor groove of the DNA. As a result, this bulky amino group acts as a steric hindrance for minor groove targeting small molecules. However, the amino group can also potentially participate in hydrogen bonding if small molecules can be effectively developed to selectively recognize such systems[14a]. There has been very little success so far in developing small molecules that can sequence specifically interact with GC containing sequences, until recently. The Dervan's and Lown's group have developed a broad class of lexitropsin derivatives for selectively targeting both AT and GC base-pairs [1, 6b, 6d, 31]. NMR studies of polyamides have shown ligands sequence specifically interacting with the G-NH2 groups, and upon complex formation, the G-NH2 protons undergo small downfield chemical shift changes, providing indirect evidence for the specific hydrogen bonds to the acceptors on the ligands. We have also recently reported a heterocyclic diamidine, DB293, to sequence-specifically interact with a GC-containing sequence [23-24]. Systematic mutation studies of the binding site residues critically highlighted the requirement of the central GC base-pair for the cooperative dimerization of DB293. DB1878 has been reported here to form an antiparallel stacked dimer with same sequence, with the central furan rings of the two molecules pointing in the opposite directions. NOESY spectra of free oligo2-1 in 90% water sample shows that the G5 amino resonance completely broadened due to the exchange through the rotation about the N-C bond (spectra not shown). The amino protons generally show a broad cross-peak in the range of 7.0 to 8.0 ppm with the

neighboring imino proton of the guanine, while the basepaired amino proton generally downfield shifted than the non-basepaired proton. Figure 6.13 shows the expanded imino-amino correlation NOESY spectra of the 2:1 complex of DB1878 with oligo2-1. A strong NOE crosspeak is observed to the G5 imino proton from the amino region. The crosspeak of the amino proton is significantly downfield shifted by more than 1 ppm (9.16 ppm) when compared to the broad resonance (7-8 ppm) in the free DNA. This extreme downfield shift of the amino proton is due to the consequence of a direct hydrogen-bonding with the furan oxygen that is pointed into the minor groove. The resonance at 9.16 ppm also has a strong NOE crosspeak with the H1' sugar proton of G5 (not shown). Since the H1' and the NH2 protons point into the minor groove of the DNA, the strong crosspeak to the H1' proton from the 9.16 ppm resonance further confirms the resonance belonging to the amino proton of G5.

6.3.9 Docking Studies

Preliminary docking studies with SYBYL software were conducted using the NMR data. The two DB1878 molecules were carefully positioned in the grooves of the oligo2-1 sequence to qualitatively visualize the NMR results and the ligands were energy minimized while holding the DNA fixed. The final structure of the complex is presented using VMD software (Figure 6.14). A complete detailed NMR structural analysis of the complex is in progress.

6.4 CONCLUSION

DNA minor groove binders have biological activities that range from anti-opportunistic infection to anticancer properties. These compounds have also provided a wealth of fundamental information about nucleic acid recognition properties, and they continue to be important models in the study of nucleic acid complexes. Until recently, the lexitropsins have remained the only class known to recognize both strands of the DNA double helix through

stacked dimer formation. A dimer motif that recognizes both the strands allows much greater selectivity in targeting desired sequences of DNA. DB293 was the first unfused dication that was shown to form stacked dimers with GC-containing sequences. Structurally related dications have been found to have a range of therapeutic properties and, therefore, development of the new dimer motif is of importance. DB1878, an indole analog of DB293, is reported here to bind very strongly and cooperatively to specific sequences of DNA that contain GC base pairs. The unique recognition mode of DB1878 offers new possibilities for development of agents for recognition of mixed base pair sequences of DNA and can help in establishing new set of recognition rules for sequence specific DNA recognition.

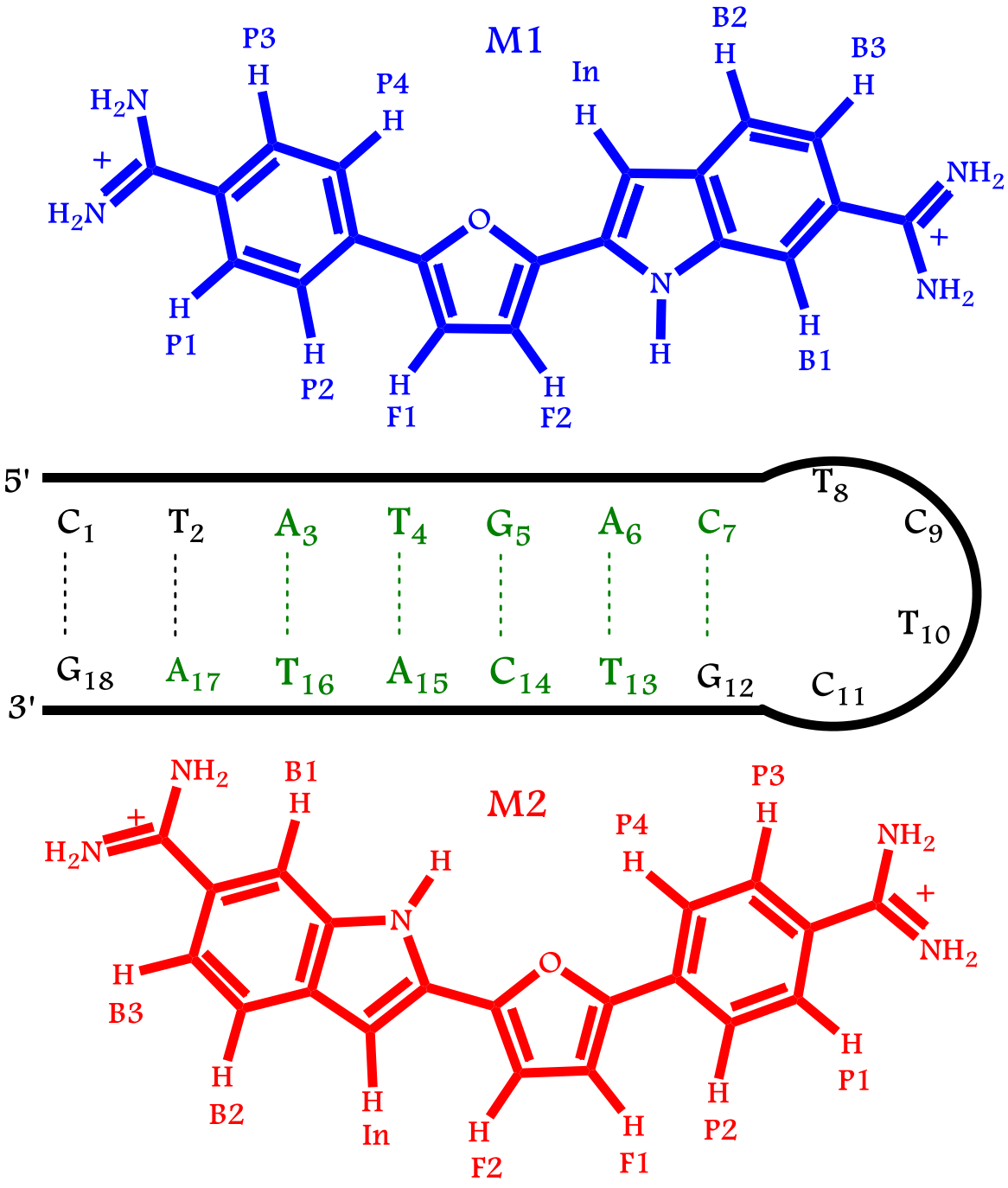


Figure 6.1: Chemical structure of DB1878 with the atom naming and coloring schemes used in this study and the hairpin duplex, Oligo2-1. The possible binding site of DB1878 is highlighted in green.

(A) (B)

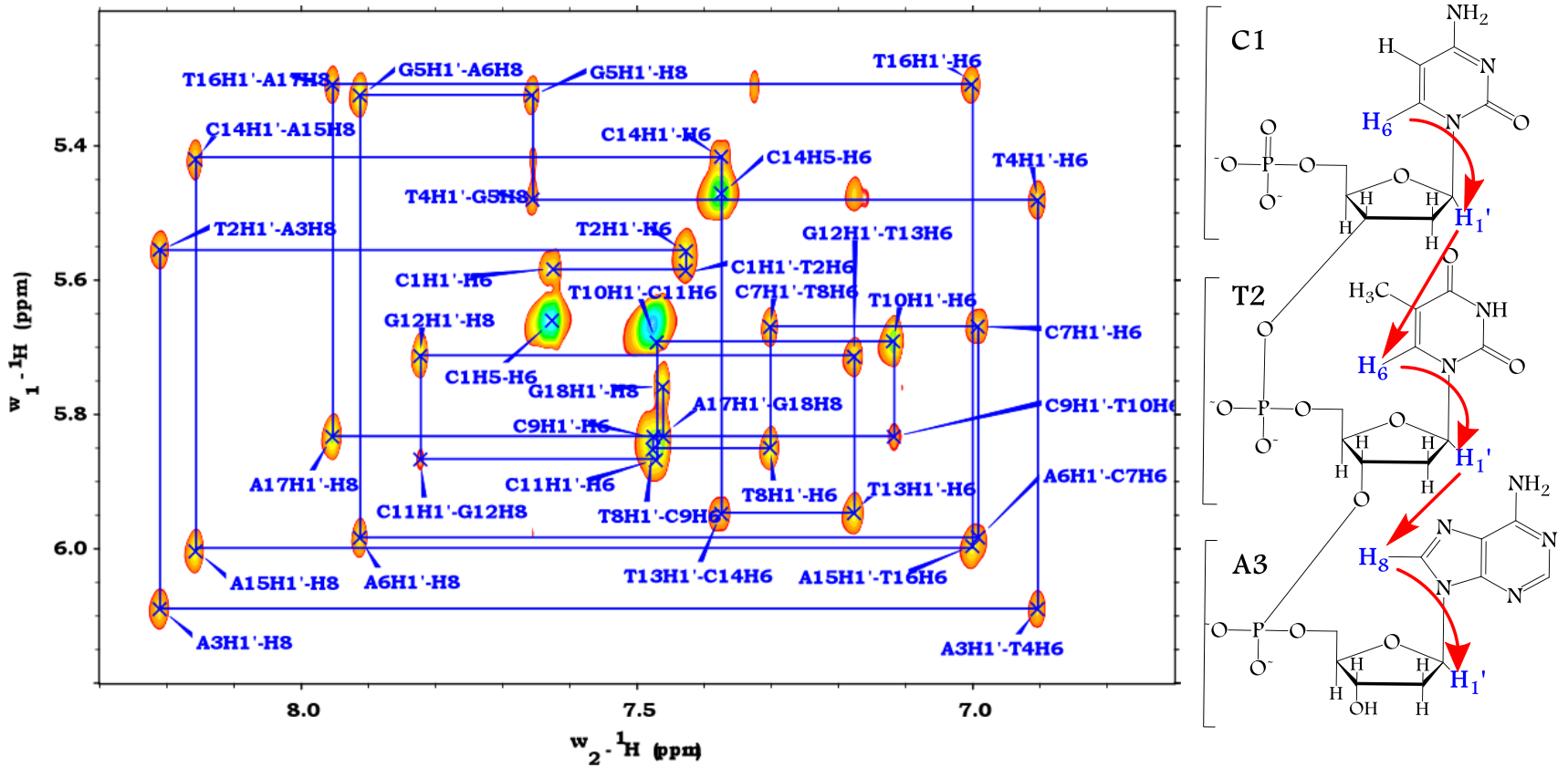


Figure 6.2: (A) Aromatic (H6/H8) to H1' backbone connectivity region of oligo2~1 at 5 °C. (B) Schematic H6/H8~H1' connectivity observed.

Similar connectivity is also observed for H2'/H2" and H3' sugar protons with the aromatic H6/H8 protons.

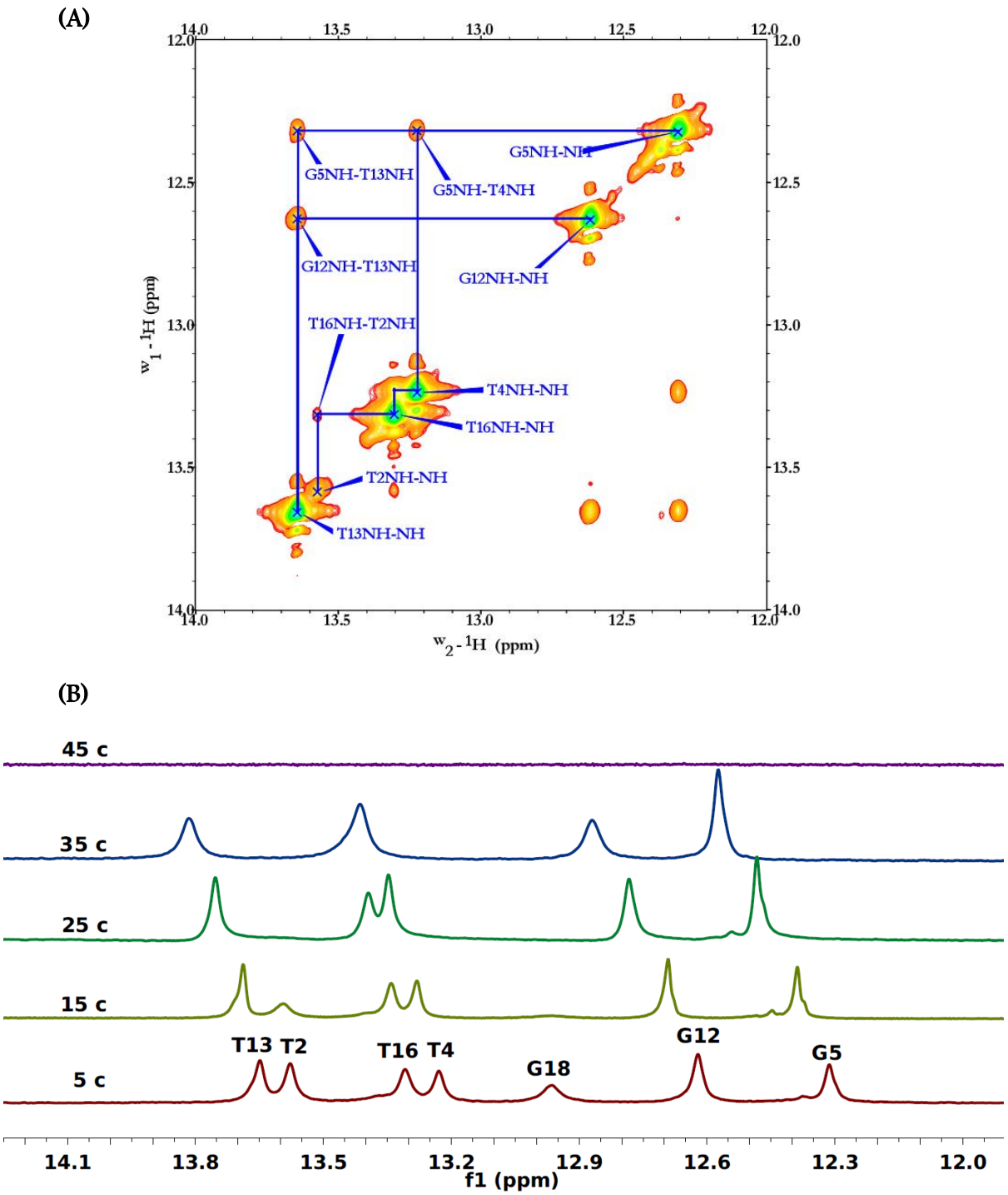


Figure 6.3: Expanded region of the observed exchangeable proton 2D NOESY spectrum of free oligo2-1 at 5 °C (A), and imino proton spectra of free oligo2-1 as a function of temperature (B).

NOESY spectrum conditions: mixing time: 300 ms, water suppression with 1-1 pulse sequence, relaxation delay: 2.5 seconds, temperature: 5 °C and [DNA]=1.5 mM

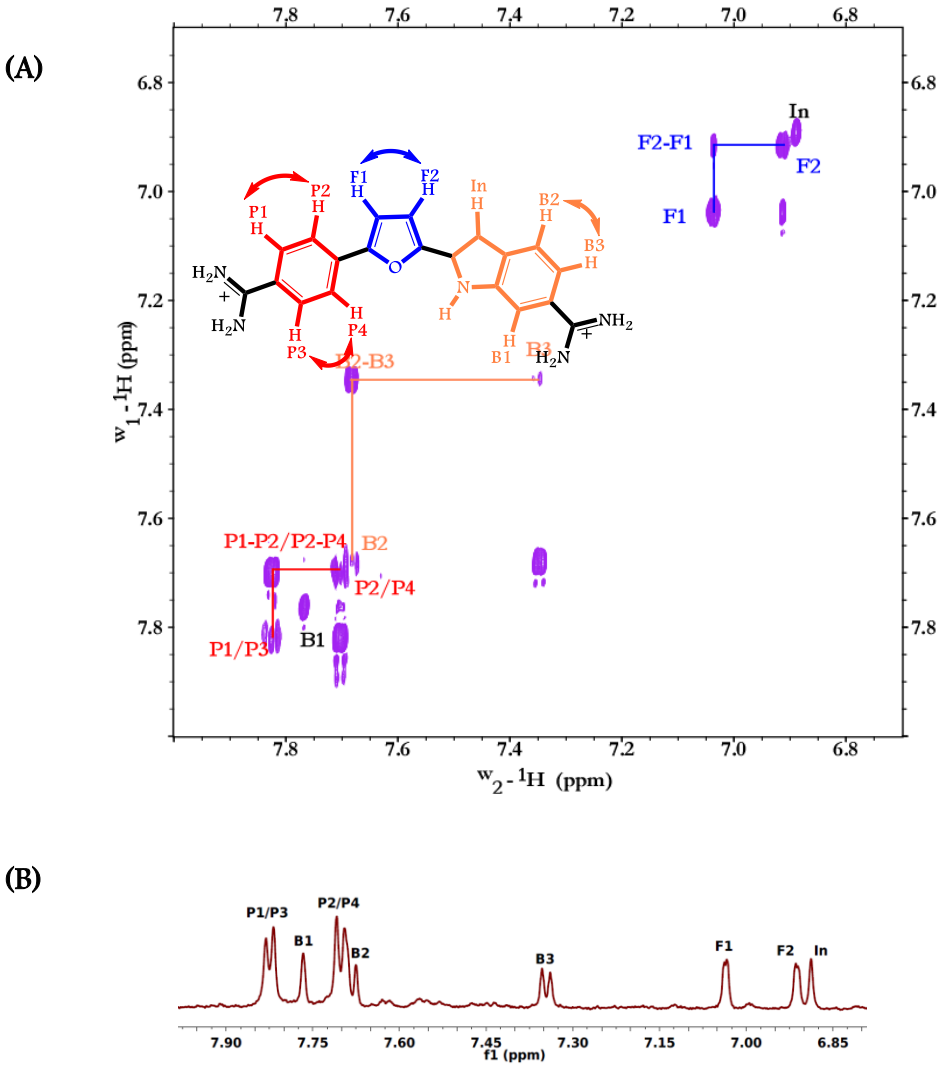


Figure 6.4: The expanded aromatic region of COSY spectrum of DB1878 in D_2O at 25 °C (A), and the corresponding 1D proton spectra (B).

The three bond coupling observed between the protons in the ligand is shown as an inset in the COSY spectrum.

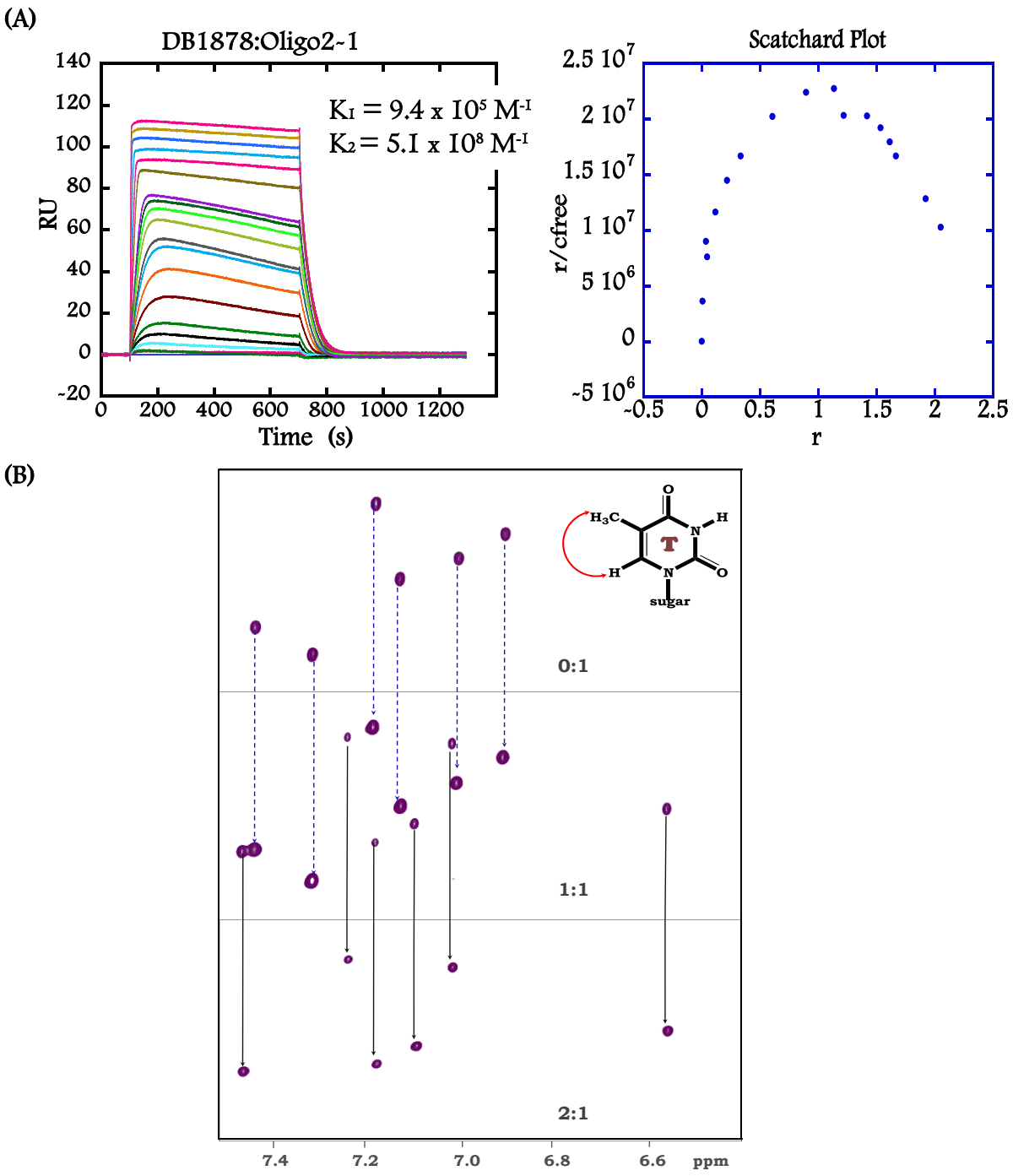


Figure 6.5: SPR sensorgrams of DB1878 binding to oligo2~1 with the corresponding Scatchard plot (A), and the 2D COSY spectra of the T: CH3~H6 region for DB1878~oligo2~1 interaction at 5 °C with listed ratios (B).

From the Scatchard plot and the COSY spectrum it is clear that the binding of DB832 to oligo2~1 is positively cooperative.

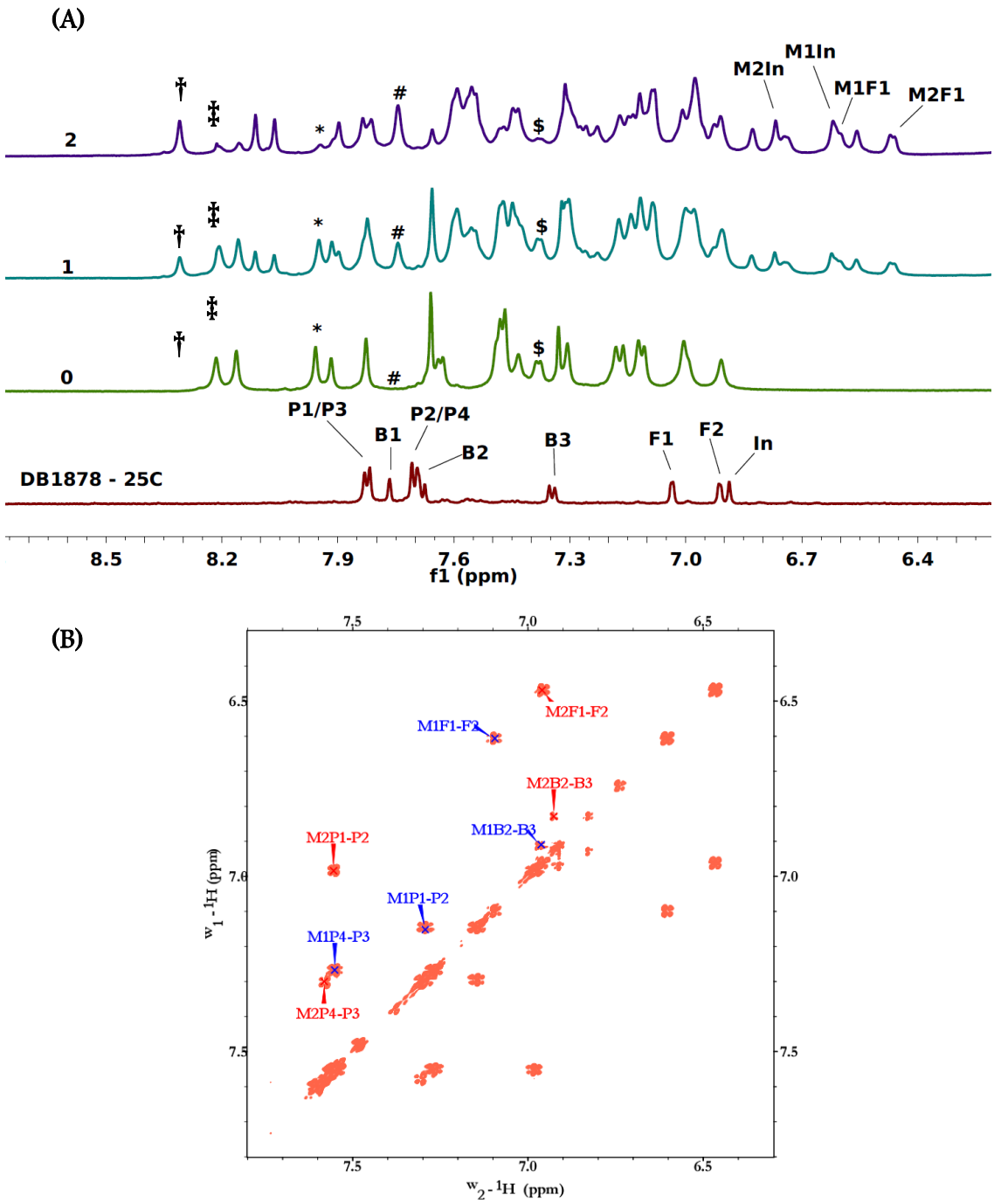


Figure 6.6: Aromatic proton spectra of DB1878 with oligo2-1 at 5 °C (A), and the aromatic region of the COSY spectrum of 2:1 complex at 5°C (B).

*Peaks labeled with † and ‡ indicates the shift of bound and unbound A3H8 respectively, and the peaks labeled with # and * indicates the shift of bound and unbound A17H8 respectively. The peak labeled with $ belong to C14H6 of the free DNA and is significantly reduced at saturation ratios of DB1878. All cross peaks in the aromatic region of the complex correspond to the bound DB1878 molecules.*

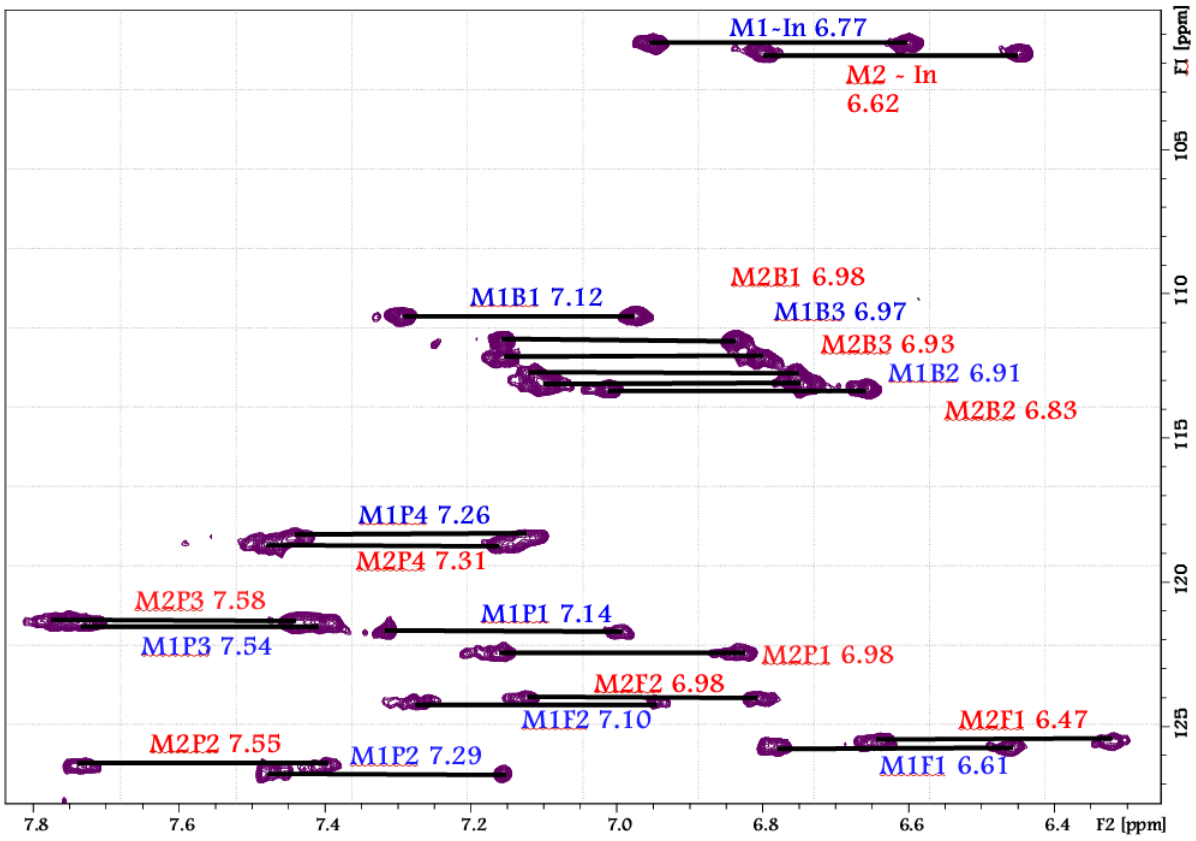


Figure 6.7: Expanded region of the ^{1}H-^{13}C-HSQC spectrum of oligo2-1 with 2:1 ratio of DB1878 at 5 °C without decoupling.

A total of 20 decouple C-H correlation pairs (10 for each bound molecule) are observed accounting for both of the drug molecules bound to the DNA. This experiment was performed at Vanderbilt University.

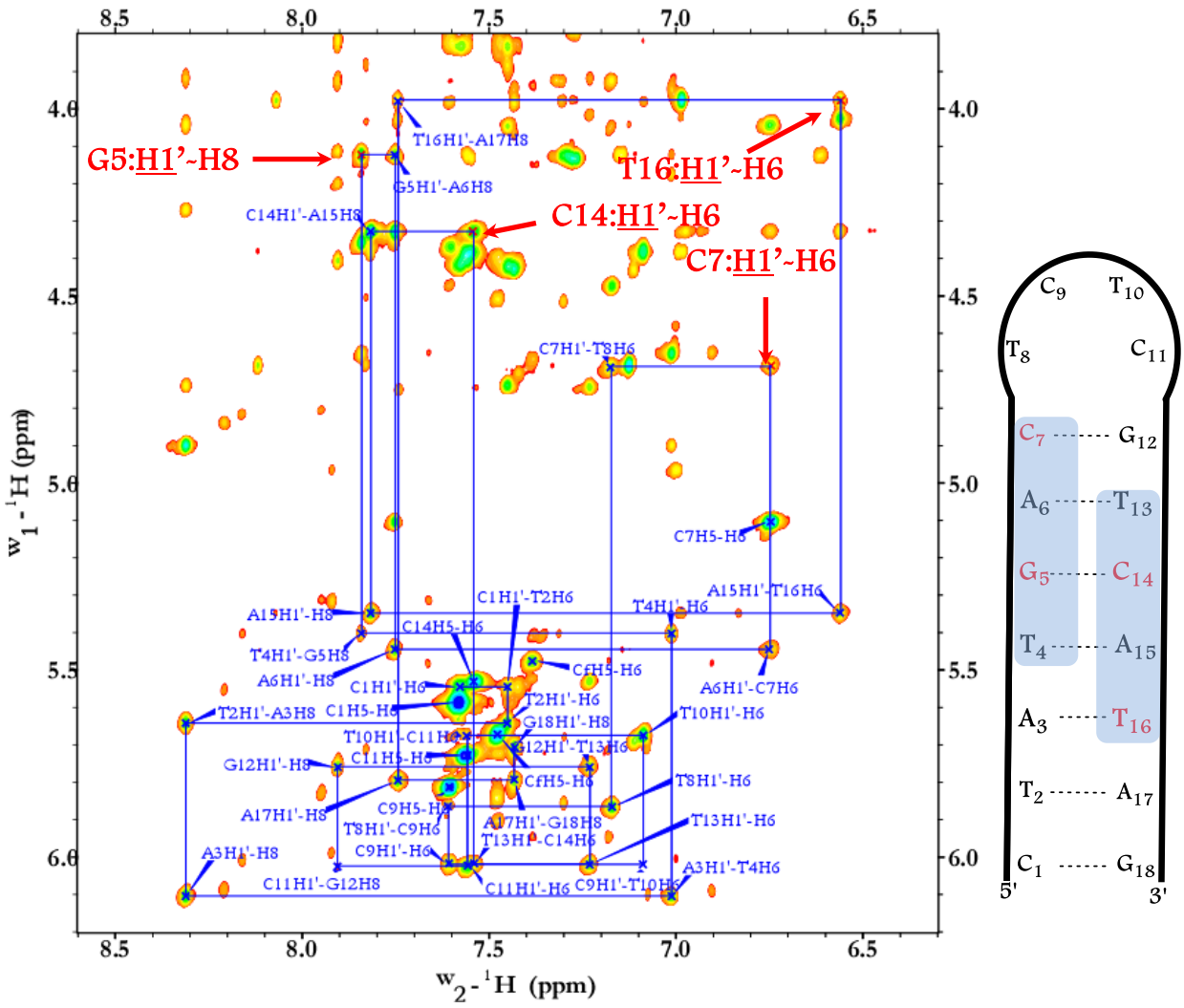
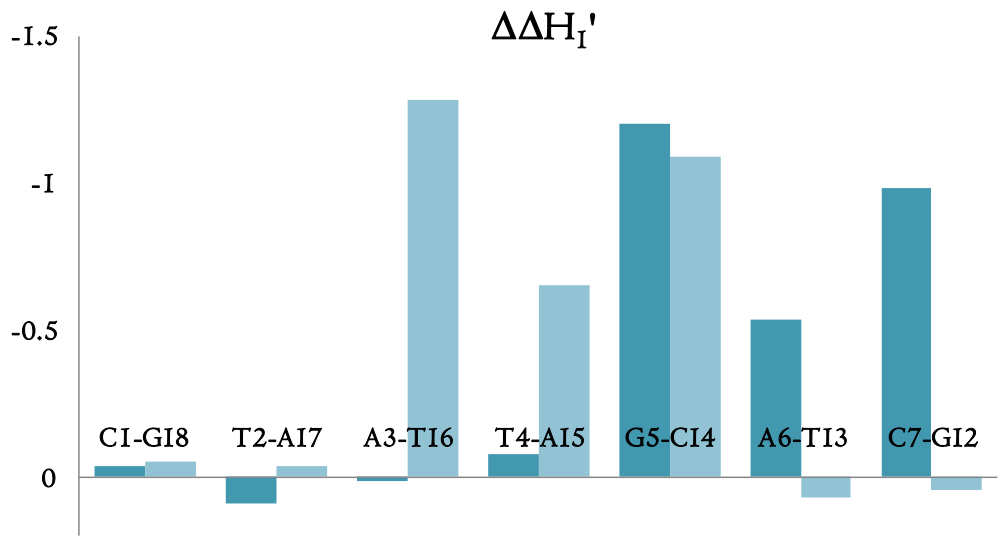


Figure 6.8: The aromatic to H1' backbone connectivity region of 2:1 complex of DB1878 with oligo2-1 at 5 °C.

Some of the sugar H1' protons that exhibit significant upfield shifts are labeled in red. H1' protons of the sugar deeply reside in the minor groove of duplex DNA, and minor groove binding ligands can significantly perturb the chemical environment of these protons as observed. A schematic representation (right) of the possible binding site of two DB1878 molecules in the minor groove of oligo2-1 based on the H1' chemical shift perturbations.

(A)



(B)

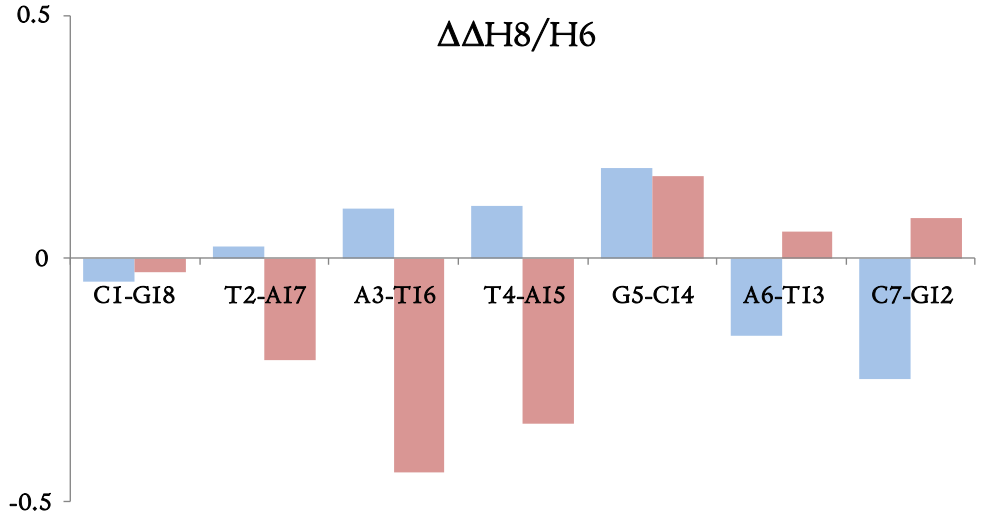


Figure 6.9: Plot of ligand induced changes for H1' (A), and H6/H8 (B) in the 1H-NMR chemical shifts for the 2:1 complex of DB1878 with oligo2-1. The loop residues are not included.

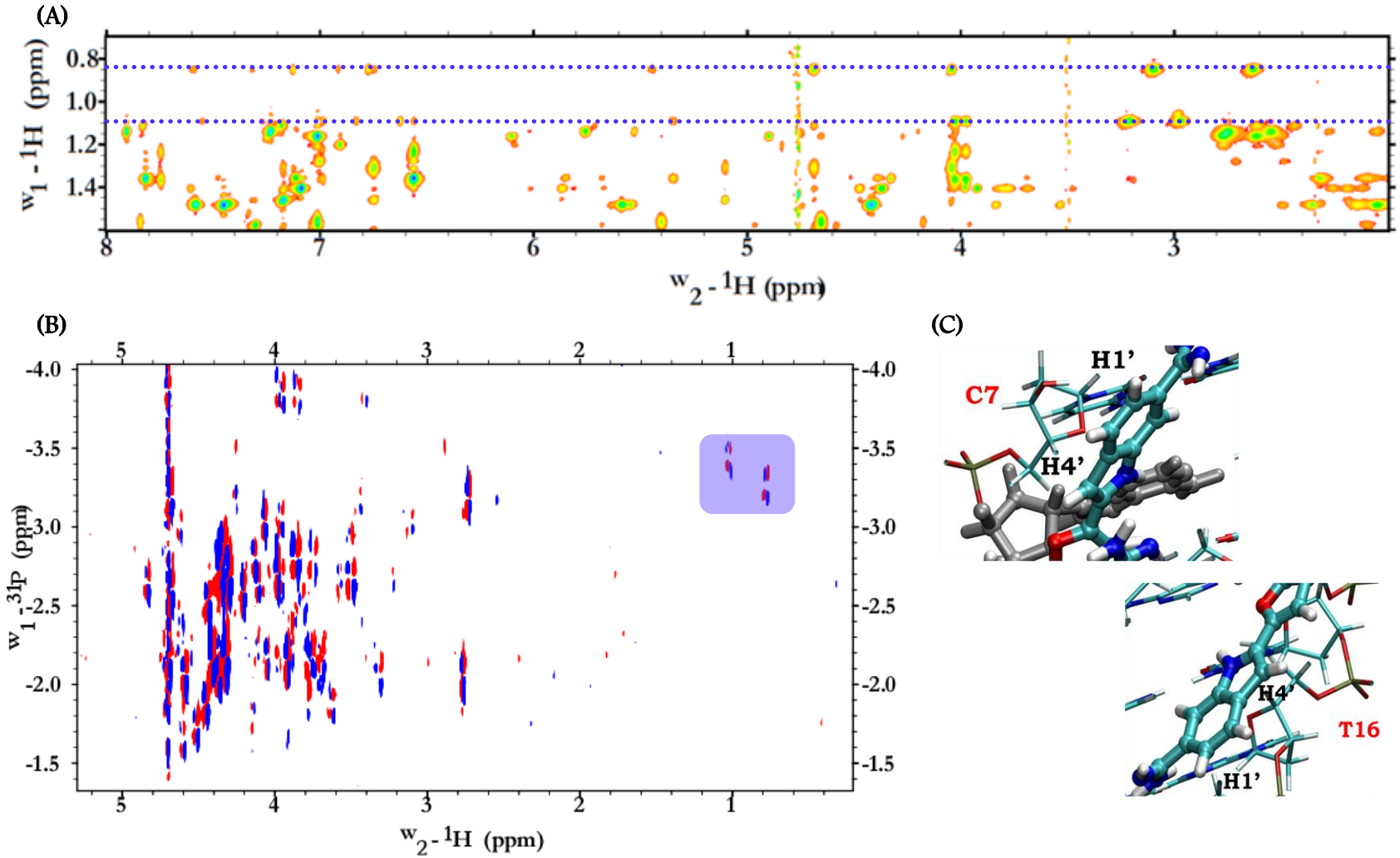


Figure 6.10: Expanded NOESY spectrum of the upfield region (A), and ^{31}P-^{1}H correlation spectrum (B) of 2:1 complex of DB1878 with oligo2-1 at 5 °C.

The blue dotted lines in (A) indicate extreme upfield shifts exhibited by C7H4' (0.85 ppm) and T16H4' (1.09 ppm). The purple box in (B) indicates the same two proton signals in the ^{31}P-^{1}H correlation spectrum. A schematic representation (C) of a possible ring current effect of the two DB1878 molecules on the two H4' protons to exhibit such extreme upfield shifts.

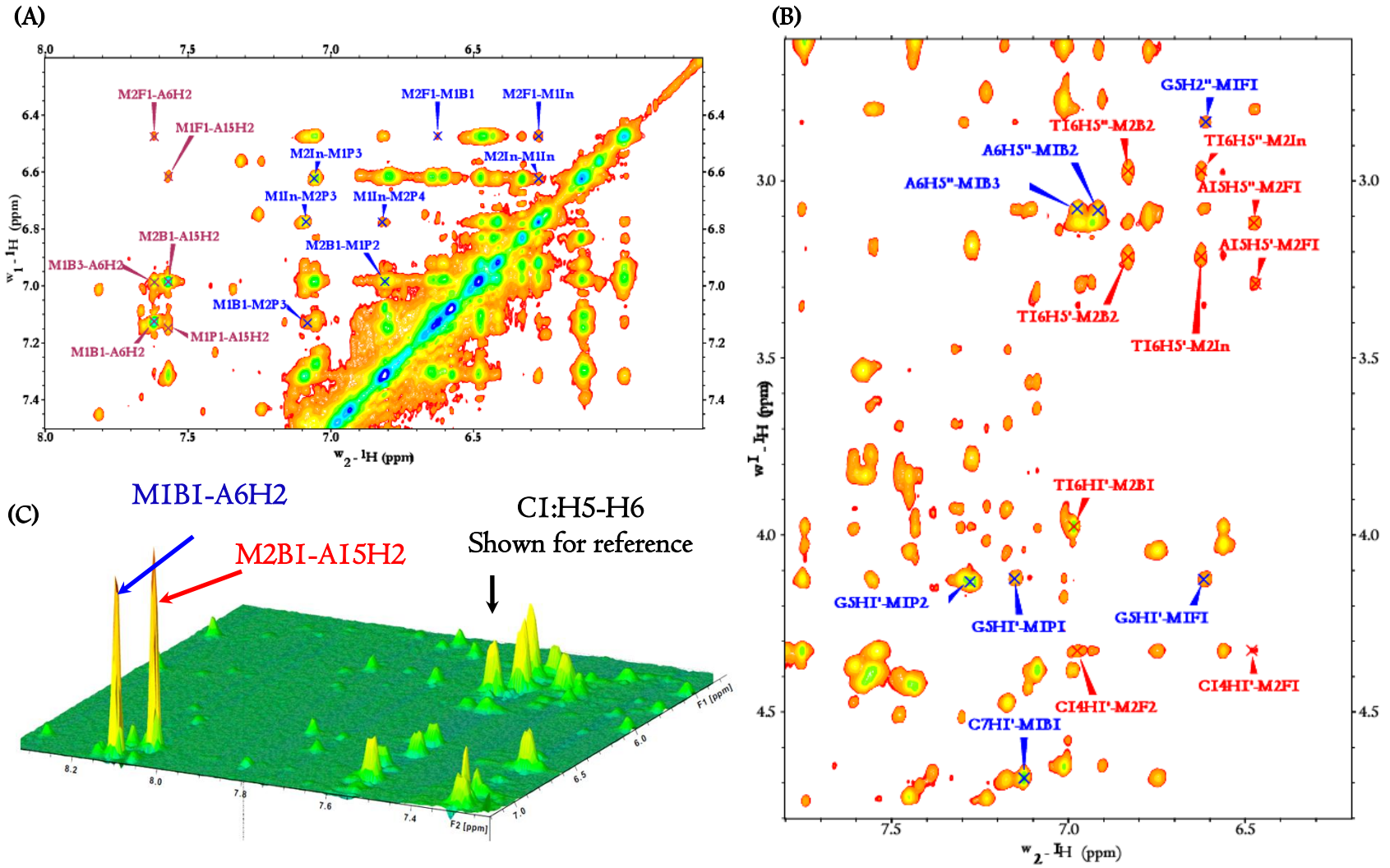


Figure 6.11: Intermolecular crosspeaks between the two DB1878 molecules and between DB1878 and DNA protons in the aromatic region (A), H1' region (B) of the NOESY spectrum of the 2:1 complex at 5 °C. The 2D stacked plot illustrating the strong interactions between the B1 protons and adenine residues are in (C).

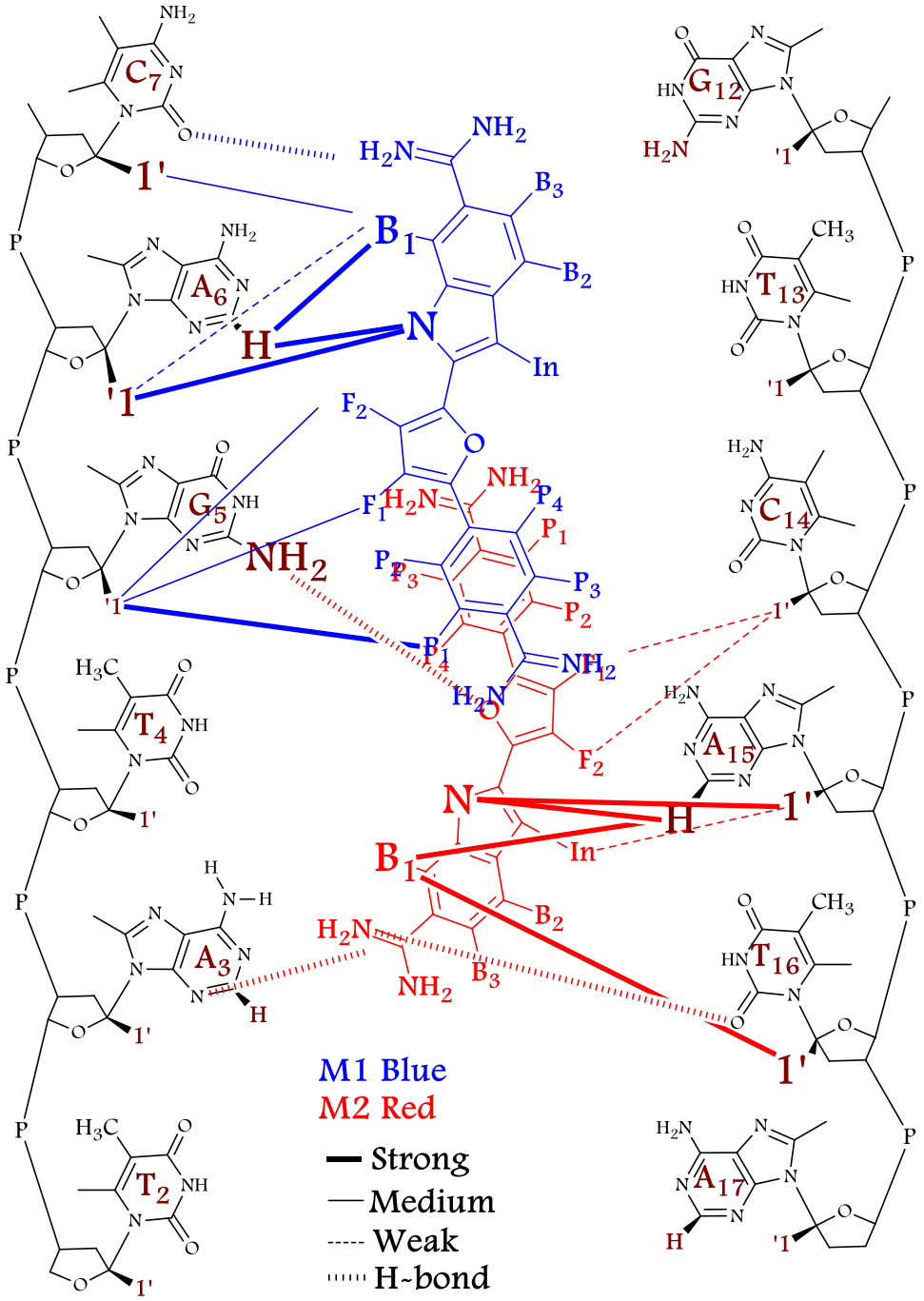


Figure 6.12: Some of the major NOE interactions observed between the two drug protons and the DNA.

Strong interactions are observed between the NH and B1 protons of the two molecules with AH2 and some of the H1' protons of the DNA. The G5NH2 is proposed to form a hydrogen-bond with the furan oxygen pointed into the minor groove based on the chemical shift analysis if G5-NH2. The terminal diamidines are also proposed to potentially participate in hydrogen bonding with acceptors on the DNA bases.

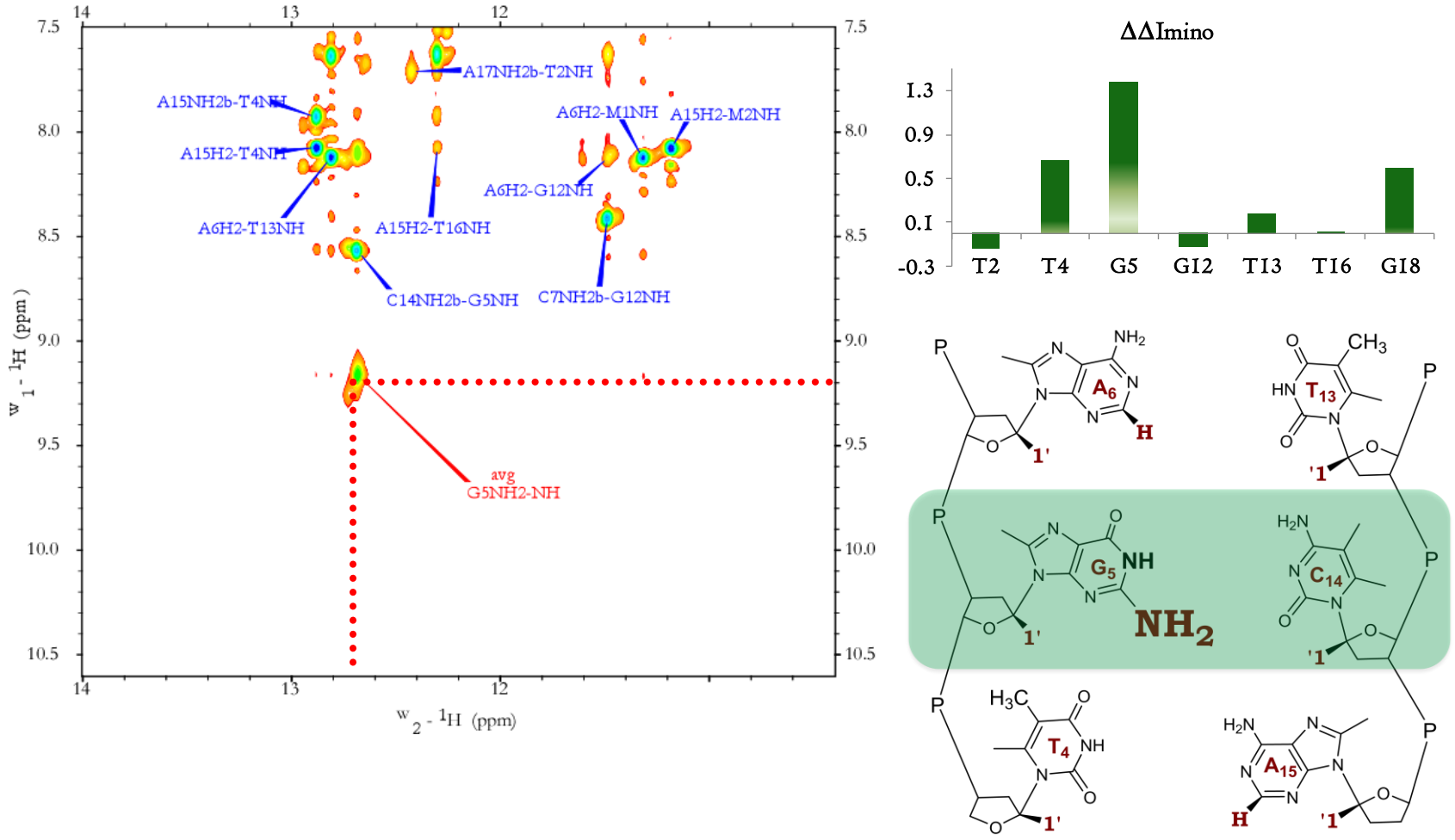


Figure 6.13: Expanded NOESY spectrum of the imino proton region of 2:1 complex of DB1878 with oligo2-1 at 5 °C.

The red dotted lines connect the G5-NH2 proton with G5-NH proton. The G5-NH2 proton is proposed to be involved in hydrogen-bonding with the furan oxygen of DB1878. A schematic representation of the G•C base-pair important for the dimer formation by DB1878.

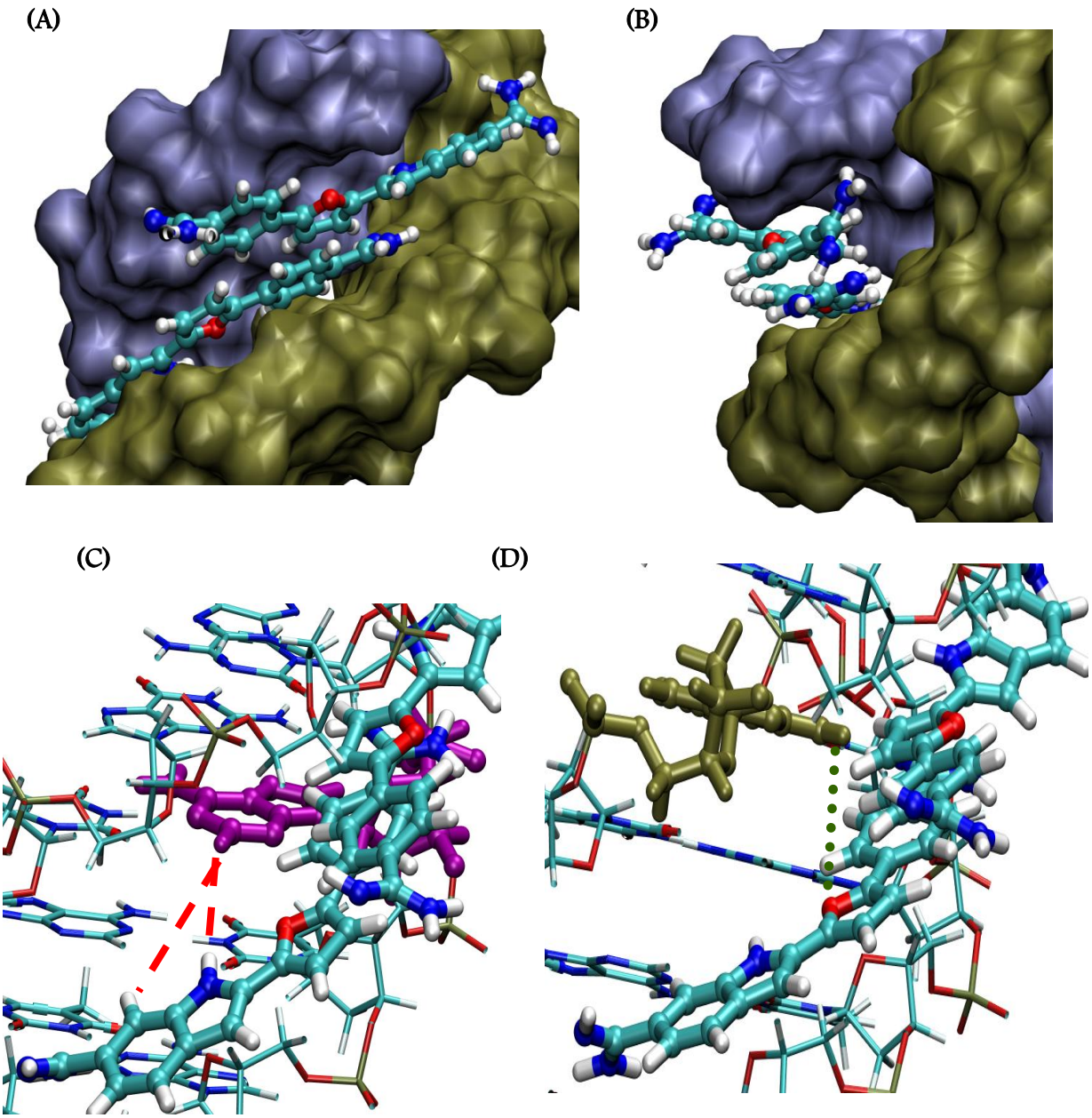


Figure 6.14: Preliminary docking studies of DB1878 complexed with oligo2-1 based on NMR data. (A) View into the groove of oligo2-1 with the two DB1878 molecules (B) View along the groove of the helix (C) interaction between M2B1 and M2NH with A15H2 and (D) possible H-bond between G5NH2 and furan oxygen.

Table 6.1: Exchangeable and non-exchangeable proton assignment of free oligo2-1 at 5 °C.

Base	H1'	H2'	H2"	H3'	H4'	H6/H8	H2	H5/CH3	NH
C_1	5.58	2.01	2.34	4.42	3.88	7.63	--	5.66	--
T_2	5.55	2.11	2.41	4.70	3.97	7.43	--	1.49	13.58
A_3	6.09	2.52	2.77	4.90	4.22	8.21	7.10	--	--
T_4	5.48	1.73	2.14	4.65	4.05	6.90	--	1.21	13.23
G_5	5.32	2.46	2.51	4.79	4.12	7.66	--	--	12.31
A_6	5.98	2.38	2.61	4.96	4.18	7.91	7.65	--	--
C_7	5.67	1.81	2.08	4.57	3.91	6.99	--	4.97	--
T_8	5.85	1.85	2.10	4.51	4.01	7.30	--	1.58	--
C_9	5.83	1.90	2.26	4.37	3.86	7.61	--	5.84	--
T_{10}	5.69	1.86	2.04	4.42	3.80	7.12	--	1.35	--
C_{11}	5.87	1.89	2.26	4.49	3.96	7.47	--	5.67	--
G_{12}	5.71	2.43	2.61	4.58	4.05	7.82	--	--	12.62
T_{13}	5.95	2.01	2.36	4.69	4.02	7.12	--	1.11	13.65
C_{14}	5.42	1.94	2.25	4.66	3.90	7.38	--	5.84	--
A_{15}	6.00	2.48	2.71	4.81	4.12	8.16	7.32	--	--
T_{16}	5.31	1.77	2.09	4.62	3.84	7.00	--	1.28	13.31
A_{17}	5.83	2.46	2.66	4.80	4.13	7.95	7.15	--	
G_{18}	5.76	2.20	2.02	4.40	3.91	7.46	--	--	12.96

Table 6.2: Assignments of the free drug and the bound drugs in the complex.

	DB1878 (25 °C)	Complex (at 5 °C)	
		M1	M2
F1	7.035	6.615	6.476
F2	6.912	7.105	6.978
B1	7.767	7.126	6.981
B2	7.683	6.915	6.831
B3	7.346	6.975	6.932
P1	7.825	7.146	6.985
P2	7.701	7.296	7.552
P3	7.825	7.549	7.588
P4	7.701	7.267	7.312
In	6.888	6.775	6.625
NH	~~~~	12.32	12.18

Table 6.3: Exchangeable and non-exchangeable proton assignment of the 2:1 complex of DB1878 with oligo2-1 at 5 °C.

Base	H1'	H2'	H2"	H3'	H4'	H5'	H5"	H6 /H8	H2	H5 /CH3	NH
C_1	5.54	2.04	2.27	4.43	3.98			7.58		5.59	
T_2	5.64	2.15	2.49	4.74	4.05			7.45		1.48	13.43
A_3	6.10	2.62	2.77	4.90	4.27			8.31	7.31		
T_4	5.40	1.56	1.85	4.65	4.18			7.01		1.16	13.88
G_5	4.12	2.31	2.83	4.37	3.98			7.84			13.69
A_6	5.45	1.96	2.28	4.33		3.29	3.08	7.75	8.11		
C_7	4.69	1.31	1.63	4.04	0.85	3.09	2.63	6.75		5.10	
T_8	5.86	1.94	2.19	4.47	3.92			7.17		1.46	
C_9	6.02	1.90	2.26	4.37				7.61		5.81	
T_{10}	5.68	1.80	2.12	4.39	3.64			7.09		1.41	
C_{11}	6.02	1.95	2.27	4.41				7.56		5.73	
G_{12}	5.76	2.55	2.75	4.74				7.91			12.49
T_{13}	6.02	2.05	2.27	4.75				7.23		1.14	13.81
C_{14}	4.33	2.00	2.06	4.39				7.54		5.53	
A_{15}	5.35	1.95	2.32	4.33		3.29	3.12	7.82	8.07		
T_{16}	3.97	1.23	1.37	4.03	1.09	3.21	2.97	6.56		1.36	13.31
A_{17}	5.79	2.29	2.60	4.75				7.74	7.08		
G_{18}	5.71	2.21	2.02	4.43	3.85			7.43			12.95

Table 6.4: Major intermolecular crosspeaks between two DB1878 molecules.

M2B1 6.98	M1P2 7.29	Strong
M2In 6.62	M1P3 7.55	Strong
M1In 6.77	M2P3 7.89	Strong
M1In 6.77	M2In 6.62	Medium
M2F1 6.47	M1In 6.62	Medium
M1B1 7.13	M2F1 6.47	Weak
M1In 6.77	M2P4 7.31	Medium
M1B1 7.13	M2P3 7.59	Medium

Table 6.5: Major intermolecular crosspeaks between the two ligands and DNA protons.

M2B1 6.98	A15H2 8.06	Strong
M1B1 7.13	A6H2 8.11	Strong
M1F1 6.61	A15H2 8.06	Weak
M2F1 6.47	A6H2 8.11	Weak
M1B3 6.97	A6H2 8.11	Weak
M1P1 7.14	A15H2 8.06	Weak
M1P1 7.14	G5H1' 4.12	Medium
M1P2 7.29	G5H1' 4.12	Strong
M1F1 6.61	G5H1' 4.12	Medium
M2F1 6.47	C14H1' 4.32	Very Weak
M2F2 6.97	C14H1' 4.32	Medium
M2B1 6.98	T16H1' 4.02	Medium
M1B1 7.12	C7H1' 4.68	Medium
M1B3 6.97	A6H5" 3.08	Weak
M1B2 6.91	A6H5" 3.08	Weak
M1F1 6.61	G5H2" 2.83	Weak
M2B2 6.83	T16H5" 2.96	Weak
M2In 6.62	T16H5" 2.96	Weak
M2F1 6.47	A15H5" 3.12	Weak
M2F1 6.47	A15H5' 3.29	Weak
M2In 6.62	T15H5' 3.21	Medium
M2B2 6.83	T16H5' 3.21	Medium
M1NH 12.32	A6H2 8.11	Strong
M1NH	A6H1' 5.44	Strong
M1NH	C7H1' 4.68	Medium
M1NH	G5H1' 4.12	Weak
M1NH	A6H5' 3.29	Medium
M1NH	A6H5" 3.07	Medium
M1NH	A6H2' 1.96	Medium
M1NH	A6H2" 2.27	Medium
M1NH	C7H4' 0.85	Medium
M2NH 12.18	T16H1' 3.97	Medium
M2NH	A15H1' 5.34	Strong
M2NH	T16H5' 3.21	Medium
M2NH	T16H5" 2.96	Medium
M2NH	A15H2' 1.95	Medium
M2NH	A15H2" 2.32	Medium
M2NH	T16H4' 1.09	Medium
M2NH	A15H2 8.06	Strong

6.5 REFERENCES

[1] B. S. P. Reddy, S. M. Sondhi, J. W. Lown, *Pharmacology & Therapeutics* **1999**, *84*, 1.

[2] a)S. Nelson, L. Ferguson, W. Denny, *Cell & Chromosome* **2004**, *3*, 2; b)M. Gniazdowski, W. Denny, S. Nelson, M. Czyz, *Curr Med Chem* **2003**, *10*, 909 ; c)M. Gniazdowski, W. A. Denny, S. M. Nelson, M. Czyz, *Expert Opinion on Therapeutic Targets* **2005**, *9*, 471.

[3] a)A. V. Vargiu, P. Ruggerone, A. Magistrato, P. Carloni, *Biophysical Journal* **2008**, *94*, 550; b)G. Singhal, M. R. Rajeswari, *Journal of Molecular Structure* **2009**, *920*, 208.

[4] P. G. Baraldi, A. Bovero, F. Fruttarolo, D. Preti, M. A. Tabrizi, M. G. Pavani, R. Romagnoli, *Medicinal Research Reviews* **2004**, *24*, 475.

[5] D. M. Herman, E. E. Baird, P. B. Dervan, *Journal of the American Chemical Society* **1998**, *120*, 1382.

[6] a)W. C. Tse, D. L. Boger, *Chemistry & Biology* **2004**, *11*, 1607; b)P. Dervan, B. Edelson, *Curr Opin Struct Biol* **2003**, *13*, 284 ; c)R. Bremer, J. Szewczyk, E. Baird, P. Dervan, *Bioorg Med Chem* **2000**, *8*, 1947 ; d)P. B. Dervan, R. W. Bürli, *Current Opinion in Chemical Biology* **1999**, *3*, 688.

[7] S. M. Nelson, L. R. Ferguson, W. A. Denny, *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* **2007**, *623*, 24.

[8] a)G. Orphanides, T. Lagrange, D. Reinberg, *Genes & Development* **1996**, *10*, 2657; b)W. D. Wilson, B. Nguyen, F. A. Tanious, A. Mathis, J. E. Hall, C. E. Stephens, D. W. Boykin, *Current Medicinal Chemistry - Anti-Cancer Agents* **2005**, *5*, 389.

[9] a)M. Sands, M. A. Kron, R. B. Brown, *Reviews of Infectious Diseases* **1985**, *7*, 625; b)R. D. PEARSON, E. L. HEWLETT, *Annals of Internal Medicine* **1985**, *103*, 782; c)M. P. Hutchinson, H. J. C. Watson, *Transactions of the Royal Society of Tropical Medicine and Hygiene* **1962**, *56*, 227; d)E. A. ELAMIN, A. M. HOMEIDA, S. E. I. ADAM, M. M. MAHMOUD, *Journal of Veterinary Pharmacology and Therapeutics* **1982**, *5*, 259; e)F. Doua, T. W. Miezan, J. R. S. Singaro, F. B. Yapo, T. Baltz, *Am J Trop Med Hyg* **1996**, *55*, 586; f)A. J. Nok, *Parasitology Research* **2003**, *90*, 71.

[10] a)J. Woynarowski, *Biochim Biophys Acta* **2002**, *1587*, 300 ; b)T. A. Beerman, J. M. Woynarowski, R. D. Sigmund, L. S. Gawron, K. E. Rao, J. W. Lown, *Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression* **1991**, *1090*, 52; c)R. M. Brosh, Jr, J. K. Karow, E. J. White, N. D. Shaw, I. D. Hickson, V. A. Bohr, *Nucl. Acids Res.* **2000**, *28*, 2420; d)J. W. George, S. Ghate, S. W. Matson, J. M. Besterman, *Journal of Biological Chemistry* **1992**, *267:*, 10683; e)J. M. Woynarowski, R. D. Sigmund, T. A. Beerman, *Biochemistry* **1989**, *28*, 3850.

[11] a)M. Yuki, V. Grukhin, C.-S. Lee, I. S. Haworth, *Archives of Biochemistry and Biophysics* **1996**, *325*, 39; b)M. M. Patel, T. J. Anchordoquy, *Biophysical Chemistry* **2006**, *122*, 5; c)N. Schmid, J. P. Behr, *Biochemistry* **1991**, *30*, 4357.

[12] a)S. Lindemose, P. E. Nielsen, N. E. Mollegaard, *Nucl. Acids Res.* **2005**, *33*, 1790; b)K. D. Stewart, *Biochemical and Biophysical Research Communications* **1988**, *152*, 1441.

[13] a)M.-H. David-Cordonnier, C. Gajate, O. Olmea, W. Laine, J. de la Iglesia-Vicente, C. Perez, C. Cuevas, G. Otero, I. Manzanares, C. Bailly, F. Mollinedo, *Chemistry & Biology* **2005**, *12*, 1201; b)M. D'Incalci, C. M. Galmarini, *Molecular Cancer Therapeutics* **2010**, *9*, 2157.

[14] a)M. J. Waring, C. Bailly, *Gene* **1994**, *149*, 69; b)C. Baily, M. J. Waring, *Nucl. Acids Res.* **1995**, *23*, 885.

[15] D. Goodsell, R. E. Dickerson, *journal of medicinal chemistry* **1986**, *29*, 727.

[16] a)C. L. Kielkopf, E. E. Baird, P. B. Dervan, D. C. Rees, *Nat Struct Mol Biol* **1998**, *5*, 104; b)S. White, J. W. Szewczyk, J. M. Turner, E. E. Baird, P. B. Dervan, *Nature* **1998**, *391*, 468; c)W. L. Walker, M. L. Kopka, D. S. Goodsell, *Biopolymers* **1997**, *44*, 323.

[17] C. Bailly, in *Advances in DNA Sequence-Specific Agents, Vol. Volume 3* (Eds.: B. J. Graham, P. Manlio), Elsevier, **1998**, pp. 97.

[18] a)B. Plouvier, R. Houssin, B. Hecquet, P. Colson, C. Houssier, M. J. Waring, J.-P. Henichart, C. Bailly, *Bioconjugate Chemistry* **1994**, *5*, 475; b)P. L. James, E. E. Merkina, A. I. Khalaf, C. J. Suckling, R. D. Waigh, T. Brown, K. R. Fox, *Nucl. Acids Res.* **2004**, *32*, 3410.

[19] a)M. E. Filipowsky, M. L. Kopka, M. Brazil-Zison, J. W. Lown, R. E. Dickerson, *Biochemistry* **1996**, *35*, 15397; b)A. Alfieri, Animati, F., Arcamone, F., Bailly, C., Dentini, M., Felicetti, P., Iafrate, E., Lombardi, P., Manzini, S., Rossi, C., and Waring, M. J. , *Antiviral Chemistry Chemotherapy* **1997**, *8*.

[20] a)S. M. Rahmathullah, J. E. Hall, B. C. Bender, D. R. McCurdy, R. R. Tidwell, D. W. Boykin, *journal of medicinal chemistry* **1999**, *42*, 3994; b)D. W. Boykin, A. Kumar, J. Spychala, M. Zhou, R. J. Lombardy, W. D. Wilson, C. C. Dykstra, S. K. Jones, J. E. Hall, *journal of medicinal chemistry* **1995**, *38*, 912; c)M. N. Soeiro, E. M. De Souza, C. E. Stephens, D. W. Boykin, *Expert Opinion on Investigational Drugs* **2005**, *14*, 957; d)W. D. Wilson, F. A. Tanious, A. Mathis, D. Tevis, J. E. Hall, D. W. Boykin, *Biochimie* **2008**, *90*, 999; e)I. Francesconi, W. D. Wilson, F. A. Tanious, J. E. Hall, B. C. Bender, R. R. Tidwell, D. McCurdy, D. W. Boykin, *journal of medicinal chemistry* **1999**, *42*, 2260.

[21] C. A. Laughton, F. Tanious, C. M. Nunn, D. W. Boykin, W. D. Wilson, S. Neidle, *Biochemistry* **1996**, *35*, 5655.

[22] a)T. Wenzler, D. W. Boykin, M. A. Ismail, J. E. Hall, R. R. Tidwell, R. Brun, *Antimicrob. Agents Chemother.* **2009**, *53*, 4185; b)I. Midgley, K. Fitzpatrick, L. M. Taylor, T. L. Houchen, S. J. Henderson, S. J. Wright, Z. R. Cybulski, B. A. John, A. McBurney, D. W. Boykin, K. L. Trendler, *Drug Metabolism and Disposition* **2007**, *35*, 955; c)J. K. Thuita, S. M. Karanja, T. Wenzler, R. E. Mdachi, J. M. Ngotho, J. M. Kagira, R. Tidwell, R. Brun, *Acta Tropica* **2008**, *108*, 6.

[23] L. Wang, C. Bailly, A. Kumar, D. Ding, M. Bajic, D. W. Boykin, W. D. Wilson, *Proceedings of the National Academy of Sciences of the United States of America* **2000**, *97*, 12.

[24] C. Bailly, C. Tardy, L. Wang, B. Armitage, K. Hopkins, A. Kumar, G. B. Schuster, D. W. Boykin, W. D. Wilson, *Biochemistry* **2001**, *40*, 9770.

[25] a)B. Nguyen, F. A. Tanious, W. D. Wilson, *Methods* **2007**, *42*, 150; b)F. A. Tanious, B. Nguyen, W. D. Wilson, in *Methods in Cell Biology, Vol. Volume 84* (Eds.: J. C. Dr. John, Dr. H. William Detrich, III), Academic Press, **2008**, pp. 53.

[26] a)D. Marion, K. Wüthrich, *Biochemical and Biophysical Research Communications* **1983**, *113*, 967; b)M. Rance, O. W. Sørensen, G. Bodenhausen, G. Wagner, R. R. Ernst, K. Wüthrich, *Biochemical and Biophysical Research Communications* **1983**, *117*, 479.

[27] a)M. Kadkhodaei, T. L. Hwang, J. Tang, A. J. Shaka, *Journal of Magnetic Resonance, Series A* **1993**, *105*, 104; b)L. R. Brown, B. C. Sanctuary, *Journal of Magnetic Resonance (1969)* **1991**, *91*, 413; c)C. Griesinger, G. Otting, K. Wuethrich, R. R. Ernst, *Journal of the American Chemical Society* **1988**, *110*, 7870.

[28] a)A. Kumar, R. R. Ernst, K. Wüthrich, *Biochemical and Biophysical Research Communications* **1980**, *95*, 1; b)R. Richarz, K. Wüthrich, *Journal of Magnetic Resonance (1969)* **1978**, *30*, 147; c)G. Otting, K. Wüthrich, *Journal of Magnetic Resonance (1969)* **1989**, *85*, 586.

[29] K. Wüthrich, *NMR of Proteins and Nucleic Acids*, John Wiley & Sons, **1986**.

[30] a)L. G. Marzilli, D. L. Banville, G. Zon, *Journal of the American Chemical Society* **1986**, *108*, 4188; b)M. S. Masaaki Tabata, Katumi Yoshioka ,Hiroaki Kodama, *Analytical Sciences* **1990**, *6*, 5.

[31] a)B. Reddy, S. Sharma, J. Lown, *Curr Med Chem* **2001**, *8*, 475 ; b)J. Gottesfeld, C. Melander, R. Suto, H. Raviol, K. Luger, P. Dervan, *J Mol Biol* **2001**, *309*, 615