University of the Pacific Theses and Dissertations

Graduate School

2020

STUDIES ON THE SIZE AND NON-PLANARITY OF AROMATIC STACKING MOIETY ON CONFORMATION SELECTIVITY AND THERMAL STABILIZATION OF G-QUADRUPLEXES

Mandeep Singh
University of the Pacific

STUDIES ON THE SIZE AND NON-PLANARITY OF AROMATIC STACKING MOIETY
ON CONFORMATION SELECTIVITY AND THERMAL STABILIZATION OF
G-QUADRUPLEXES

By

Mandeep Singh

A Dissertation Submitted to the

Graduate School

In Partial Fulfillment of the

Requirements for the Degree of

DOCTOR OF PHILOSOPHY

Thomas J. Long School of Pharmacy and Health Sciences
Pharmaceuticals and Chemical Sciences

University of the Pacific
Stockton, California

2020

STUDIES ON THE SIZE AND NON-PLANARITY OF AROMATIC STACKING MOIETY
ON CONFORMATION SELECTIVITY AND THERMAL STABILIZATION OF
G-QUADRUPLEXES

By

Mandeep Singh

APPROVED BY:

Dissertation Advisor: Liang Xue, Ph.D.

Committee Member: Andreas Franz, Ph.D.

Committee Member: Jianhua Ren, Ph.D.

Committee Member: Georgios Pantouris, Ph.D.

Committee Member: Geoffrey Lin-Cereghino

Department Co-Chair: Jianhua Ren, Ph.D.

Department Co-Chair: Jerry Tsai, Ph.D.

STUDIES ON THE SIZE AND NON-PLANARITY OF AROMATIC STACKING MOIETY ON CONFORMATION SELECTIVITY AND THERMAL STABILIZATION OF G-QUADRUPLEXES

Copyright 2020

By

Mandeep Singh

DEDICATION

This dissertation is dedicated to my daughter Kav Neer Kaur, my universe.

ACKNOWLEDGEMENTS

First, thanks to god for providing me the strength and courage through all these years.

I am grateful to my advisor Dr. Liang Xue for his guidance, help, and support to conduct research and dissertation writing. He helped me growing as an independent researcher. My sincerest thanks to him. This landmark in my life would not have been possible without his help. I am sure that we will have a lifelong company.

I would like to thank my dissertation committee members Dr. Jianhua Ren, Dr. Andreas Franz, Dr. Georgios Pantouris and Dr. Geoffery Lin-Cereghino for agreeing to serve and provide valuable feedback on my dissertation.

I am thankful to Ms. Sue McCann. She made my stay at Pacific very comfortable during my early years. I always felt better after talking to her.

I want to thank Dr. Brett Williams, who has been an exceptional caretaker of graduate students. Without his willingness to help, it would have been very difficult for me to conduct research. For me, he is a silent engine of the department that keeps things running.

I would like to express my sincere gratitude to Dr. Franz and Dr. Samoshin for their support with the NMR facility.

Thanks to the MS facility people Yadwinder, Yuntao, and Michael for their help in running my mass spectrometry samples.

Special thanks to Krege Christison for his help with the HRMS of my compounds.

I made friends from multiple nations with different backgrounds. I am grateful to them for creating such an inclusive environment here at the Pacific.

I want to express my appreciation for my current and former lab mates, Vanessa, Landy, Dr. Siwen Wang and Brock Allen for their willingness to help.

My friends Marcos and Oscar both share a common trait of care and helpfulness. I will always be indebted to them for their support like showing up, especially for me to run my NMR on weekends. Thanks to Boris for listening to my jokes and laugh without understanding.

Thanks Carim for the helpful chemistry discussion.

My friend Anoopjit has helped me whenever I needed it the most. Thanks to him. Thanks to Anoop's mom for taking care during my initial stay in the USA. I will never be able to pay back for the care and love I received from Anoop's family. I also thanks Yadwinder for his support in the department and outside.

I cannot express in words the support I received from my father Swarn Singh, my mother Manjit Kaur and my wife Kanwal Kaur. My daughter Kav Kaur has been an inspiration for me to finish this task. I am thankful to my sisters Simpy and Rimpy for having faith in me. I am thankful to my brother Damandeep for his continued support in the USA from last 5 years in every respect of my daily life.

Finally, I am grateful to all the people who have crossed my path during Ph.D.

STUDIES ON THE SIZE AND NON-PLANARITY OF AROMATIC STACKING MOIETY
ON CONFORMATION SELECTIVITY AND THERMAL STABILIZATION OF
G-QUADRUPLEXES

Abstract

By Mandeep Singh

University of the Pacific
2020

Targeting DNA has the advantage over proteins for cancer remediation because of the fewer copies of the ligands required for the desired therapeutic effect. Traditionally, covalent DNA binders like alkylating agents have been used to induce genetic instability through the formation of DNA lesions and strand breaks, leading to cellular apoptosis. The primary drawback of this treatment is the non-specific binding that affects both cancerous and non-cancerous cells.

G-quadruplexes are the DNA secondary structures that are present in abundance near the promoter regions of the oncogenes and are involved in the regulation of their activities. A ligand-mediated stabilization of G-quadruplexes in the promoter regions and down-regulation of the associated oncogenes have been validated. In contrast to alkylating agents, G-quadruplex ligands induce genetic stabilization through non-covalent interactions. They can be designed to interact specifically with G-quadruplex DNA over duplex DNA, which reduce side effects arising from the off-targeting.

G-quadruplex ligands invariably have the large planar aromatic moiety to interact with G-quadruplexes through π-π stacking interactions. For determining the size effect of the aromatic moiety on stabilization of G-quadruplexes, a series of ligands were synthesized by conjugating

nucleobases or 1,10-phenanthroline with an aminoglycoside, neomycin. The resulting conjugates increased the binding affinity synergistically and enabled us to study the effect of the stacking moiety required for G-quadruplex stabilization. Nucleobase-neomycin conjugates did not show stabilization stabilize of human telomeric G-quadruplex. 1,10-Phenanthroline-neomycin conjugate (**7b**) on the other hand binds to human telomeric G-quadruplex with a K_a of $(8.9\pm2.4)\times10^8$ M^{-1} and inhibits telomerase activity at 1.56 µM probably through G-quadruplex stabilization.

Moving forward, we further enlarged the aromatic moiety by tethering two 1,10-phenantholine molecules together through a five-atom linker. The resulting molecule (2-Clip-phen) was conjugated with various amino-containing side chains. 2-Clip-phen derivatives showed at least 30 times weaker binding to duplex DNA over G-quadruplex DNA. In addition, compounds showed a preference for the antiparallel G-quadruplex conformation over parallel and hybrid G-quadruplex conformations, as shown in the CD spectroscopy studies. Ligands **11** and **13** induced the formation of an antiparallel G-quadruplex from random coils and stabilize it to 60 ºC (T_m) in a salt-free condition. Mass spectrometry study showed the formation of a two-tetrad G-quadruplex with the 2-Clip-phen ligand. Docking study showed that the ligand interacts most favorably with antiparallel G-quadruplex conformation, which is supported further by the larger thermal stabilization effect on antiparallel G-quadruplex compared with other G-quadruplex conformations. Our study suggests that 2-Clip-phen can be used as a scaffold for designing G-quadruplex binding ligands that preferentially bind to antiparallel G-quadruplexes, which has never been reported before.

TABLE OF CONTENTS

LIST OF TABLES

Table

LIST OF FIGURES

Figure

LIST OF SCHEMES

Scheme

LIST OF ABBREVIATIONS

A	Adenine
APH	Aminoglycoside phosphotransferase
BCL	B-cell lymphoma
2-CP	2-Clip-phen
CD	Circular dichroism
DCM	Dichloromethane
DMAP	4-Dimethylaminopyridine
DMF	Dimethylformamide
DSBs	Double strand breaks
ESI	Electrospray ionization
FID	Fluorescence intercalator displacement
G4	G-quadruplex
ITC	Isothermal titration calorimetry
Htelo	Human telomeric
IS	Internal standard
mAbs	Monoclonal antibodes
MIC	Minimum inhibitory concentration
MTT	3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide
NMR	Nuclear magnetic resonance
PCR	Polymerase chain reaction
Phen	1,10-Phenanthroline

TCDP	1,1'-thiocarbomul-di-2(1H)-pyridone
TKI	Tyrosine kinase inhibitor
TPS-Cl	2,4,6-triisopropylbenzenesulphonyl chloride
TRAP	Telomeric repeat amplification protocol
T_m	Melting temperature
UV	Ultravoilet

CHAPTER 1: INTRODUCTION

1.1 Cancer Is a Genetic Disease

The central dogma of molecular biology[1] (governs the flow of information in the cells. The genetic code in the DNA is transcribed into the mRNA in the nucleus, which is then transported into the cytoplasm, where it is translated into the polypeptides through the agency of ribosomes. Therefore, the information originally stored in the genes (genetic code) is expressed in the cells through the polypeptides/proteins.
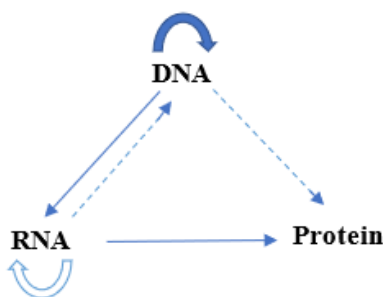
Figure 1.1. The central dogma of molecular biology. Solid arrows present general transfers, dotted and unfilled arrows represent special transfers.

Three decades ago, the prevalent belief was that some viruses were the source of cancer in humans; however, Varmus and Bishop's work[2] proved that this phenomenon could only account for nearly 20% of humans' cancers. In 1982, evidence was found that single nucleotide mutation in H-ras (G$\underline{G}$C → G$\underline{T}$C) leads to the incorporation of valine instead of glycine at the 12th position of p21 protein derived from T24 bladder carcinoma oncogene, resulting in the conversion of this proto-oncogene to its virulent form[3]. Wigler and his co-workers published a similar study on single-nucleotide mutations that lead to the conversion of N-ras proto-oncogene

a year later[4]. These pioneering works suggested that changes in the DNA (mutations) could lead to activation of the proto-oncogenes and inactivation of tumor suppressor genes (TSGs), factors that might be involved in development of cancer.

1.2. Chemotherapy of the Cancer

Chemotherapy for cancer treatment dates back to the 1950s. The first generation of chemotherapeutic agents is non-selective with severe side effects. For instance, nitrogen mustard is recognized as a non-selective DNA alkylating agent and an antimetabolite and tubulin-binding drug. In recent years, the main focus in this field is to develop targeted cancer therapy. Understanding the genetic origin of cancer on a molecular level has paved the way for studying biochemical pathways responsible for diseased states. Targeting these pathways leads to the development of targeted cancer therapy, one of the most significant milestones achieved in medical sciences. But this is not without caveats. These biochemical pathways are often involved in signal transduction, causing drug resistance via activation of the alternative pathways. In addition, mutations frequently occur in the proteins involved in these pathways, which changes the binding site of the drug to proteins, making it ineffective. Therefore, there is always a need for new alternative strategies to combat cancer. The following sections summarize the different chemotherapy drugs and the problems associated with their use.

1.2.1. Alkylating Agents

Alkylating agents are one of the oldest classes of chemotherapy drugs for the treatment of cancer. Their therapeutic effect was discovered serendipitously. During World War I and II, the observations made on the men exposed to mustard gas showed extended depletion of lymph nodes and bone marrow aplasia. In light of this observation, clinical studies with less toxic mechlorethamine (Mustragen) showed its remarkable effectiveness against Hodgkin's disease[5].

However, remissions were short and incomplete. This study is considered the dawn of modern cancer chemotherapy. Further research for safer analogs led to the development of the less toxic chlorambucil and cyclophosphamide activated in situ by liver enzymes.

The alkylating agents can be categorized into the following structural classes[6] (**Figure 1.2**)

1. **Nitrogen mustards** (mechlorethamine (**1**), chlorambucil (**2**), and melphalan (**3**))

2. **Oxaphosphorines mustards** (cyclophosphamide (**4**) and ifosfamide (**5**))

3. **Triazenes** (dacarbazine (**6**) and temozolomide (**7**))

4. **Nitrosoureas** (N, N'-bis(2-chloroethyl)-N-nitrosourea (BCNU) (**8**) and N-methyl-N-nitrosourea (MNNG) (**9**))

5. **Mitomycin** (mitomycin C (**10**))

6. **Platinum complexes** (cisplatin (**11**), oxaliplatin (**12**), and satraplatin (**13**))

Alkylating agents exert their functions by producing DNA cross-links between or within the DNA strands (except triazenes, which makes the DNA lesions, leading to strand breaks and apoptosis).
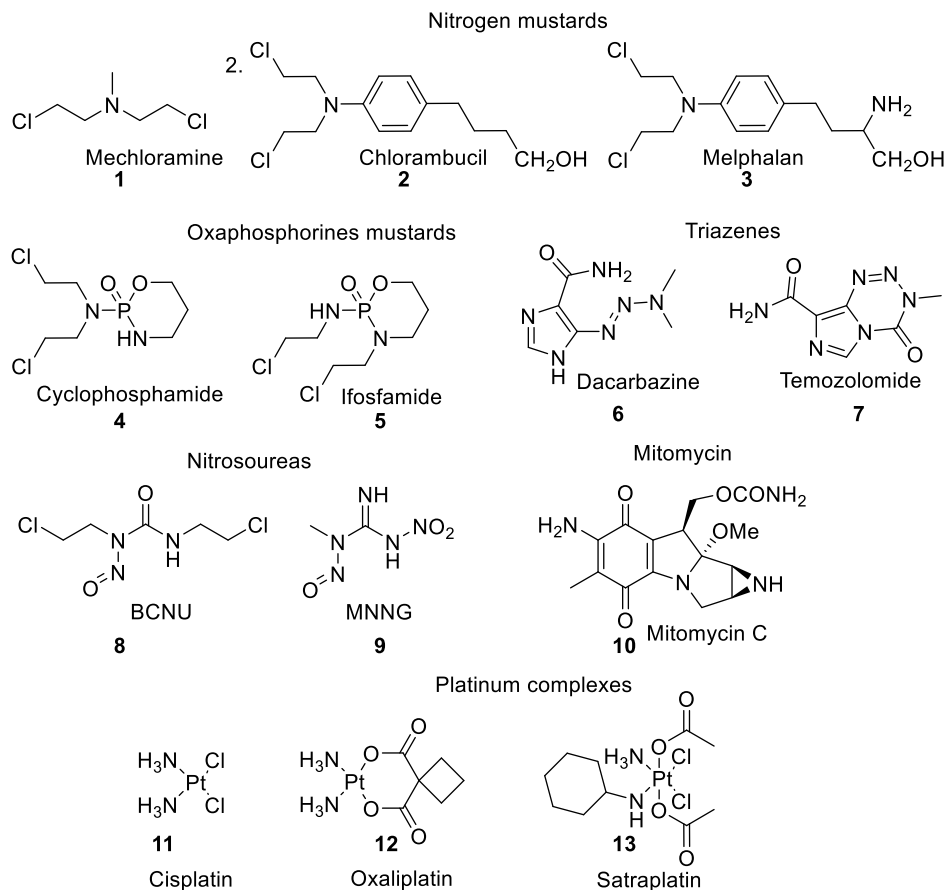
Figure 1.2. Different structural classes of alkylating agents.

The mechanism of action of mechlorethamine (nitrogen mustard) requires the generation of a strongly electrophilic aziridium ion through anchimeric assistance, which is then attacked by a weakly nucleophilic tertiary nitrogen on guanine bases (N-7). Two of such consecutive events lead to interstrand cross-links of the DNA (**Figure 1.3**), resulting in cellular apoptosis.
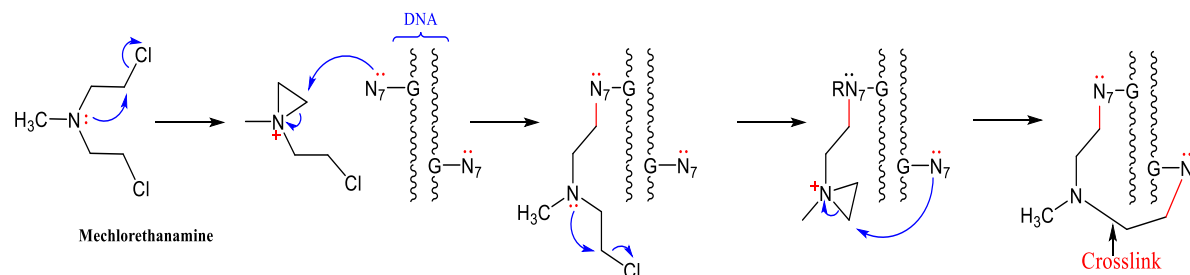
Figure 1.3. Mechanism of action of mechlorethamine (a nitrogen mustard).

1.2.2. Antimetabolites

In the 1950's nutritional scientists observed that folic acid was essential for bone marrow function. Sidney Farber[7] used a folic acid antagonist (Methotrexate (**14**)) (**Figure 1.4**) to treat leukemia in the children's medical center in Boston. The result showed temporary remission of leukemia progression in children. In 1954, Elion and Hitchings[8] carried out a series of experiments on antimetabolites using bacterial models. They observed a synergistic effect of antimetabolites targeting multiple biochemical pathways simultaneously, which ultimately led to the development of 6-mercaptopurine (**15**) and 6-thioguanine (**16**) (**Figure 1.4**). These two drugs were used to treat acute leukemia. In addition, they have been used for the treatment of herpes viral infection and as immunosuppressive agents. In 1957, fluorinated analogs of pyrimidines (e.g., 5-fluoro uracil (**17**) (**Figure 1.4**) and purines were synthesized and tested to treat solid tumors. It achieved remarkable tumor suppression in mice models[9]. These two drugs became an essential part of the combination of chemotherapy program[10] in the late 1970s.
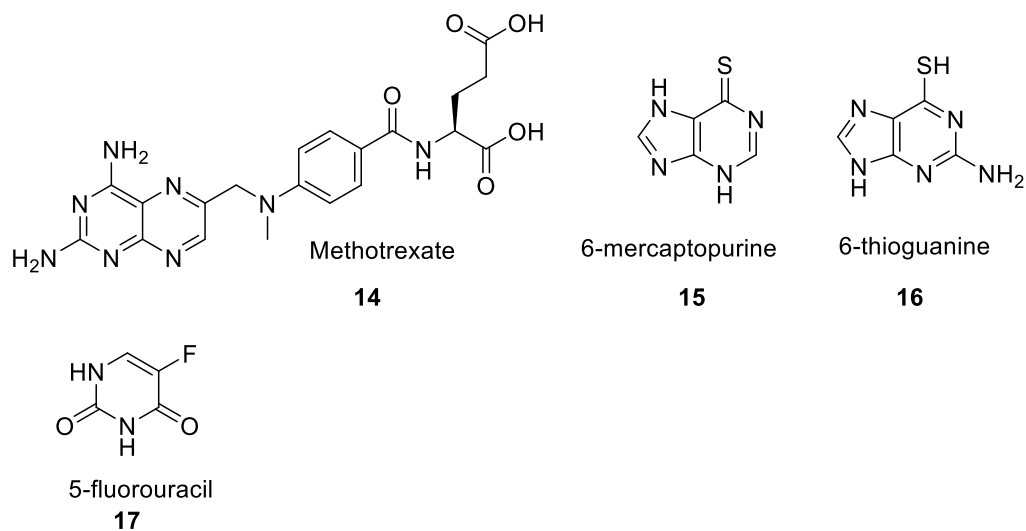
Figure 1.4. Structural of the different classes of antimetabolites.

1.2.3. Tubulin Binding Drugs

The Discovery of the anticancer activity of plant alkaloids from Vinca rosea viz

Vinblastine, vincristine, vindesine, and vinorelbine was a breakthrough in cancer treatment[11].

These alkaloids bind to the tubulin and inhibit the formation of spindle fibers (microtubules)

essential for cell division. Another drug in this class is Paclitaxel from the bark of the Pacific

yew tree Taxus brevifolia, which induces the multipolar division due to the formation of

additional spindle poles. Cell death occurs due to the missing one or more chromosome essential

for growth[12].

1.2.4. Targeted Cancer Therapy

Mutations that cause the transformation of proto-oncogenes, disruptions of tumor

suppressor genes, and cell cycle regulatory pathways are crucial factors of cancer development;

therefore, targeting the biochemical pathway/target is a practical approach to suppression of

tumor progression or growth. This approach is different from the conventional non-selective

cytotoxic chemotherapy treatment. Kinases are one of the largest family of proteins and most sought after for anticancer drug development. These proteins are involved in important cell signaling and signal transduction pathways. In the past two decades, the development of kinase inhibitors as a targeted cancer therapy had an exponential growth. In 2017 there were 201 WHO registered (INN) small molecule kinase entities[13].

Tamoxifen, the first drug developed for targeted cancer therapy, is an anti-hormonal drug that blocks an oestrogen receptor (OER) on the surface of the OER positive breast cancer cells and thus controls the growth of cancer. FDA approved the first targeted small molecule kinase inhibitor, Imatinib (nib = kinase inhibitors) in 2001 for Chronic myelogenous leukemia[14]. It executes its functions by inhibiting tyrosine kinase. Tyrosine kinase is over-activated due to interchromosomal exchange between chromosomes 9 and 22 to form a new shorter chromosome 22 (Philadelphia chromosome). The production of this fusion gene BCR-ABL1 (ABL = Abelson, BCR = breakpoint cluster) is responsible for a defective protein that renders the tyrosine kinase always on, causing the cells to divide uncontrollably.

Her2 (human epidermal growth factor receptor 2) proto-oncogene was discovered in 1984, located on chromosome 17, is overexpressed due to its amplification multiple copies[15]. It encodes for a transmembrane tyrosine kinase receptor protein overexpressed in HER2 positive breast cancers. The downstream signal transduction by this kinase leads to a cascade of events for cell proliferation and inhibition of apoptosis. In 1998, FDA approved trastuzumab (Herceptin), a humanized monoclonal antibody, to target and downregulate the HER2 overexpression (downregulate the intracellular tyrosine kinase)[16]. Even though the monoclonal antibodies (recognizes only one antigen) were first developed in 1975 by the hybridoma method

(murine origin), this was the first example of using the humanized monoclonal antibodies for targeted cancer therapy.

In addition to these targeted therapies, other targets such as CAR-T (Chimeric Antigenic Receptor) (CAR) T-Cell therapy, and immune checkpoint therapy (Nobel prize in 2018) have attracted extensive research in recent years.

1.3. Side Effects and Challenges in Current Therapies

Conventional cytotoxic alkylating agents are non-selective in their mechanisms of action. Because their primary target is the DNA, the rapidly dividing cells like bone marrow, hair forming cells, gonads, intestinal mucosal lining cells, and lymphoid tissues, along with cancer cells, are the most affected. Cytotoxic alkylating agents also tend to cause a high percentage of single base-pair mutations in the treated cells. Hence, the dosing regimen of alkylating agents must be tailored according to the patient response and recovery from the cytotoxic effect on the bone marrow[17].

Similarly, tubulin inhibitors are poised with hematological and neurological side effects limiting their clinical use[18]. The side effects associated with antimetabolites are hematological, including neutropenia, thrombocytopenia, myelosuppression, and mucositis[19].

Even targeted therapies are not without side effects, which are not severe most of the time. The side effects could result from the nature of the targeted pathway or the nature of the ligand (monoclinal antibodies or small molecules kinase inhibitors). It is noteworthy that off-targeting to other isoforms of the receptor can produce serious side effects[20].

Small molecule kinase inhibitors (tyrosine kinase inhibitors) are usually associated with the skin rash problem. A positive correlation between the response to the treatment and development of 2-4 grade skin rash has been observed in addition to the hepatotoxicity,

proteinuria, Hemorrhage, ocular toxicity. Some of these toxicities are the result of primary target interactions and while others are due to the off-targeting.

<div align="center">

1.4. Resistance to the Drugs

</div>

1.4.1. Resistance to Conventional Therapies

Drug resistance is the major factor that affects the clinical applicability of the drug. The resistance of the nitrogen-mustard class of alkylating agents includes two observations. 1) Increased activity of the Glutathione S-transferases (GSTs) that neutralize alkylating agents via substitution reactions. 2) Increased expression of the DNA repair enzymes and increased detoxification of the intermediates.

In the case of oxazaphosphorines, there is an increased expression of aldehyde dehydrogenase (ALDH1), which neutralizes the side product aldehydes and increases the activity of the GST and levels of Glutathione (GSH). There is also increased activity of the O^6-alkylguanine alkyltransferase (AGT), which helps for the repair of damaged DNA.

The resistance of triazenes mainly results from the AGT over-activation leading to dealkylation of a methyl group at the O6 of guanine.

The resistance to Mitomycin C is associated with its reduction to an inactive form by two-electron reduction with DT-diaphorase (DTD) or Xanthine dehydrogenase. Additionally, resistance to mitomycin C has been related to the overactivation of multidrug resistance-1 (MDR1) proteins.

The resistance to cisplatin comes from the overactivation of the enzyme involved in DNA damage repair ERCC2. Cisplatin also reacts with other nucleophilic species in the blood (albumin) and in the cytosol, such as glutathione and metallothionein. The resulting products are actively effluxed out of the cells by glutathione-S-conjugate pumps[6].

The resistance of tubulin-binding drugs results from the overexpression of efflux pumps such as P-glycoprotein and multidrug resistance-associated protein-1. These are ATP-dependent family of proteins located in the cell membrane. Mutations in drug binding sites and changes in tubulin isotypes are also responsible for observed drug resistance to tubulin-binding drugs[18].

1.4.2. Resistance to Targeted Cancer Therapy

It includes 1). Resistance to the small molecule kinase inhibitors. 2) Resistance to the monoclonal antibody (mAbs).

1). Resistance to kinase inhibitors in targeted therapy is primarily due to the mutation at the active site. For example, Glivec, the first tyrosine kinase inhibitor for targeting chimeric BCR-ABL1 protein-tyrosine kinase, loses its efficacy due to point mutations at the ABL tyrosine kinase domain by replacing amino acid 315 threonine with isoleucince[21]. Overexpression of multidrug resistance P-glycoproteins is also responsible for the resistance to the drug observed[22].

2). The resistance to mAbs can be divided into three categories[23].

a). It can arise due to antibody incompatibility. Some people develop antibodies against chimeric or murine origin antibodies. In some cases, the binding of the Fc site on monoclonal antibodies to FcR (Fc receptor) on the killer T cells could be ineffective due to polymorphism in the FcR (Fc receptor). This interaction is vital to recruiting killer T cells to kill cancer cells.

b). It can be of tumor-related origins. The tumors may develop resistance after initial exposure to the antibody. The blocked antigen sites with the antibody are compensated by decreasing the concentration of the antigen or increasing the number of pathways on the surface. Some tumors may induce mutation to create an additional pathway that stops the desired outcome of the events triggered by antibody-antigen combination. For instance, activation of BCL-2 gene in B cell non-Hodgkin's lymphoma stops the apoptotic pathway triggered by the antibody interaction.

Alternatively, drug resistance could result from the generation of mucin in response, which may mask the potential binding site (epitope masking)[24] on the cell surface. Also, the conformational changes of the binding site (at tumor surface) due to mutation may significantly reduce the efficacy.

c). It can result from the host-related factors. Drugs may behave differently in people of different ethnicity. For example, Japanese people have a higher concentration of circulating CD20 (cCD20). Because CD20 on the cell surface is the natural target of Rutiximab, the higher concentration of cCD20 may significantly reduce the efficacy

of mAbs-based drugs. Some ethnic groups can develop human anti-chimeric (HACA) and human anti-mouse antibodies (HAMA) against therapeutic antibodies. The development of personalized human mAbs has solved this problem to a significant extent.

At the current stage, we are not able to target all disease-related biochemical pathways. Many of the pathways still need further understanding. The combination of different therapies can effectively sensitize cancer cells to chemotherapy. In addition, simultaneous administrating drugs with different mechanisms of action increases the probability of remission significantly.[10] Therefore, despite the cytotoxic and secondary mutagenic side effects of alkylating agents, they are still used as popular chemotherapeutic agents against cancer.

There is always an urgent need for new compounds that can match the cytotoxicity of alkylating agents and the benefits of targeted cancer therapy. In recent years, ligands that target G-quadruplexes have attracted much attention. They can induce synthetic lethality in cancer cells deficient in the repair enzymes[25], and trigger genetic instability and apoptosis in cells[26,27] with minimum toxic effects[28]. Additional benefits of targeting DNA G-quadruplexes over the protein targets are described as follows[29].

1). Not all the proteins are druggable, but their corresponding genes can be targeted.

2). Point mutations at the drug binding site may make the drug ineffective[30] for the treatment. Targeting at the gene level (promoter region) can circumvent this problem.

3). Targeting at the gene level is an efficient approach instead of targeting the overexpressing oncogenic proteins. Fewer copies of the target mean less concentration of the ligand needed.

4). DNA ligands can differentiate between protein and DNA targets. Most of them don't follow Lipinksi's rule of five. DNA ligands can differentiate between the duplex DNA and tetraplex DNA like G-quadruplex.

The next part of the introductory chapter summarizes the basics of G-quadruplexes and G-quadruplex binding ligands.

1.5. G-quadruplexes

1.5.1. The Fundamentals of G-quadruplexes

G-quadruplexes are helical, non-canonical, dynamic, and polymorphic nucleic acid structures. The basic unit of the G-quadruplexes is the G-tetrad, in which four guanine molecules self-associate through the Hoogsteen hydrogen bonds[31] and form a planar structure. The formation of such a construction was revealed in x-ray diffraction of 5'-guanine monophosphate (GMP) by Gellert *et al* in 1962. G-tetrads formed in nucleic acid strand(s) can stack on top of each other to form a helical structure (G-quadruplex), stabilized through van der Waals interactions[32] (**Figure 1.5**)
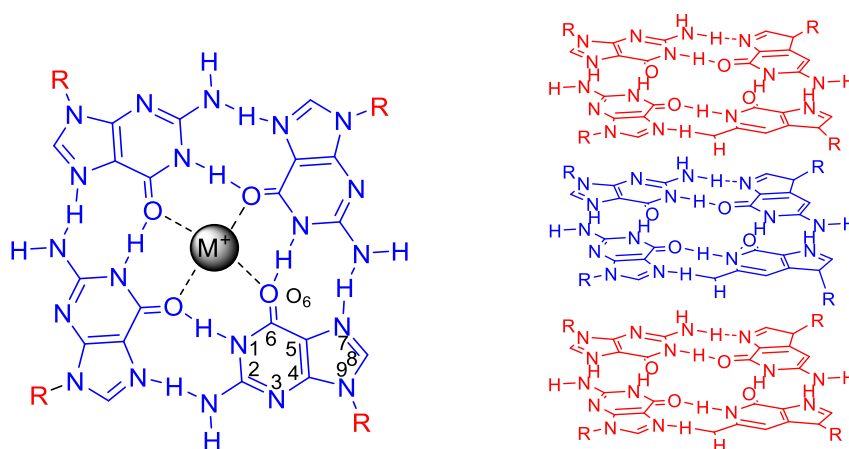
Figure 1.5. Structure of G-tetrad (four strands are labeled in red) (left) and overlap of G-tetrads due to π- π stacking interaction (right).

The cavity in the middle of G-tetrads is electron-rich, commonly occupied by monovalent metal ions such as Na^+ and K^+ under physiological conditions. The interactions between four oxygen atoms at the 6-position of guanine tetrads and metal ions are the key factor in the stabilization of G-quadruplexes. Due to its larger ionic radius (1.33 Å), K^+ is sandwiched

between the two G-tetrads and octa-coordinates with the O6 of guanines in the two G-tetrads (**Figure 1.5**). Na$^+$ ions (ionic radius 0.95 Å) can move freely along the central cavity[33]. The nature of the metal ion can affect the topology of G-quadruplex DNA. For instance, studies showed that the human telomeric G-rich DNA sequence 5'-AGGG(TTAGGG)-3' adopts a hybrid conformation in K$^+$ and a basket-type antiparallel conformation in Na^{+}[34,35]. Detailed information on G-quadruplex topologies will be discussed in the following sections.

Topology of the G-quadruplexes. G-quadruplexes are highly polymorphic structures. Typically, a DNA sequence containing more than four consecutive guanine tracks (three or more guanines in each track) can adopt an ensemble of conformation in dynamic equilibrium with each other[36,37]. G-quadruplexes are composed of stacked G-tetrads with negatively charged phosphate backbones that delimit cavities called grooves. Depending on relative strand orientations, G-quadruplexes can adopt parallel, antiparallel, and hybrid topologies. G-quadruplex strands can orient in four possible ways[38,39]. 1). All strands are oriented in the same direction. 2). Three strands are oriented in one direction, and the fourth is in the opposite direction. 3). Two neighboring strands are oriented in one direction, and the other two are oriented in the opposite direction (AABB like arrangement). 4). Two neighboring strands are oriented in opposite directions alternatively (ABAB like arrangement) (**Figure 1.6**). The strand orientations are related to the glycosidic conformation of the guanines, which can be either syn or anti (**Figure 1.7**). The relative orientations of the guanines in the three possible strand orientation is given below in Table 1.1[41]
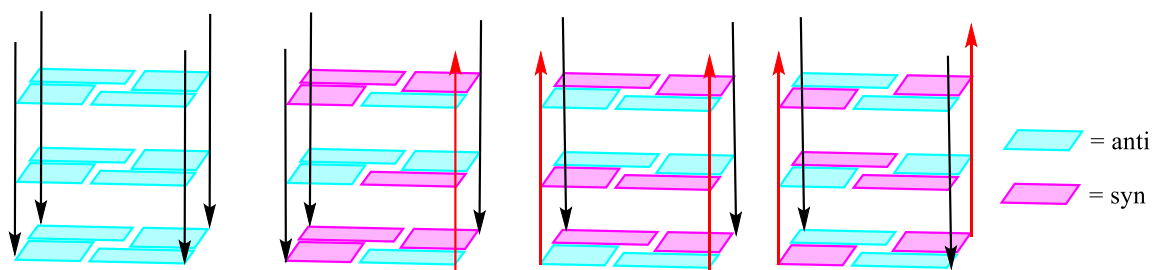
Figure 1.6. Four possible type of strand orientations in G-quadruplexes.
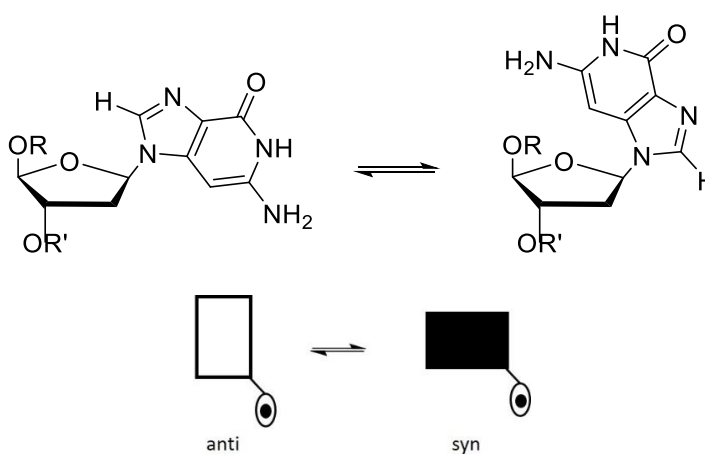
Figure 1.7. Syn and anti-glycosidic bond conformations of the guanine nuleoside[40].

Table 1.1
The Relationship Between the Relative Strand Orientation and Allowed G-Tetrad Configuration.

	Strand orientation	Allowed G-tetrad configuration	Type
1.	All strands oriented in the same direction	anti . anti. anti . anti syn . syn . syn . syn	Parallel
2.	Three strands are oriented in direction, and one is in the opposite direction	syn . anti . anti . anti anti . syn . syn . syn	Hybrid (3+1)
3.	Two neighboring strands are oriented in one direction, and the other two are in another direction	syn . syn . anti . anti	Antiparallel
4.	Two neighboring strands are oriented in opposite directions	syn . anti . syn . anti	Antiparallel

The grooves (3 types) between strands can be narrow (n), wide (w), or medium, depending on the directionality of the strands. A medium groove is formed between two parallel strands with the same orientations of guanines (syn . syn or anti . anti). A wide or narrow groove is formed between two antiparallel strands with the opposite orientations of guanines (syn . anti) (**Figure 1.8**[41] and **1.9**[40]).
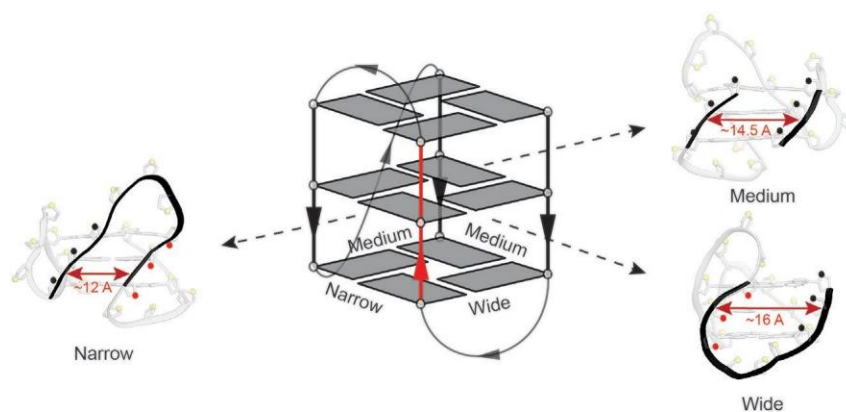
Figure 1.8. The cartoon of 3+1 hybrid G4 showing the three types of grooves. The groove dimensions are mentioned with red double-headed arrow[41].
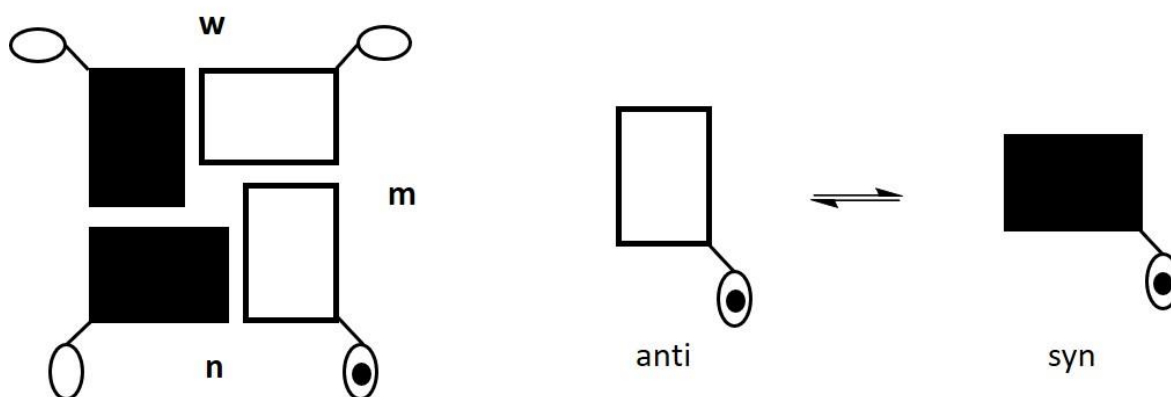
Figure 1.9. Syn and anti-glycosidic bond conformations of the guanine determine the groove width of G-quadruplexes[40].

G-quadruplexes have three kinds of loops known as the lateral loop, diagonal loop, and edge reversal loop (propeller loop). Loops have a remarkable effect on the stability of the G-quadruplexes[42]. Theoretically, a G-quadruplex can have three loops and three loop types, giving rise to 27 possible loop combinations. The sequence and number of nucleotides in the loops can affect the type of loop that can formed[40]. A propeller loop of a G-quadruplex with two or three tetrads commonly contains a single nucleotide. A single-nucleotide loop is also possible as a lateral loop, but it is not long enough to form a diagonal loop[40] (**Figure 1.10**).
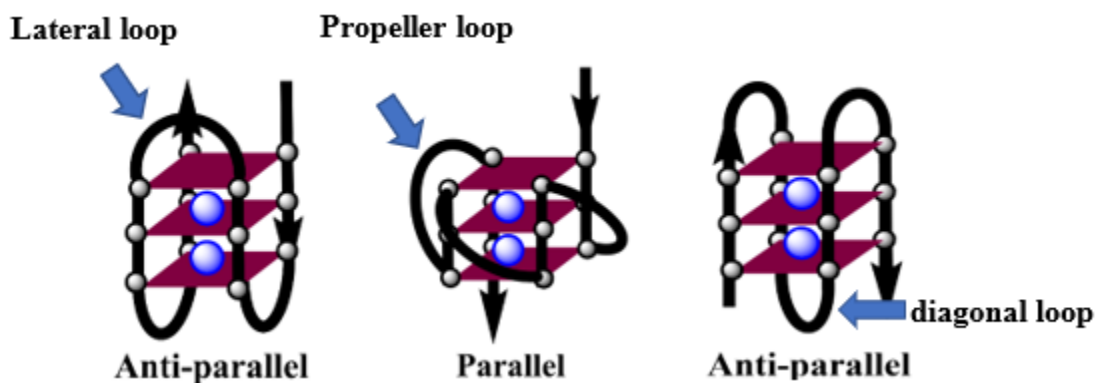
Figure 1.10. Common loop types in the G-quadruplexes.

The π-π base stacking interactions between G-tetrads play a major role in the stabilization of the G-quadruplexes. A G-tetrad can overlap with the bottom tetrad in three common ways[43] (**Figure 1.11**) based on the analysis of many G-quadruplexes in the Protein Data[40].
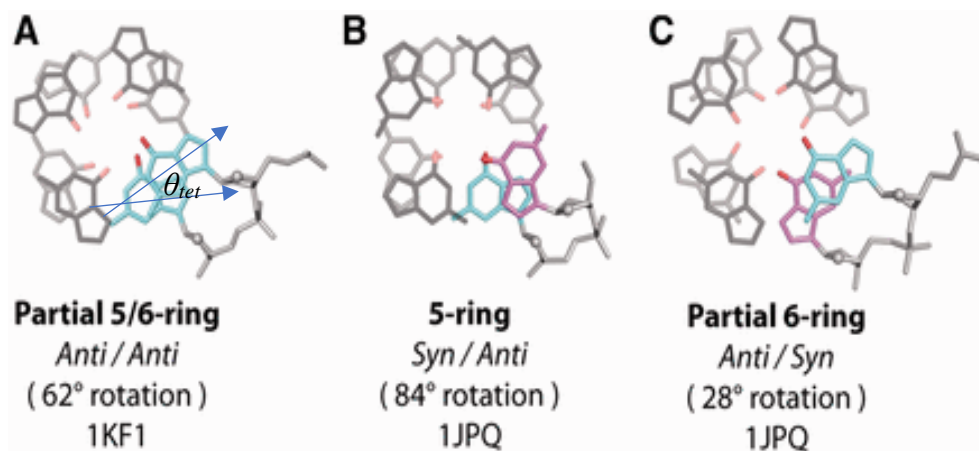
Figure 1.11. Different ways a G-tetrad can overlap with other G-tetrads. The blue lines are carved along the C8-H8 bond. The angle between them is relative rotation angle rotation (θ_{tet}) from 5' to 3'.

a). A partial 5/6 ring overlap is found in the Anti/Anti base step, in which the imidazole ring of one guanine overlaps with the pyrimidine ring of the other guanine.

b). A 5-ring overlap is found in the Syn/Anti base step, in which the imidazole ring of one guanine overlaps with the imidazole ring of the other guanine.

c). A partial 6-ring is found in the Anti/Syn base step, in which the pyrimidine of one guanine overlaps with the pyrimidine of the other guanine.

Importance and relative abundance of G-quadruplexes in the genome. The presence of G-quadruplexes in vitro led to the development of algorithms that can scan the genome for the putative G-quadruplex (pG4) forming sequences. The algorithms developed can search for the contiguous stretches of three or more guanines with intervening non-guanines such as $GGG_{3+}N_{1-7}GGG_{3+}N_{1-7}GGG_{3+}N_{1-7}GGG_{3+}N_{1-7}$. A recent study predicted approximately ~700,000 pG4 sequences present in the human genome[44] considering bulges and longer loop regions[45] in the G-quadruplexes.

Studies have shown that the pG4 sequences are mostly concentrated near or in the promoter regions of regulatory genes. These regions are involved in the signal transduction,

proliferation, cell-cell signaling, growth factor activity, regulation of cell cycles involving other than the normal housekeeping genes, and tumor suppressor genes[46]. The pG4 sequences are typically located in the downstream (transcription starting site +500 base pairs) of the genes with significant transcriptional activities (**Figure 1.12**). This phenomenon is conserved throughout the animal kingdom[29], implicating the evolutionary role of the G-quadruplexes in the regulation of gene expression[47,48].
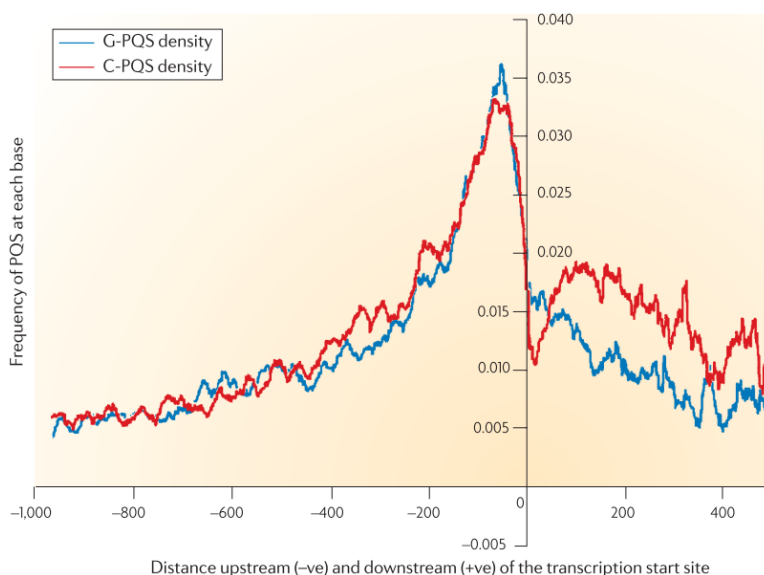
Figure 1.12. Relative abundance of potential quadruplex forming sequences (PQS) in the human genome near transcription start site[29].

Evidence of the role of G-quadruplexes in regulating gene expression also came from the following experimental data. The concentration of G-quadruplexes in cells is cell-cycle dependent. The highest concentration of G-quadruplexes was found in the S phase during DNA replication[48]. At the S-phase stage, DNA is relaxed from the super helical stress[29] allowing for winding into G-quadruplexes with a physiological concentration of K^+. G-quadruplex specific

helicases have also been discovered, which unfold G-quadruplexes during replication, transcription, and translation to prevent genomic instability[49]. Mutations in G-quadruplex specific helicases like BLM and WRN[50] cause the loss of their biological functions, leading to pathological conditions such as Bloom's and Werner's syndromes[51]. Further studies suggest these diseases are related to precocious aging, significant DNA damage, and a high propensity for cancer development. A recent finding of XPD and XPB G-quadruplex specific helicases also link their role in assisting gene transcription[52].

Telomeres structure and their function. The abundance of pG4 sequences in the telomeric regions and their proximity to the promoter regions of the proto-oncogenes has sparked considerable interest in G-quadruplexes as a novel class of targets for cancer remediation. The DNA sequences in the telomeric regions contain the highest GC content and have been found to fold into G-quadruplexes under physiological conditions as observed with G-quadruplex specific antibodies[44].

Human telomeres at the ends of chromosomes contain double-stranded 5'-AGGGTT repeats (10-15 kb long) and a single-stranded overhang[53] (50-500 nucleotides long) with the same sequence repeat at the 3'-end. This sequence motif is conserved throughout the eukaryotes. The 3' G-rich overhang is folded into the duplex through strand invasion to form a structure called T-loop[54], where it binds to the complementary C-rich strand, displacing the G-rich strand into a displacement loop (D loop). The size of T-loop varies from species to species, even from cell to cell. The formation of the T-loop (**Figure 1.13**) protects the chromosome ends, which would otherwise be recognized as double-strand breaks (DSBs) by DNA repair pathways. T-loop also prevents the chromosome ends from nonhomologous end-joining (NHEJ) and thus maintain the chromosome integrity.

The telomeric region and T-loop are stabilized by a shelterin complex comprising of six proteins, which associate specifically with the telomeric 5'-TTAGGG repeats.
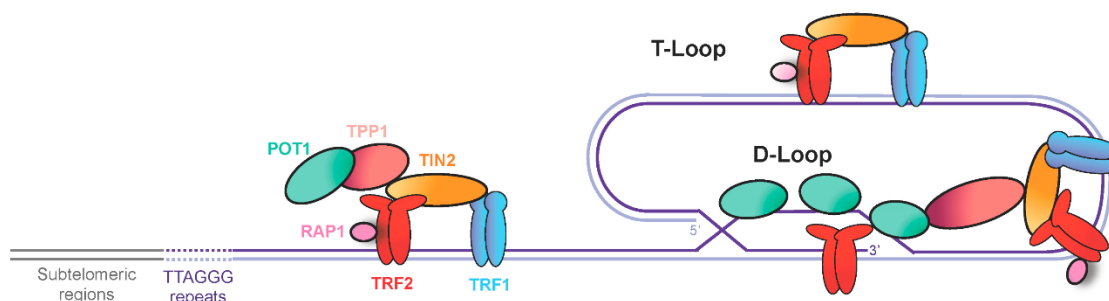
Figure 1.13. T-loop and its stabilization by shelterin complex[55].

The proteins TRF1and TRF2 bind the duplex region of the telomere in a sequence-specific manner. Studies show that TRF2 can promote the formation of T-loop like structures in telomeric DNA in vitro. POT1 recognizes the single-stranded overhang of TTAGGG repeats and the D-loop in the T-loop structure. TRF1 and TRF2 are abundant in the telomeres and can recruit other proteins to the telomeres. Rap1 is recruited to the telomeres by TRF2, but it does not directly bind to telomere DNA. The TIN2 protein is the central component of the shelterin complex, bridging TRF1, TRF2, and TPP1. It enables the complex to connect the components that bind to the single-stranded and double-stranded DNA of telomeres. TPP1 bridges POT1 to TIN2, which has one oligonucleotide/oligosaccharide-binding fold protein domain (OB) to interact with the enzyme telomerase. POT1 has two OB folds with which it recognizes ssDNA at the telomeres and helps maintain the length of the telomeres[56]. The shelterin complex at the telomeres plays a central role in maintaining the genomic integrity[56], telomere length maintenance[57], and replication regulation[55].

Under normal conditions, a cell loses a part of its DNA at the telomere during each replication cycle. This is called the end replication problem:

End replication problem. It is an intrinsic property of DNA replication machinery. DNA polymerase can only synthesize a DNA strand in the 5' to 3' direction. It requires a free hydroxyl group at the 3'-end of the existing RNA or DNA fragment to add additional nucleotides. DNA replication is semi-conservative. A daughter strand is synthesized on the template (parent) strand during the replication. One strand known as the leading strand can be synthesized continuously from 5' to 3' by the DNA polymerase. The other strand, known as the lagging strand, cannot be replicated continuously. It uses RNA primers to initiate the synthesis of Okazaki fragments (small 150-200 longs pieces of DNA). Later, these Okazaki fragments are connected with a combined activity of DNA endonuclease and DNA ligase. Because of this machinery, the lagging strand cannot be fully extended due to the lack of a template strand at the 3'-end of chromosome ends (3' single-stranded overhang) (**Figure 1.14**). The loss of the chromosome ends accumulates after each replication cycle. Such erosion of the telomere eventually reaches a critical point that triggers DNA damage responses through the shelterin complex, leading to cellular senescence and apoptosis[58].

Germ cells and cancer cells use the same replication machinery, but they are involved in literally unrestricted replication (immortality). How do they circumvent the end replication problem associated with telomere shortening? It has been found that these cells overly activate an enzyme called telomerase, a reverse transcriptase. Telomerase maintains the telomere length by adding the TTAGGG repeats to the 3' chromosome ends using its inherent RNA primer. Telomerase is overexpressed in 85-90% of the cancer cells, and its presence is negligible in somatic cells. This makes telomerase an attractive therapeutic target. In the early 1990s, it was

found that the G-quadruplexes are not the substrate of the telomerase[59]. Therefore, the induction

and stabilization of G-quadruplex at the telomeric regions can inhibit the binding of telomerase

and prohibit its ability to maintain the telomere length. Ultimately, this could trigger cancer

cells' programmed death[60].

To successfully targeting telomeric G-quadruplexes, one needs to differentiate G-

quadruplex and duplex DNA. Besides, as mentioned above, G-quadruplexes can adopt various

conformations depending on the DNA sequence and experimental conditions. Methods of

distinguishing G-quadruplexes with different conformations are ideal for potential therapeutic

applications. Using small molecules that specially bind to G-quadruplexes is a promising

approach to target G-quadruplexes. In the next section, I will discuss G-quadruplex binding

ligands and their crucial interactions and structural motifs.
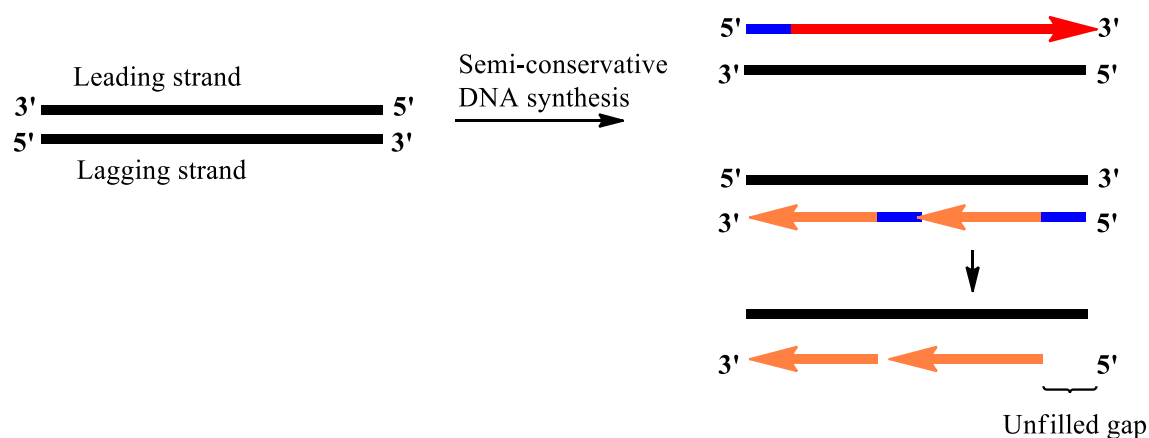
Figure 1.14. An illustration of end replication problem.

1.6. G-quadruplexes and Their Interactions With Small Molecules

G-quadruplexes are more prevalent than thought before in the gene promoters. Under physiological conditions, the G-quadruplex sequences can fold into different conformations that equilibrate between each other. How this interplay[37] between different conformations leads to the difference in gene expression is not fully understood. G-quadruplexes adopt different topologies; however, there is one inherent similarity between different conformations - the omnipresence of planar G-tetrads. The planarity of G-tetrads allows ligands with an extended aromatic surface to stack, a key factor to consider when designing G-quadruplex binding ligands. Small molecules that stack with G-tetrads increase the thermal stabilization of the G-quadruplex. However, such a stacking mode often makes it difficult for molecules to bind selectively to G-quadruplexes with different topologies.

In the following sections, different G-quadruplexes and their specific ligands are reviewed. Common G-quadruplexes with different conformations are shown in Table 1.2. Many of them are found in the promoter regions of many oncogenes[29]. They are related to the regulation of oncogene expression[61]. Most G-quadruplexes under physiological conditions adopt parallel conformation except BCL-2, which adopts a hybrid conformation. HRAS1 G-quadruplex adopts antiparallel conformation.

Table 1.2

G-quadruplexes and Conformation Adopted by Them.

Sr. No	Gene/region	Conformation	PDB ID	Reference
1	c-MYC	Parallel	1XAV	62
2	c-KIT 1	Parallel	2O3M	63
3	c-KIT 2	Parallel	2KYP	64
4	BCL-2	Hybrid	2F8U	65,66
5	HRAS1	Antiparallel		67,68
6	htelo	Hybrid	2GKU	69

1.7. Ligands Targeting Parallel Stranded G4

1.7.1. c-MYC G-quadruplex

The expression of c-MYC in cells is difficult to measure due to the short half-life of c-MYC mRNA and low concentrations of c-MYC proteins. Additionally, c-MYC proteins have a very short half-life in proliferating cells (~20 mins[70]), and the proteins readily undergo ubiquitylation and degradation by proteases[71]. c-MYC proto-oncogene encodes a transcriptional factor that activates or suppresses a number of other genes. c-MYC also interacts with thousands of other genes[72] through its protein controlling every aspect of cellular functions (**Figure 1.15**)[70].
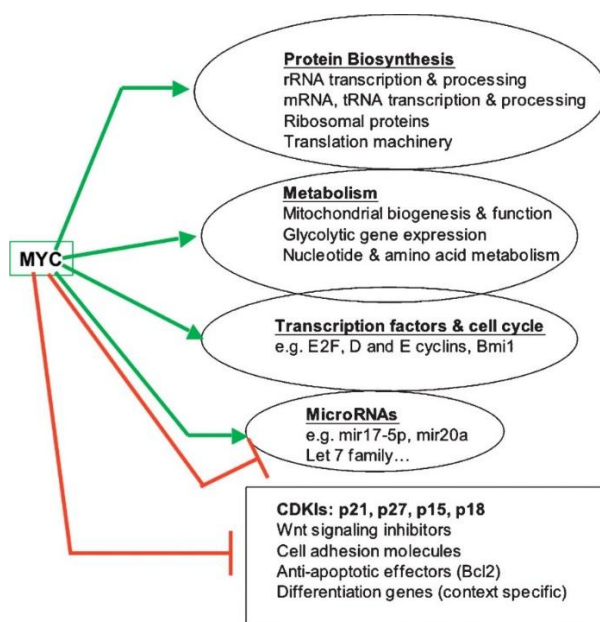
Figure 1.15. Genes whose expression is affected by overexpression of c-MYC[70].

c-MYC is a master regulator for cell growth regulation, increased metabolism[73], and proliferation of cancer cells[74]. Overexpression of c-MYC gene is observed in nearly 80% of the

solid tumors of breast, gastrointestinal, and ovarian cancer tumors, making it an attractive

therapeutic target for cancer remediation.

The transcription activity of c-MYC is controlled by multiple promoters such as P1 and

P2 as the predominant ones (**Figure 1.16**). Located at 142-115 bp upstream of the P1 promoter,

NHE III$_1$ is a highly conserved and guanine-rich element, responsible for 80-95% c-MYC

transcription. The 27 mer G-rich sequence (pu27) is shown in **figure 1.15**, which contains five

consecutive runs of guanines (3 or more guanines in each run). The pu27 sequence can fold into

a G-quadruplex with a parallel conformation in K$^+$ under physiological conditions[75]. DNA foot

printing[61] and NMR[62] showed that the G-quadruplex is composed of the 3' II, III, IV, and V runs

of guanines in pu27 (**Figure 1.16**). The formation of this G-quadruplex in NHE III$_1$ promotes

the downregulation of the c-MYC expression[76] and has been correlated with antitumor activity.

The c-MYC G-quadruplex formed in vivo can be unwound by theNM23H2, a non-metastasis

family of protein. The resulting single-stranded DNA is stabilized by a nucleic acid-binding

protein (CNBP) and a CCHH zinc-type zinc finger to facilitate further transcription[29].
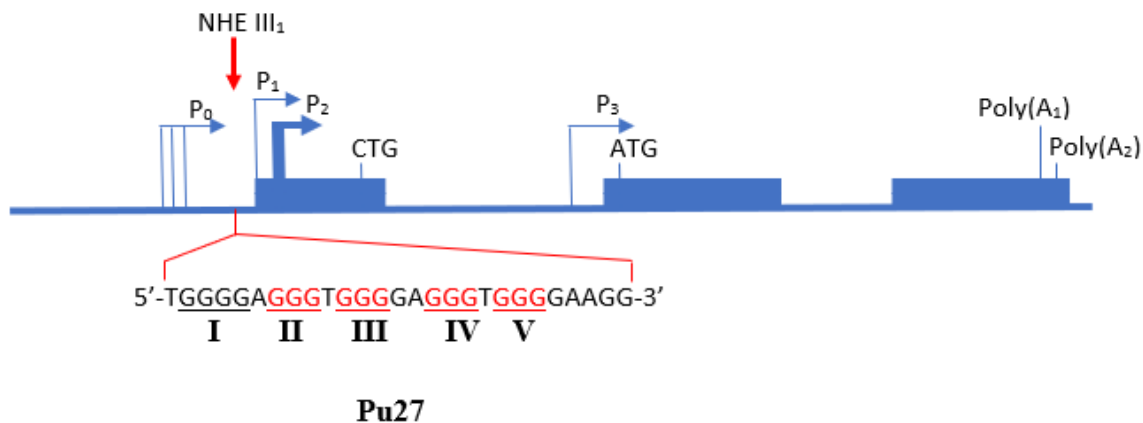
Figure 1.16. Promoter structure of the c-myc gene with pu27 in the NHE III1 of promoter P1[36].
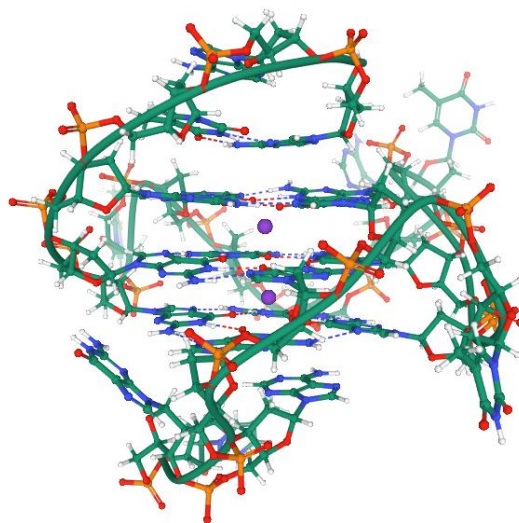
Figure 1.17. NMR structure of the c-MYC G-quadruplex (PDB:1XAV).

Stabilization of the parallel G-quadruplex (**Figure 1.17**) formed in the NHE III$_1$ can prevent the NM23 from binding and unwinding of the G-quadruplex and thus silence the expression of the c-MYC gene[76]. Over the years, several classes of G-quadruplex binding ligands have been developed to interact with the c-MYC G-quadruplex.
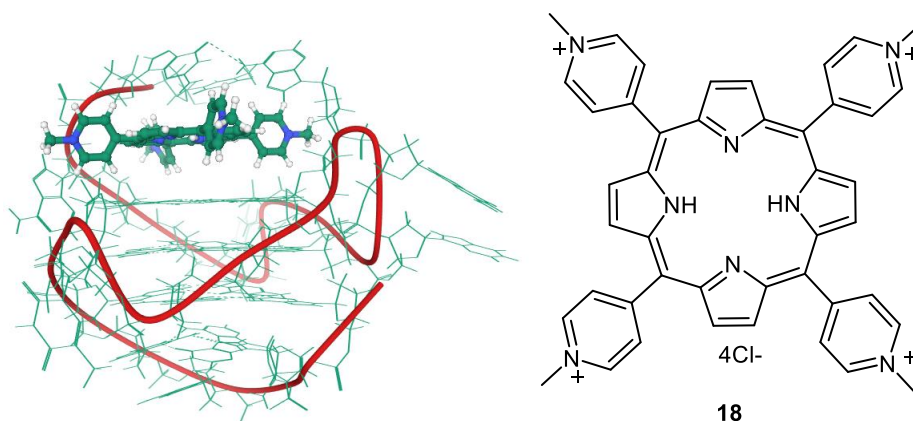
Figure 1.18. TMPyP4 bound by end stacking mode to the c-MYC G-quadruplex (NMR solution structure, PDB ID 2A5R[77]).

TMPyP4 (**18**) (**Figure 1.18**) consists of a planar porphyrin ring with 4-methyl substituted cationic pyridinium rings as side chains. NMR studies showed that TMPyP4 end stacks on the terminal G-tetrad of the c-MYC G-quadruplex[78] and side chains interact with the loops[79]. TMPyP4 has a moderate selectivity for intermolecular quadruplex[80] over duplex DNA[81,82]. Its binding affinity to G-quadruplex DNA is 0.25 ± 0.1 and 1.9 ± 0.2 µM for 3' and 5'-tetrad of 23mer htelo human telomeric G-quadruplex[83].

The strong binding of TMPyP4 to c-MYC G-quadruplex DNA comes from its shape complementarity with the G-tetrad. Its large aromatic surface stacks with the G-tetrad via pi-pi interactions. Four positive charges of TMPyP4 are responsible for its diminished selectivity for G-quadruplex over duplex DNA. It has been demonstrated experimentally by Ruan *et al* [84] that a TMPyP4 analog with only three positive charges has improved selectivity over duplex DNA.

Quindoline (**19**) derivatives have also been extensively studied for interactions with c-MYC G-quadruplex DNA. Quindolines are derived from the parent compound cryptolepine (**20**) (**Figure 1.19**). Cryptolepine is a natural product from Cryptolepis sanguinolenta and has been used as a traditional Ghanian folk medicine for the treatment of a variety of ailments. It has also been used as a vasodilator[85] and antihyperglycemic medication[86].

Quindoline
19
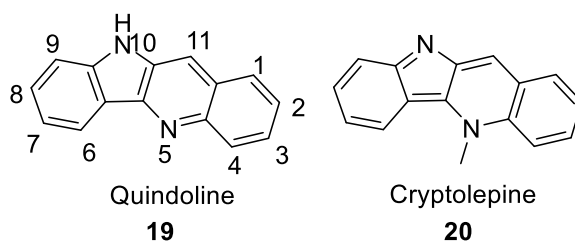
Cryptolepine
20

Figure 1.19. Cryptolepine (left) and Quindoline (right).

Ou *et al*[87] synthesized a series of quindoline derivatives to target c-MYC G-quadruplex and its downregulation in vivo (**Figure 1.20**). The rationale of the design is to optimize the electronic properties of quindoline. The electron density of the nitrogen at position 5 is increased by adding an electron-donating substituent at position 11. Hence, the pKa values of the NH group at the 5-position of quindoline derivatives indeed increase from 4.2 to ~8.2[88]. Under physiological conditions, this nitrogen could be readily protonated, enhancing its interaction with the electron-rich center of the G-quadruplex. Compound **29** with an N-diethyl group as an end moiety at position 11 showed a mild selectivity for parallel c-MYC G-quadruplex over a hybrid G-quadruplex DNA as determined by competition dialysis assay. It stabilized c-MYC G-quadruplex (wild type Pu27 5µM) by over 25 °C at 5 µM concentration[87].
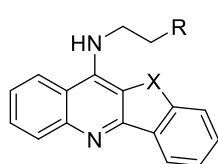
21 X =NH, R = OH	**22** X =O, R = OH
23 X =NH, R = CH$_2$OH	**24** X =O, R = CH$_2$OH
25 X =NH, R = N(CH$_3$)$_2$	**26** X =O, R = N(CH$_3$)$_2$
27 X =NH, R = CH$_2$N(CH$_3$)$_2$	**28** X =O, R = CH$_2$N(CH$_3$)$_2$
29 X =NH, R = N(CH$_2$CH$_3$)$_2$	**30** X =O, R = N(CH$_2$CH$_3$)$_2$

Figure 1.20. 11-substituted Quindoline derivatives.

The G-tetrads at the 3' and 5' ends are different depending on their flanking bases[89]. The effect of flanking bases on ligand binding was studied on a "truncated c-MYC G-quadruplex sequence" (Pu22) using solution NMR spectroscopy (**Figure 1.21**). The DNA sequences used in the study are shown below. Pu27 is the native sequence of c-MYC. TGA (-3, -2 and -1 to the first guanine involved in G-tetrad formation, labeled in red) are the flanking bases at the 5'-end of Pu22, and TAA (+1, +2, and +3 to the adjacent guanine) are the flanking bases at the 3' end of Pu22.

```
              1 2 3  4  5 6  7  8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27
Pu27     5'-TGGGGAGGGTGGGGAGGGTGGGGAAGG-3'
Pu22            TGAGGGTGGGTAGGGTGGGTAA-3'
```

Figure 1.21. Sequence of the wild type c-MYC DNA Pu27 and the truncated version Pu22.

Compound **29** forms a 2:1 ligand to G-quadruplex complex. Its binding to the 5'-end of the c-MYC G-quadruplex leads to a dramatic rearrangement of the flanking bases (5'-TGA). The adenine (-1A, A6) is recruited by Compound **29** to form a new plane at the top of 5'-G-tetrad, which is completely by wrapped a guanine residue (-2G, G5) (**Figure 1.20**). Binding of the ligand at the 3' end also showed the involvement of the flanking bases (5'-TAA). Compound **29** recruits the thymine base (+1T, T23) to form a new plane at the 3' end of the G-quadruplex, which is capped by the adenine residue (+2A, A24) (**Figure 1.22**).
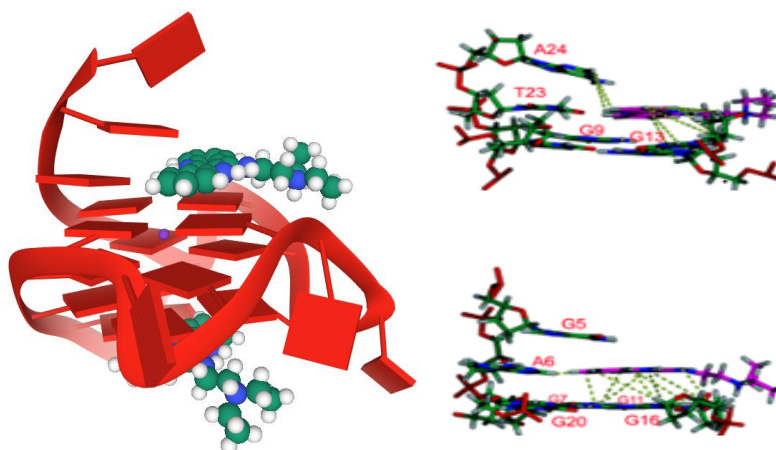
Figure 1.22. The ribbon structure of compound **29** bound to the c-MYC G-quadruplex (PDB:2L&V)(left) and c-MYC bound to the terminal G-tetrad at 5'-position (bottom-right) and at the 3'end (top right) involving the flanking sequences.

Compound **29** is protonated at pH 7 (pK$_a$, ~8.3). NOE constrained MD calculations

showed the formation of hydrogen bonding between the O4 of thymine (T23) and the N-H of

Compound **29** (N1H) (**Figure 1.23** (right)). The presence of hydrogen bonding interactions is

supported by the increased thermal stabilization of G-quadruplex at low salt concentrations. On

the other hand, binding at the 5' end is not affected by the salt concentration as dramatically as

that at the 3' end. The 5'-end is more hydrophobic; therefore, stacking interactions at this end

have more effect on the binding. Mutational analysis showed that guanine (G5) that form the

capping structure at the top of quindoline-i and adenine plane at the 5' end is more critical for the

binding. The overlapping area at the 5'-end is larger than that at the 3'-end, also suggesting the

importance of hydrophobic stacking interactions (**Figure 1.23**).
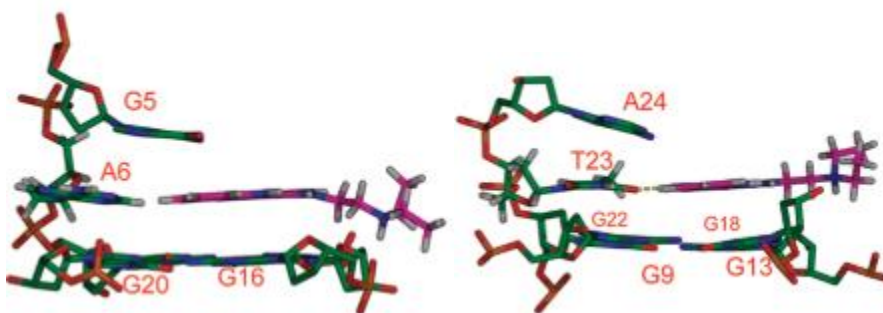
Figure 1.23. Compound **29** overlapped at the 5' G-tetrad and 3'G-tetrad of the c-MYC G-quadruplex.

To improve the binding selectivity of quindoline derivatives for c-MYC G-quadruplex,

Funke *et al*[90] designed and synthesized phenyl-substituted indolo[3,2-b]quinolines. The

compound **31** (**Figure 1.24**) with a larger hydrophobic stacking area showed a remarkable

stabilization effect (ΔT$_m$ is over 35 ^{o}C) on c-MYC G-quadruplex. This compound also showed

an order of magnitude of selectivity for c-MYC G-quadruplex over human telomeric (htelo) G-

quadruplex. The NMR titrations showed an increased preference for binding to the 5' G-tetrad than the 3' G-tetrad. The better stacking interaction of **31** with the 5' G-tetrad is believed to be the driving force.[90]
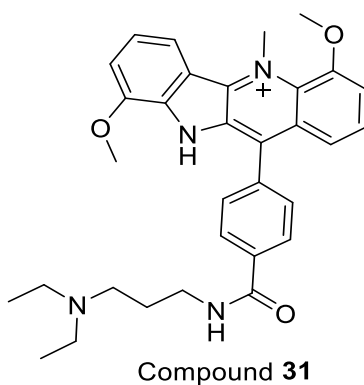
Compound **31**

Figure 1.24. Structure of 11-phenyl substituted indolo[3,2-b]quinoline (**31**).

Trisubstituted 11-piperazinylquindoline derivatives (**Figure 1.25**) were developed for tight and selective binding to c-MYC G-quadruplex. These compounds contain three substituents on the quindoline ring. As an example, incorporation of a N, N-dimethyl ethyl linker at R_1 in compound **32** increased its selectivity for c-MYC G-quadruplex over duplex DNA and other parallel G-quadruplexes. Thermal denaturation monitored by CD showed more than 16.5 °C increase in the T_m of c-MYC G-quadruplex when incubating compound **32** and DNA with a 1:1 molar ratio. Unfortunately, the ability to stabilize c-MYC G-quadruplex by these trisubstituted derivatives is not correlated well with their antitumor activity against HCT-116 (colon) and Raji (Lymphoma) cell lines. A CA46 exon-specific test suggested that the stabilization of c-MYC G-quadruplex is not the sole reason for the inhibition of cancer cell growth by these derivatives[91]. It is noteworthy that the stabilization of DNA might not always

have a positive correlation with the inhibition of c-MYC expression. Similar results were

observed with the famous G-quadruplex binding ligand Quarfloxin[29].

11-piperazinylquindoline derivatives   Compound **32**
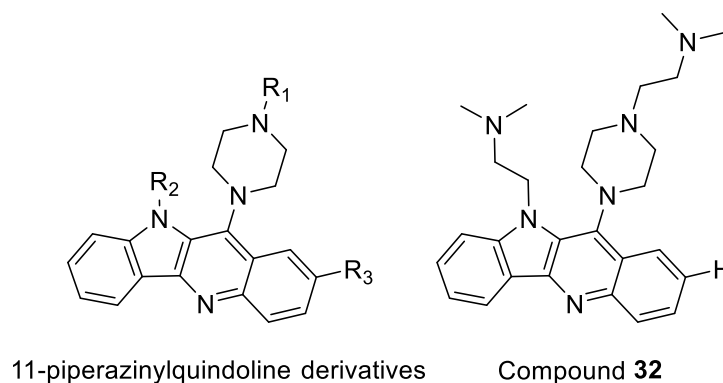
Figure 1.25. The general structure of trisubstituted Quindoline derivatives (left) and compound **32** (right).

Ming-Hao Hu and his co-workers reported the development of tetrasubstituted ligands

with excellent binding affinity and selectivity over other parallel stranded G-quadruplexes such

as BCL-2, c-KIT 1, and KRAS. A representative compound IZCZ-3 (**33**), a 1-methylpiperazine

substituted diaryl imidazole and carbazole conjugate (**Figure 1.26**), was able to differentiate

between hybrid-type telomeric G-quadruplex DNA and antiparallel HRAS G-quadruplex DNA.
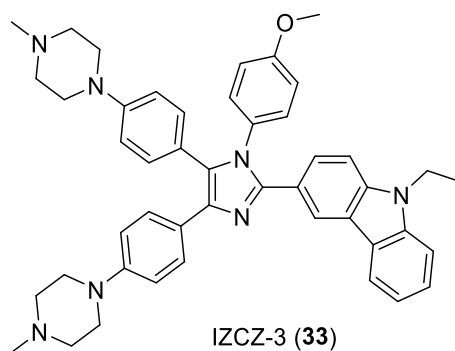
IZCZ-3 (**33**)

Figure 1.26. The structure of tetrasubstituted clover leaf-like molecule IZCZ-3 (**33**).

The compound **33** inhibited cell growth and induced apoptosis by reducing the level of mRNA and c-MYC proteins. The CA46 exon assay indicated that the observed biological effect of the compound is correlated with its stabilization of c-MYC G-quadruplex[91]. This G-quadruplex interacting compound showed comparable activity (effectiveness) with no non-specific toxicity as compared to the other non-specific DNA binding drug Doxorubicin[92].

The binding modes of compound **33** with G-quadruplex were determined by docking studies (**Figure 1.27**). Results suggested that compound **33** stacks closely onto the terminal G-tetrads of c-Myc G-quadruplex. The positively charged imidazole ring of compound **33** sits on top of the central cavity of the G-tetrad and electrostatically interacts with it. The 1-methylpiperazinyl side chain makes close contacts with the edgewise reversal loop. In addition, the flanking bases of c-MYC G-quadruplex are intrinsically flexible, which accommodates a better ligand fit. As a comparison, in the binding of this compound with htelo G-quadruplex, the lateral loop of the G-quadruplex can hinder the close contact between the ligand and G-tetrad, leading to the relatively higher binding energy. The docking data mentioned here can be used to explain the better binding affinity and selectivity of this ligand towards c-MYC G-quadruplex.[89]
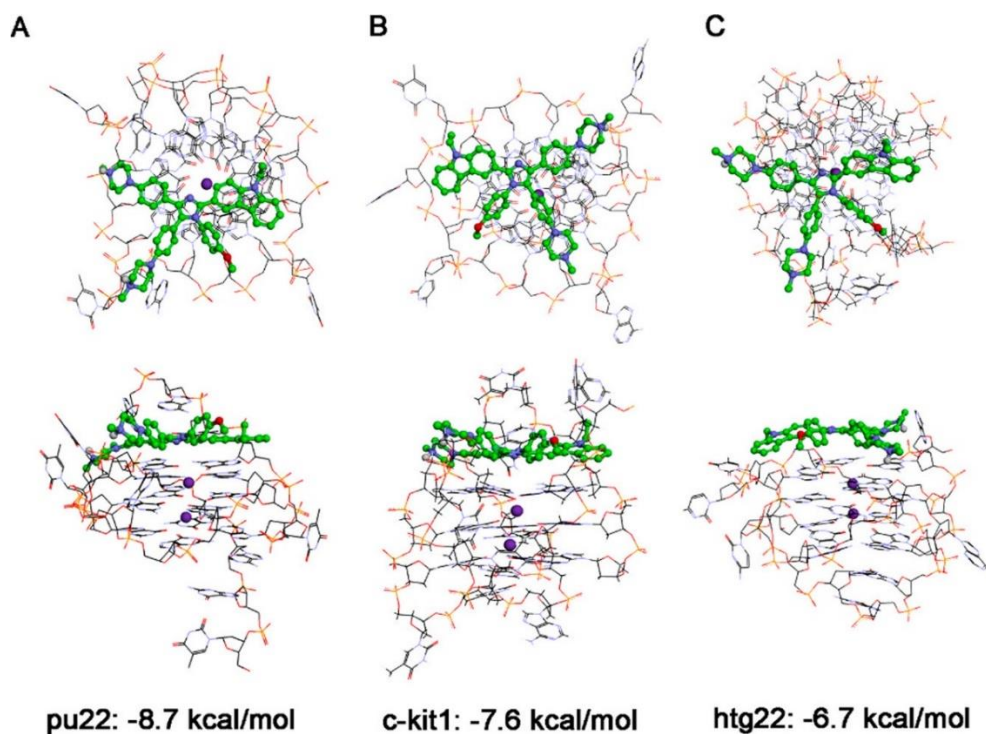
Figure 1.27. Top view (upper panel) and side view (lower panel) of the binding models of compound **33** with (A) the c-MYC G-quadruplex pu22 (PDB ID: 2L7V), (B) the c-KIT G-quadruplex c-KIT 1 (PDB ID: 4WO3), and (C) the telomeric G-quadruplex htg22 (PDB ID: 2MB3).

Other ligands with moderate affinities and relatively better selectivity to c-MYC G-quadruplex DNA include 11-triazole benzofuruquinoline[93], bis-triazolylcarbazole[94], some natural products like Fonsecin B[95], Berberine derivatives[96] and metal complexes like Iridium (III)[97] and Platinum (II) complexes[98] (**Figure 1.28**).
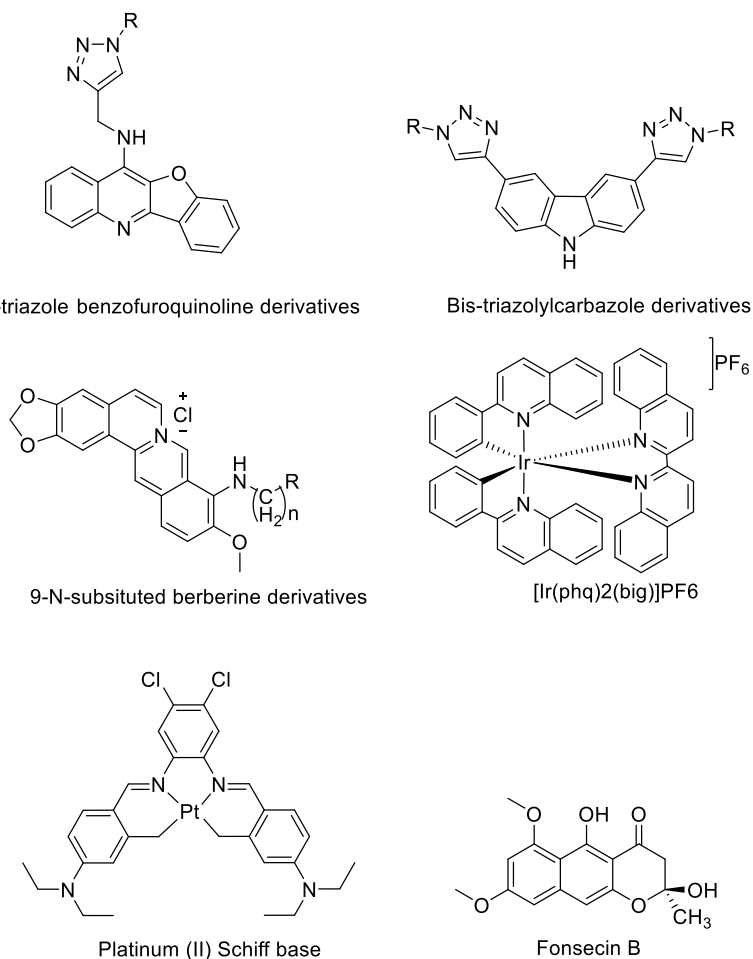
Figure 1.28. Different classes of c-MYC G-quadruplex ligands.

1.7.2. c-KIT G-quadruplex

The c-KIT gene contains codes for a transmembrane protein product. This protein acts as a tyrosine kinase receptor that helps to relay signals from the cell surface to the interior of the cell. Binding of this c-KIT receptor with Stem cell factor (SCF) induces its dimerization and promotes transphosphorylation. Transphosphorylation triggers a cascade of downstream reactions affecting hematopoietic cellular differentiation, proliferation, and development. Mutations in the c-KIT gene produce proteins that have defects in the juxtamembrane region and kinase domain. Mutant proteins cause SCF independent autophosphorylation and support cell

proliferation. An abnormal level of c-KIT expression is correlated with small cell lung carcinona, melanoma, and more than 80% cases of gastrointestinal stromal tumors[99] (GISTS). In recent years, c-KIT pathways have become promising targets for cancer remediation. An FDA approved drug, Imatinib[14], has been used as an effective treatment against c-KIT positive GISTs by inhibiting the tyrosine kinase activity of c-KIT.

Two potential G-quadruplex forming sequences Kit 1 and Kit 2 are present in the promoter regions of the c-KIT gene. c-KIT 1 extends from -87 to -109 base pairs, and c-KIT 2 extends from -140 to -160 base pairs located upstream of the transcription initiation site. The region containing c-KIT 2 was found to be a core regulatory element of transcription[100]. Interestingly, this region is highly conserved in the human, rat, mouse, and chimpanzee.

c-KIT G-quadruplexes adopt a parallel G-quadruplex conformation under physiological conditions[63,101–103]. The structure of c-KIT 1 G-quadruplex is unusual. A two-nucleotide lateral loop snaps back another five nucleotide loop, forming a large pocket between the bottom G-tetrad and lateral loop (**Figure 1.29**, blue arrow)[103]. Solution NMR studies[63] showed the formation of base pairing between adenine-1 and thymine-12 at the top of 5'-G tetrad and a Hoogsteen interaction between adenine and guanine in the five-nucleotide lateral loop. Such interactions contribute to the stability of the c-KIT 1 G-quadruplex.
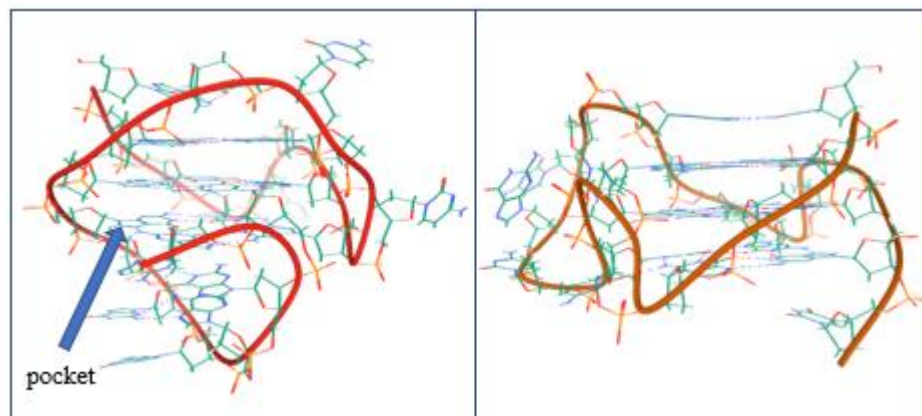
Figure 1.29. NMR solution structure of c-KIT 1 (left) and c-KIT 2 solution structure (right).

The c-KIT 2 sequence also adopts a parallel topology containing two one-nucleotide double chain reversal loops and one five-nucleotide double chain reversal loop. A non-canonical base pairing between cytidine-1 and adenine-13 at the top of guanine-14/guanine-18G segment of the 5' G-tetrad was found.[64]

To search for specific ligands that bind to c-KIT G-quadruplexes, Rocca *et al*[104] used a computational method to screen a library of 462,045 molecules against two c-KIT G-quadruplexes (PDB ID: AWO2 and AWO3). All molecules were filtered following Lipinski's rule of five and selected for favorable ADME properties. Three compounds (**Figure 1.30**) were chosen, all of which can bind to both c-KIT 1 and c-MYC G-quadruplexes. According to the dock studies, these compounds are groove binders. Circular dichroism studies showed a modest decrease of the absorption band at 265 nm with increasing ligand concentrations, indicating the lack of π-π stacking interactions. The primary binding site of the ligands to c-KIT 1 G-quadruplex was in the cleft (**Figure 1.29**, blue arrow). According to the docking data, the piperidine residue in all the best-fitted structures is oriented towards the negatively charged phosphate backbone. Compound **36** forms a salt bridge between the G4 nucleobases and an

adenine residue (A19). None of the compounds showed the ability to stabilize duplex DNA.

This study showed a new method to target c-KIT 1 by utilizing phosphate backbones bordering

the pocket (**Figure 1.29**. blue arrow). The ligands bound to the c-KIT 1 G-quadruplex are shown

in **figure 1.31**.
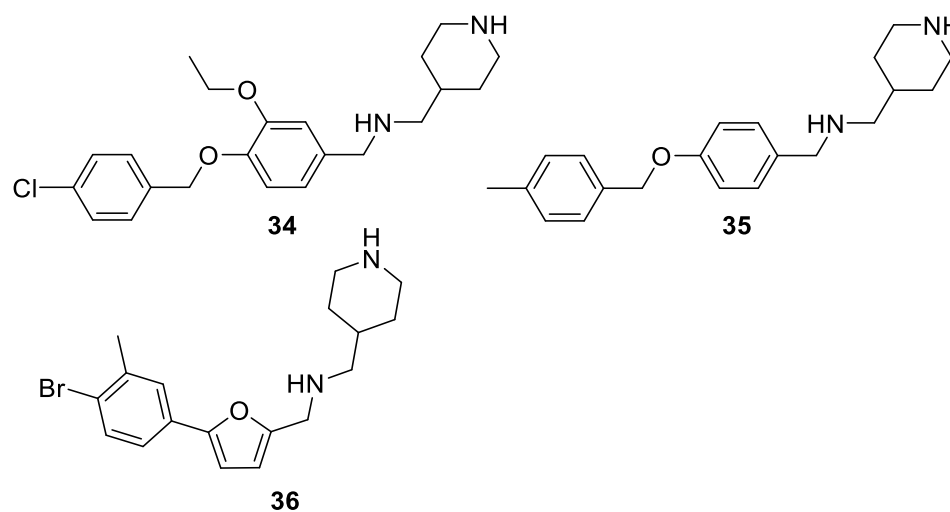
Figure 1.30. Structure of dual c-KIT and c-MYC binders.
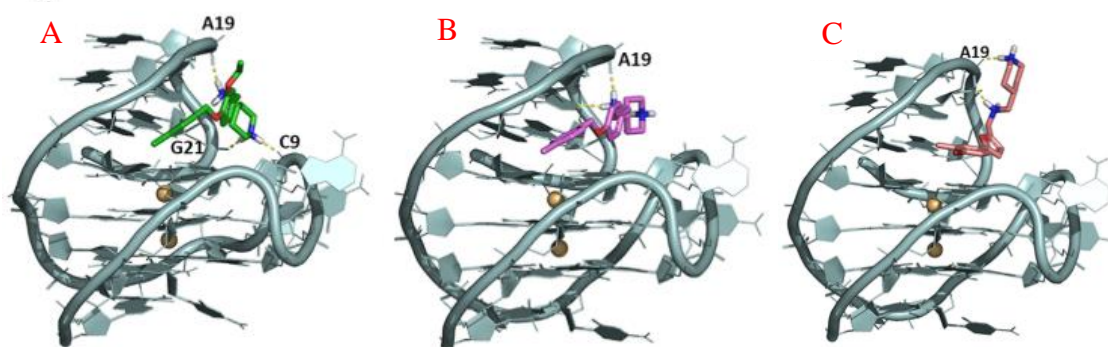
Figure 1.31. Representations of the compounds (A) **34**, (B) **35** and (C) **36** docked to c-KIT 1 G-quadruplex[104].

In a separate study, Zorzan *et al*[105] screened an in-house library of compounds based on

five common DNA binding motifs: Anthraquinone (AQ), anthracene (AN), phenanthroline

(Phen), naphthalene diimide (NDI), and heterocyclic diamidines (HAD). Anthraquinone- and

anthracene-based ligands showed the best thermal stabilization and selectivity for the c-KIT 1

and c-KIT 2 G-quadruplex over telomeric G-quadruplex and duplex DNA. The FID assay

revealed three potential binding ligands (**37**-**39**), as shown in **figure 1.32**. Out of the three, AQ1

(**37**) was the most potent ligand that binds to c-KIT 1 and c-KIT 2 G-quadruplexes.

Table 1.3
*FID $^{G4}DC_{50}$ and K_d Values of the Hits (**37-39**) Against c-KIT 1 and c-KIT 2.*

Ligand	FID $^{G4}DC_{50}$		K_d (μM)	
	c-KIT 1	c-KIT 2	c-KIT 1	c-KIT 2
AQ1 (**37**)	0.32 ± 0.05	0.35 ± 0.05	1.99 ± 0.15	1.01 ± 0.15
AQ7 (**38**)	0.24 ± 0.04	0.25 ± 0.02	3.04 ± 0.43	2.29 ± 0.28
AN6 (**39**)	4.11 ± 0.70	3.63 ± 1.00	71.5 ± 20.1	25.5 ± 4.01

Compound **18** showed low cytotoxicity against c-KIT positive cell lines HGC27 (human

gastric carcinoma cell line) and MCF7 (Michigan cancer foundation-7, a breast cancer cell line)

even though it is a potent c-KIT G-quadruplex ligand based on the in vitro studies such as

polymerase stop assay.

Compound **39** did not affect the mRNA level in HGC27 cells but showed a weaker effect

in MCF7 cells. By contrast, Compound **17** reduced the c-KIT expression by 2 to 10 folds in

HGC27 cells and 37 folds in MCF7 cells. Compound **37** showed anti-proliferative activity

against imatinib-resistant cell lines (HMC1.2: Human mast cells), suggesting the potential of this

compound to be effective against any c-KIT mutation in the protein-coding region. The

cytotoxicity of compound **37** was much less against c-KIT deficient cells, implicating that **37**

interacts with the c-KIT G-quadruplex. It is noteworthy that **37** is the best G-quadruplex binding

ligand developed for downregulation of the c-KIT expression.
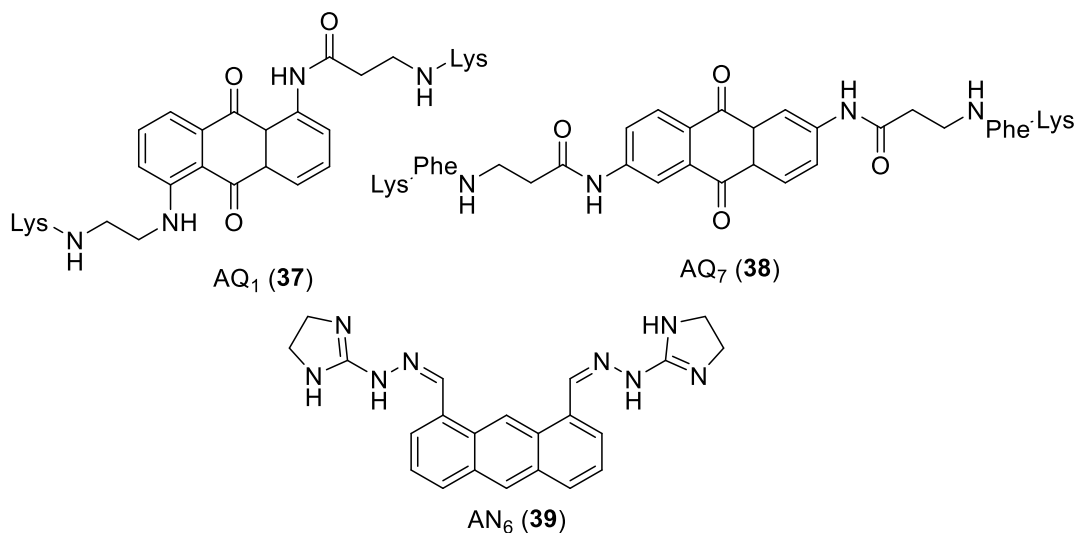
Figure 1.32. Structure of Anthraquinone derivatives AQ1 (**37**) (top left), AQ7 (**38**) (right) and
Anthracene derivative AN6 (**39**) (bottom). Lys = lysine.

All the ligands mentioned above have limited selectivity for c-KIT G-quadruplex over c-

Myc G-quadruplex DNA because they have parallel topologies. In fact, till today, many G-

quadruplex binding ligands can effectively differentiate between G-quadruplex and duplex DNA

and distinguish G-quadruplexes with different conformations (parallel, antiparallel, and hybrid).

However, reports on targeting G-quadruplexes with parallel conformations (e.g., c-KIT G-

quadruplex and c-Myc G-quadruplex DNA) are scarce. Bejugam *et al* [106] synthesized a series of

trisubstituted isoalloxazines derivatives for selective recognition of c-KIT G-quadruplex DNA.

Impressively, several compounds (**Figure 1.33**) were able to differentiate between the c-KIT 1

and c-KIT 2 parallel G-quadruplexes and a telomeric hybrid G-quadruplex DNA. Compound **40**

was the most potent compound with a K_d value of 3 µM and 6.4 µM for c-KIT 2 and c-KIT 1,

respectively. By contrast, the K_d values of compound **43** for c-KIT 2 and c-KIT 1 were 9.7 µM and 31 µM, respectively. The K_d value of compound **40** binding to the hybrid telomeric G-quadruplex was 43 µM, and no meaningful K_d value of compound **43** for telomeric G-quadruplex could be detected. Compound **43** is the only ligand containing a fluorine atom at the R_2 and showed the most selective response against c-KIT 1 and 2 over the telomeric hybrid G-quadruplex. Results also suggested that extra cationic side chains can lead to non-specific binding interactions, which may lead to loss of selectivity of ligand to different G-quadruplex species[107]. Moreover, the presence of fluorine atom would increase the hydrophobicity of the compound, which may interact better with non-polar G-tetrad better. More detailed binding process can be investigated using rate of association and dissociation in SPR studies. It is noteworthy that compound **43** showed the better downregulation of c-KIT expression (31 fold in 6h) in the MCF-7 cell lines.
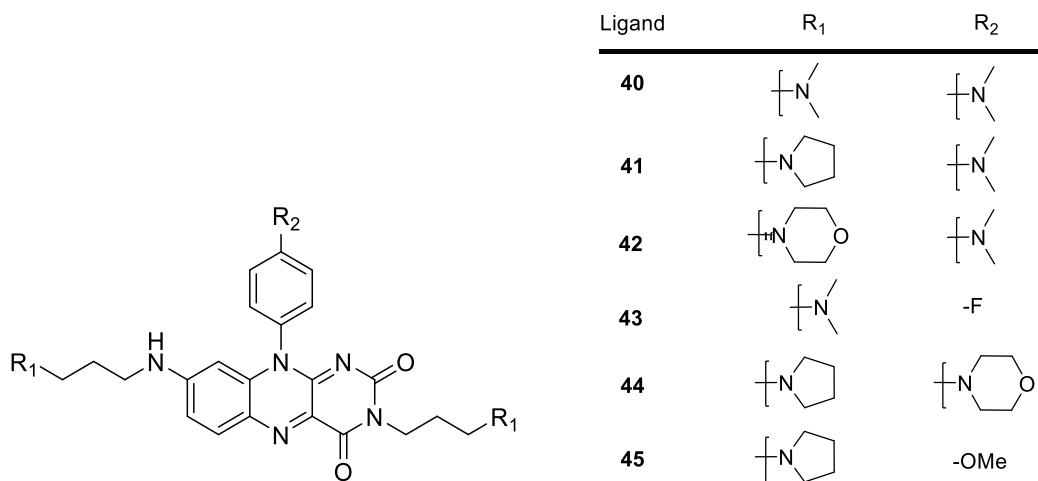
Figure 1.33. Structures of isoalloxazine ligands **40-45**.

The same group also developed a new c-KIT G-quadruplex binding scaffold,

Benzo[a]phenoxazines[108] (**Figure 1.34**). Both ligands showed greater affinities for c-KIT 2 over

c-KIT 1. In particular, compound **46** binds to c-KIT 2 9.5 times more selectively over c-KIT 1,

and compound **47** binds to c-KIT 2 7.5 times more selectively over c-KIT 1. Both compounds

bind to c-KIT G-quadruplexes more preferentially (about 2-5 times) than c-MYC and telomeric

G-quadruplexes. In addition, both compounds bind more selectively to G-quadruplexes than

duplex DNA. The K_d value of compound **46** binding to G-quadruplex is 10 fold higher than

duplex DNA, and the K_d value of compound **47** to G-quadruplex is more than 50 fold lower than

duplex DNA, as determined using surface plasma resonance (SPR).
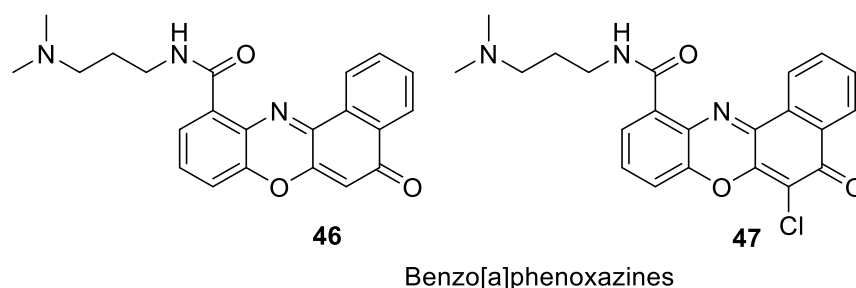
Figure 1.34. Structure of benzo[a]phenoxazines ligand **46** (left) and **47** (right).

Other ligands that are developed for recognition of parallel c-KIT G-quadruplexes over

hybrid conformation of telomeric G-quadruplex include triarylpyridines derivatives[109],

indolylmethyleneindanone derivatives[110], oxazole peptide macrocycles[111] shown in **Figure 1.35**.

Triaryl pyridine derivatives

Indolyl methyleneindanone derivatives

1	R	Stereochemistry
a	$(CH_2)_4NH_2$	S,S,S
b	$(CH_2)_4NH_2$	R,R,R
c	$(CH_2)_4NH_2$	S,S,R
d	CH_2NH_2	S,S,S

Oxazole based peptide macrocycles
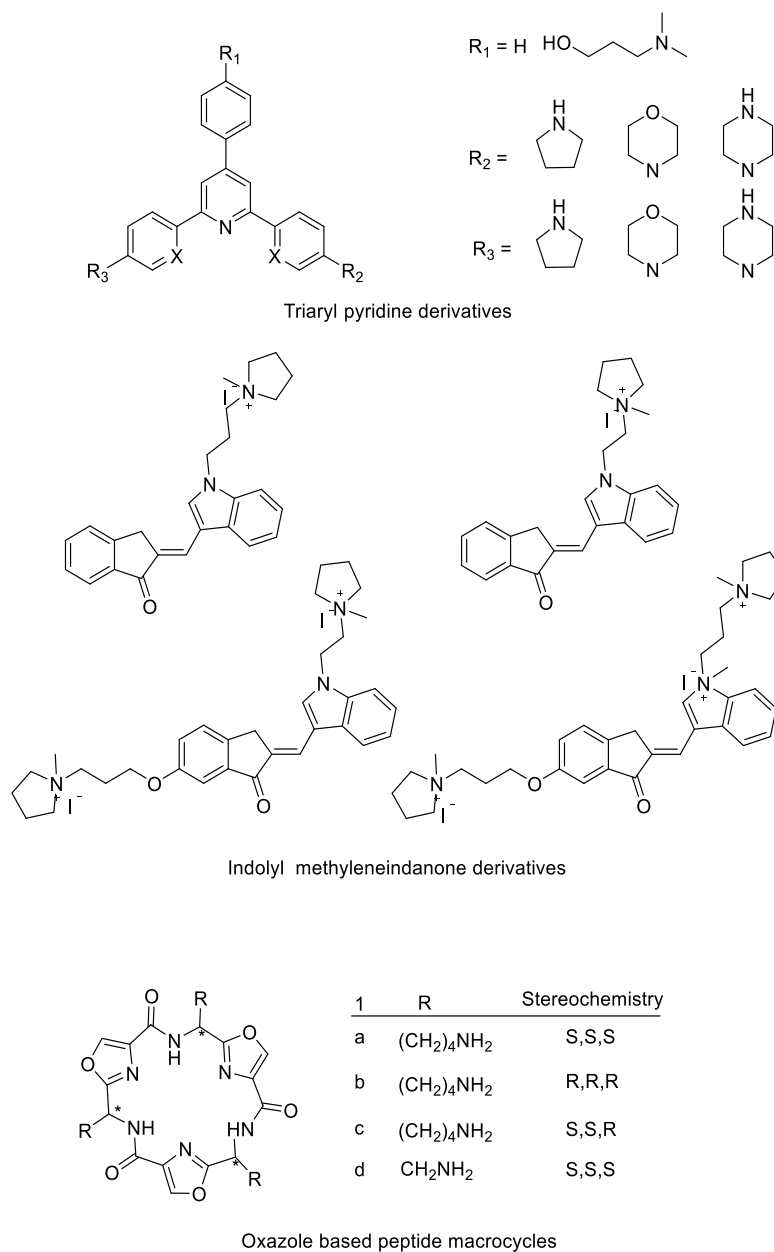
Figure 1.35. Scaffold of various other classes of ligands developed for c-KIT G-quadruplex.

1.8. Ligands Targeting Antiparallel G-quadruplex

1.8.1. HRAS/KRAS

The H-ras, N-ras, and K-ras were the first human oncogenes discovered more than three decades ago, belonging to the members of the Ras GTPase family. They are crucial in relaying

signals downstream from the tyrosine kinase membrane receptors to a cascade of parallel

phosphorylation reactions that culminate with nuclear transcription factors. This signal

transduction is involved in normal cell cycle progression, growth, proliferation, apoptosis, and

senescence. Mutations in ras genes have been associated with various developmental disorders

and cancer[112]. Ras oncogene mutations are primarily concentrated within 2 hot spots, codons

12,13 and codon 61, resulting in 1 in 19 possible point mutations for each RAS isoform[113]. Val,

ASP, CYS are the dominant mutant alleles of the K-ras gene. These mutations lock the protein

in the active conformation. Currently, the Sanger Center keeps and updates the nature and

frequency of ras mutations in human tumors. About 30% of all human tumors screened carry

some form of mutations in ras genes, and most of these mutations are detected in K-ras oncogene

(25-30% of all the tumors screened)[114], along with its involvement in 95% of pancreatic tumors,

50% colon tumors, and 30% of lung adenocarcinomas[112]. The rates of oncogenic mutations in

N-ras and H-ras family members are much less prevalent (8% and 3% of samples screened,

respectively).

So far, RAS genes are undruggable at the protein level due to the tissue-specific nature of

the up/stream signal transduction pathways. Interestingly, HRAS gene has two G-rich sequences

HRAS1 (-434 to -462) and HRAS2 (-509 to -530) locate at the upstream of the transcription start

site of the HRAS[68]. HRAS1 was found to adopt an antiparallel conformation in solution[68,115].

On the other hand, HRAS2 adopts parallel G-quadruplex conformation. Several classes of

compounds (**Figure 1.36**) have been developed to target HRAS G-quadruplexes.

Anthrathiophenedione-guanidino derivative (**48**, ATPD-1) (**Figure 1.36**) can bind to both

parallel and antiparallel HRAS G-quadruplexes and consequently downregulate the expression of

the gene[68].

In another report, tetrakis-(diisopropyl-guanidine) phthalocyanine[116] (**49**, DIGP) (**Figure 1.36**) was shown to stabilize the antiparallel HRAS1 G-quadruplex and strongly halt the formation of full-length product in a polymerase stop assay. A similar preference for antiparallel telomeric G-quadruplex was found for this class of compounds by the Balasubramanian group[117].

Liu *et al*[118] synthesized a series of para-substituted bis-(vinylquinolinium) benzene derivatives (**50**) (**Figure 1.36**), some of which can increase the T_m of HRAS1 G-quadruplex by more than 9.3 °C. However, the compounds showed low selectivity for HRAS1 over duplex DNA (~factor of 2) and limited selectivity toward antiparallel G-quadruplexes. The lack of selectivity might result from their linear shape, which can fit into the duplex grooves with a slight twist of the molecules. On the other side, its preference for antiparallel G-quadruplexes could result from the interactions with the lateral loop residues.

Other quinolinium-based ligands that can selectively bind to HRAS1 G-quadruplex over hybrid and parallel G-quadruplexes include a fluorescent triphenylamine-quinolinium[119] probe (**51**) and a series of benzothiazole-substituted benzofuroquinolinium (**52**) compounds[120] (**Figure 1.36**). These compounds also show good selectivity over dsDNA and ssDNA.

Different G-quadruplex topologies have different groove characteristics. So far, there are not many ligands[121] that can bind into the grooves of G-quadruplexes specifically. A carbazole based G-quadruplex groove binding ligand (**53**) for HRAS1 antiparallel G-quadruplex was developed[122] (**Figure 1.36**). The lead compound showed 6 to 20 times selectivity over other G-quadruplex topologies.

ATPD-1 (**48**)

DIGP (**49**)

p_BQ-OHEtP (**50**)

Triphenylamine-quinolinium (**51**)

Benzothiazole-subsituted benzofuroquinoliniium (**52**)

Substituted styrene-carbazole derivatives (**53**)
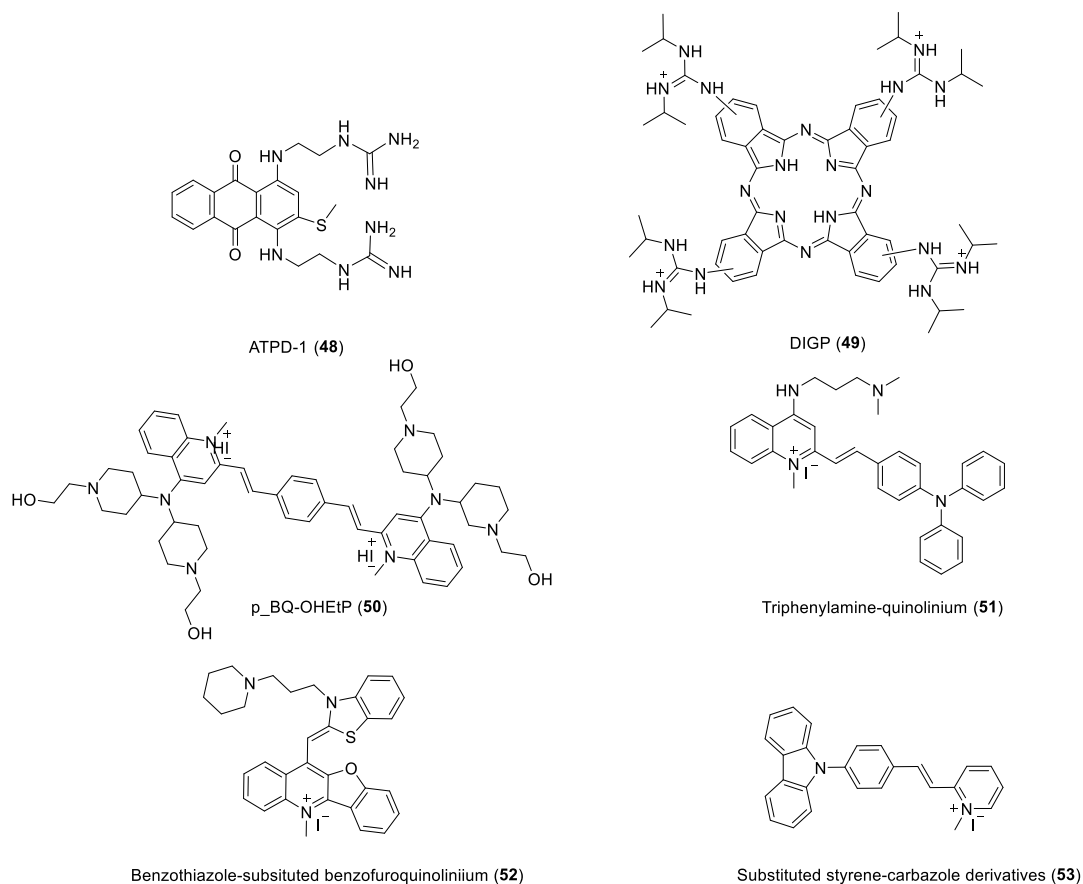
Figure 1.36. Classes of ligands that target HRAS1 antiparallel G-quadruplex.

1.9. Ligands Targeting Hybrid Conformation of G-quadruplex

1.9.1. BCL-2 G-quadruplex

The BCL-2 gene codes for mitochondrial membrane proteins that facilitate the oncogenesis by avoiding programmed cell death. BCL-2 gene expression is often significantly enhanced in pancreatic cancer in addition to the activation of other oncogenes [28], which has led to inadequate responses to chemotherapy and drug resistance in the treatment of human pancreatic cancers.

So far, three G-rich sequences upstream of the BCL-2 transcription site have been identified[45,65,66,123], which are Pu 32 [-1875 to -1906 base pairs from the transcription start site

(TSS)], Pu39 (-1451 to -1489 base pairs upstream of the TSS) and P1G4 (-1412 to -1439 bp upstream of the TSS) (**Figure 1.37**). Pu32 can adopt a hybrid G-quadruplex conformation as observed from spectroscopic data. However, the detailed structural elucidation and its relevance to BCL-2 biological functions have not been elucidated. Pu39 can adopt multiple conformations of G-quadruplexes, depending on the sequence tracts used for the G-quadruplex formation. The most stable G-quadruplex of Pu39, confirmed with DMS DNA footprinting and NMR studies, is a parallel G-quadruplex conformation[45], with an unconventional 13 nucleotide propeller loop. A loop containing 13 nucleotides is unusual. Its discovery broadens our knowledge on G-quadruplex-folding in solution. Additionally, Pu39 can also adopt a hybrid conformation using four G-tracts in the middle of the sequence with a 7-nucleotide lateral loop. The latter (PDB ID: midG4)[65] is the key conformation used for research on BCL-2 G-quadruplex. The guanine-rich P1G4 sequence was recently discovered, which is located immediately above the BCL-2 P1 promoter. The P1G4 G-quadruplex shows a significant effect on the transcriptional regulation of the BCL-2 gene. It adopts two hairpin-like G-quadruplex conformations with two edge reversal loops comprising of 12 nucleotides and 11 nucleotides, respectively. Both G-quadruplexes coexist in equilibrium under physiological conditions.
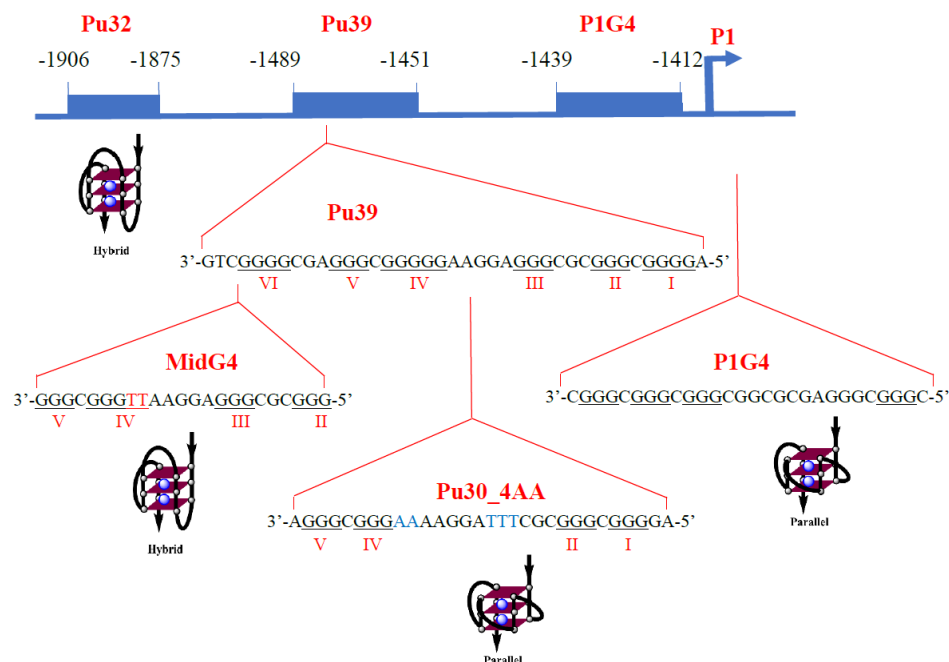
Figure 1.37. G-quadruplexes in the promoter P1 of the BCL-2 gene. The guanines used in the formation of G-quadruplex are underlined. The substitutions in the original loop sequence are colored in blue.
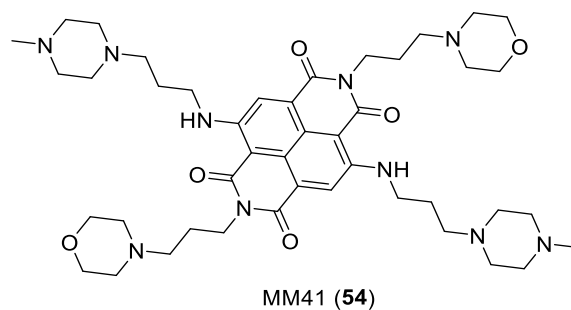
Figure 1.38. Structure of Naphthalene diimide derivative **MM41**[28].

A G-quadruplex binding agent MM41 (**54**) (**Figure 1.38**) has been used for downregulating the activity of BCL-2 along with the KRAS gene. Thermal denaturation monitored by fluorescence showed that **54** selectively stabilizes BCL-2 G-quadruplex and two KRAS G-quadruplexes by more than 20 °C. Molecular dynamics simulations showed **54** stacks

with the terminal G-tetrad through π-π interactions, and four side chains extend into the grooves. Guanine-12 of the loop sits at the top of the ligand, forming a sandwich-like structure[28] (**Figure 1.39**). **54** exhibits antitumor activity in the MIA PaCa-2 pancreatic cancer xenograft model. Interestingly, it did not affect the telomerase activity (TRAP assay) and expression of the c-MYC gene, suggesting that **54** has a limited stabilization effect on telomeric G-quadruplex and c-MYC G-quadruplex.
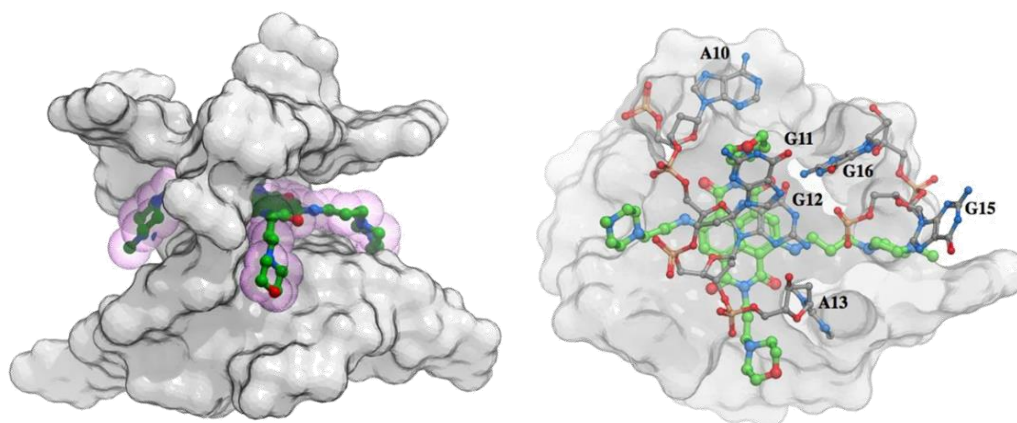
Figure 1.39. A snapshot from MD simulations. Top view of the compound **54** docked onto the terminal tetrad and interaction with G12 (left), Side chains of ligands interacting with the grooves of the BCL-2 G-quadruplex[28].

In another study, Quindoline (SYUIQ-01[87] (**55**) and SYUIQ-F05[124] (**56**)) and cryptolepine derivatives (SYUIQ-FM05[125] (**57**)) were used to downregulate the expression of BCL-2 gene (**Figure 1.40**). The binding of BCL-2 G-quadruplexes by three compounds was measured using SPR spectroscopy (Table 1.4). Results suggested that the rates of association of the ligands to the G-quadruplex and duplex are roughly the same. But the rates of dissociation of the ligands from the G-quadruplex and duplex are drastically different, especially for compounds **57** and **56**. Hence, the low K_D (k_d / k_a) values of these two ligands came from their relatively low

rate of dissociation from the G-quadruplex. **57** is a more potent G-quadruplex binding ligand,

with an impressive k_d value of $3.18 \times 10^{-8}\, M^{-1}$. Due to similar molecular structures, the

hydrophobic interactions of compounds **56** and **57** with the BCL-2 G-quadruplex DNA is

roughly the same. The major difference is the quaternary nitrogen in compound **57**. This

positively charged atom sits above the G-tetrad center and interacts with four oxygens (O6) of

guanines via electrostatic attractions (acts as a pseudo potassium ion). Such ionic interactions

could decrease the observed dissociation constant of compound **57**. It is necessary to point out

that the positive charge(s) of a ligand must be placed at a specific location in G-quadruplex DNA

to achieve binding specificity. Otherwise, non-specific interactions of a positively charged

ligand with DNA could occur, reducing binding selectivity. Ligand **57** was also found to reduce

the production of mRNA and induce apoptosis in HL-60 (human leukemia-60) cells[126].
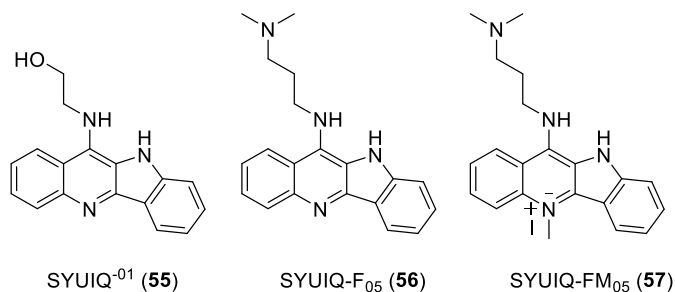
Figure 1.40. Structure of Quinoline derivatives SYUIQ-01 (**55**) (left), SYUIQ-F05 (**56**) (middle) and Cryptolepine derivative SYUIQ-FM05 (**57**) (right).

Table 1.4
Kinetic Parameters of the Ligands (57-59) Determined by SPR Spectroscopy.

	SYUIQ-FM05 (57)		SYUIQ-F05 (56)		SYUIQ-01 (55)	
	BCL-2	duplex	BCL-2	duplex	BCL-2	duplex
k_a (M^{-1} s^{-1})	2.29×10^5	2.23×10^5	1.20×10^5	1.66×10^5	8.65×10^4	5.92×10^3
k_d (s^{-1})	7.28×10^{-3}	0.06	0.02	0.08	0.13	0.02
K_D (M)	3.18×10^{-8}	2.69×10^{-7}	1.67×10^{-7}	4.82×10^{-7}	1.50×10^{-6}	3.38×10^{-6}

k_a is association constant, k_d is dissociation constant, K_D (kd/ka) is equilibrium dissociation constant

Giusiniano[127] *et al* synthesized a series of ligands for selectively targeting BCL-2 and c-MYC G-quadruplex using nucleobases as synthons. Three nucleobase-oriented scaffolds are shown in **Figure 1.41**.
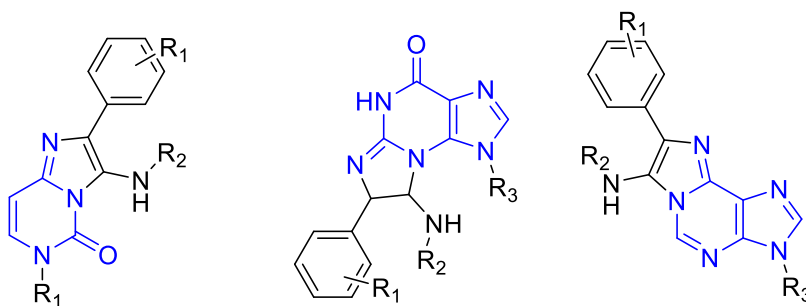
58 R1 = 2-ethyl morpholino	**59** R1= 4-chloro	**60** R1= N,N'-dimethyl
R2 = tertiary butyl	R2 = tertiary butyl	R2 = 2-ethyl morpholino
R3 = N,N'-dmethyl	R3 = 2-ethyl morpholino	R3 = 2-ethyl morpholino

Figure 1.41. Structure of cytosine, guanine and adenine-based scaffolds for G4 stabilization.

An initial survey of compounds **58**, **59,** and **60** showed that **58** and **59** were ineffective in stabilizing any G-quadruplex conformations. Compound **60** did show thermal stabilization of the c-MYC, BCL-2, and human telomeric G-quadruplexes. Based on the scaffold of compound

60, a library of 16 compounds was synthesized. Compound **61 (Figure 1.42)** was the most

potent and selective ligand for BCL-2 (hybrid) and c-MYC (parallel) G-quadruplex topologies in

the series.
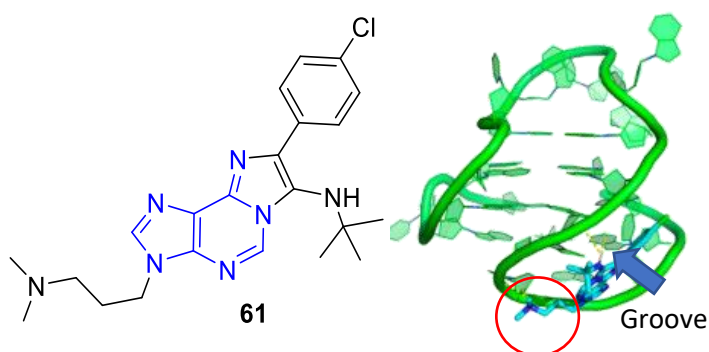
Figure 1.42. Compound **61** (left) Compound **61** docked on BCL-2 G-quadruplex, the electrostatic interaction is circled in red. PDB:2F8U (right).

Compound **61** showed selective inhibition of BCL-2 positive Jurkat cell growth (human

leukemia cell line) without having any effect on the cell viability of melanoma cells, MCF-7

cells, HCT-116 cells, and non-cancerous cells.

An NMR titration study showed that the binding of compound **61** only causes slight

perturbation of the guanines in the 3' G-tetrad and has no effect on the guanines in the 5' G-

tetrad. By contrast, docking studies showed that the compound interacts with the N-H of

guanines pointing in the grooves (top and middle tetrad) in the grooves. The amino side chains

of the ligand interact with the phosphate backbones (Circled in red in **Figure 1.42**).

1.9.2. Human Telomeric G-quadruplex

The ends of the chromosome consistently shorten after each replication cycle, known as

the end replication problem. When a cell's chromosome ends reach a critical point, it will trigger

cellular senescence. Theoretically, each cell can have around 50-70 replication cycles. On the contrary, cancer cells are immortal, having unrestricted replication cycles. Cancer cells overcome the end replication problem by overexpressing a reverse transcriptase, human telomerase. Telomerase extends the chromosome ends by adding 5'-TTAGGG repeats and thus maintain their length. It requires a single-stranded DNA (3'-G rich overhang) to pair with its embedded mRNA template to execute its function. Telomerase is overexpressed in 85-95% of cancer cells, making it a promising anticancer target. As discussed before, the guanine-rich telomeric region can fold into G-quadruplexes in vivo. Because the G-quadruplexes are not the native substrate of telomerase; therefore, the formation of G-quadruplex in the telomeric region can prevent the binding of telomerase and thus inhibit its biological activity. The use of small molecules to induce telomeric G-quadruplex is an effective approach for the inhibition of telomerase activity.

The first evidence of ligand-mediated G-quadruplex stabilization responsible for telomerase inhibition was reported in 1997. A G-quadruplex binding ligand 2,6-diamidoanthraquinone (**62**) (**Figure 1.46**) thermally stabilizes telomeric G-quadruplex and inhibits telomerase activity with an IC_{50} value of 23 μM.

Until today, the design of G-quadruplex binding ligands mainly focuses on two factors, aromatic surface and charged side chains. Ligands containing extended and fused aromatic rings can maximize π-π interactions with the planar nature of G-tetrads. Ligands with positively charged moieties can enhance their binding with and negatively charged G-quadruplex grooves and electron-rich G-tetrads. Porphyrin derivatives[79,128] (**63**) and PIPER[129] (**64**) are examples of G-quadruplex ligands with extended aromatic rings and charged side chains (**Figure 1.46**). When designing G-quadruplex ligands, one needs to consider the ligand's selectivity for G-

quadruplex over duplex DNA. The charged side chains are also a favorable factor for duplex binding ligands. Several synthetic and natural G-quadruplex binding ligands are listed below.

Neidle *et al*[130] synthesized a series of disubstituted and trisubstituted acridine molecules. The side chains were judiciously placed to interact with the grooves to maximize the binding affinity. The derivatives with three side chains showed a much weaker dissociation constant in the SPR (surface plasmon resonance) study. The acridine ring nitrogen aligns along the channel of negative electrostatic potential above the top G-tetrad (**Figure 1.43**). The compound showed 40 times selectivity over duplex DNA. One of the trisubstituted derivatives, BRACO-19 (**65**), showed the inhibition of telomerase activity with an impressive $^{tel}IC_{50}$ value of 95 nM. It inhibits the telomerase-mediated telomere extension in vivo by displacing hTERT (catalytic subunit of telomerase) from telomeres. In addition, it colocalizes with ubiquitin in the cytoplasm for proteasomal destruction[131]. Cancer cells treated with BRACO-19 showed an increased incidence of chromosome fusion[132], a hallmark of genetic instability.
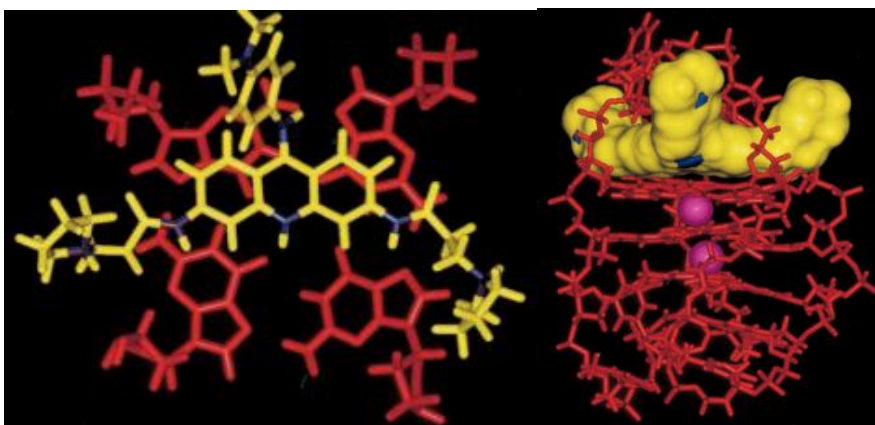
Figure 1.43. BRACO-19 (yellow color) bound to htelo G-quadruplex[130] (red color).

Pyridostatin (PDS, **66**) is another potent ligand for telomerase inhibition by displacing POT1 from telomeres[54]. This displacement triggers the DNA damage response in cells. The appearance of nuclear foci of phosphorylated histone gH2AX (an early damage response marker) has been observed in the cells treated with PDS[26]. Although PDS does not have a fused aromatic ring structure, its flexible structure can make intramolecular hydrogen bonds to result in a planar structure for binding (**Figure 1.46**). The side chains containing amine groups have electrostatic interactions with negatively charged phosphate backbones[133].

Phen-DC3 (**67**), a 1,10-phenanthroline analog, is one of the most potent telomerase inhibitors[134] via the G-Quadruplex approach[135]. It induces genetic instability in the yeast through G-quadruplex stabilization[27]. Phen-DC3 induces the conformation change of the hybrid G-quadruplex to two-tetrad antiparallel G-quadruplex or hybrid-3[136], leading to the unfavorable entropy change during binding[137]. The primary mode of binding of Phen-DC3 is stacking interactions at the outer tetrad (**Figure 1.44**) (image shown is for c-MYC G-quadruplex).
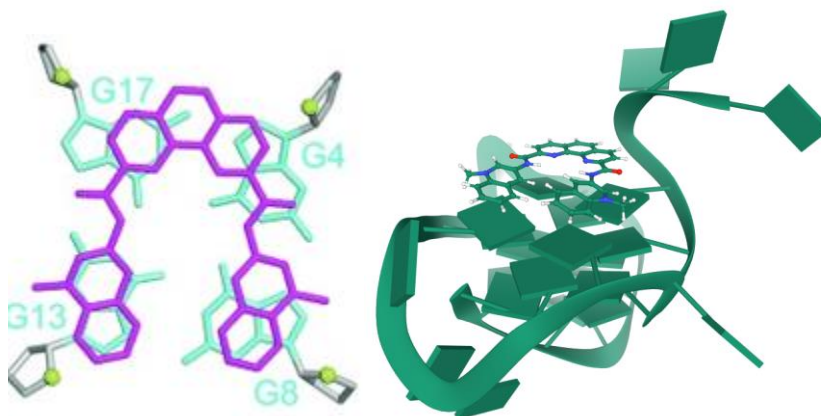
Figure 1.44. Phen-DC3 (**67**) stacked on the G-tetrad (left), Phen-DC3 bound to the G-quadruplex[138] (PDB:2MGN).

Substituted Naphthalene diimides are another class of ligands extensively developed to target telomeric G-quadruplexes. The ligands have a fused aromatic core and preferentially bind through end-stacking mode. One of the advantages of these ligands is to decorate the naphthalene diimide core with up to 4 basic/hydrogen bonding moieties stabilizing ranges of conformation in solution depending upon the nature of side chains as determined from the CD spectra, for example, a parallel conformation with N-methyl piperazine[139] (**68**), antiparallel with morpholine and N-methyl piperazine (2+2) side chains[140].

The basic side chains have a pronounced effect with the one of the most potent analogues so far can thermally stabilize telomeric G-quadruplex for over 28.3 °C (**Figure 1.46**). The $^{tel}IC_{50}$ of the compound is 50 nM. It has shown a potent POT1 in displacement activity, which act as the DNA damage response. The compound showed potent anticancer against pancreatic cancer cell lines both in vitro[139] and in vivo[141].

The compound is positively charged under physiological conditions and has a suggested binding mode similar to the MM41[142] (**54**). It prefers to binds to the parallel stranded conformation htelo conformation, which is the most common binding mode for this class of ligands[139,142,143] except a few[140].

Telomestatin (**69**), a natural product from the bacterium Streptomyces annulatus 3533-SV4[144], is one of the most potent G-quadruplex ligands, inhibits telomerase at 0.64nM using traditional TRAP assay (58 ± 17nM, according to direct primer assay)[134]. The compound has a macrocyclic structure, which is in perfect congruity with the outer G-tetrad (**Figure 1.45**). The compound has over 70-fold selectivity over the duplex DNA[80]. It inhibits the activity of telomerase at 0.64nM using traditional TRAP assay. The telomestatin derivative (**70**, **L2H**) showed binding to the 5'-G tetrad of the (3+1) human telomeric hybrid conformation[145]. The

compound stabilizes the G-quadruplex through the pi-pi interaction and side chains maintain

close contacts with the nucleobases in the backbone of G-quadruplexes (**Figure1.45**).
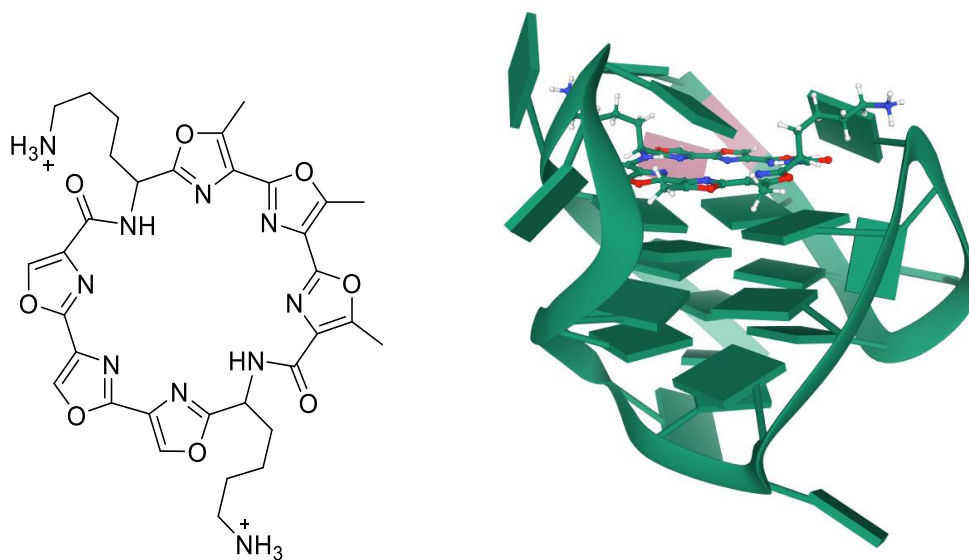
Figure 1.45. Telomestatin analogue L2H2-6M(2)OTD (**70**) (left), and its bound state to (3+1) htelo hybrid G-quadruplex.
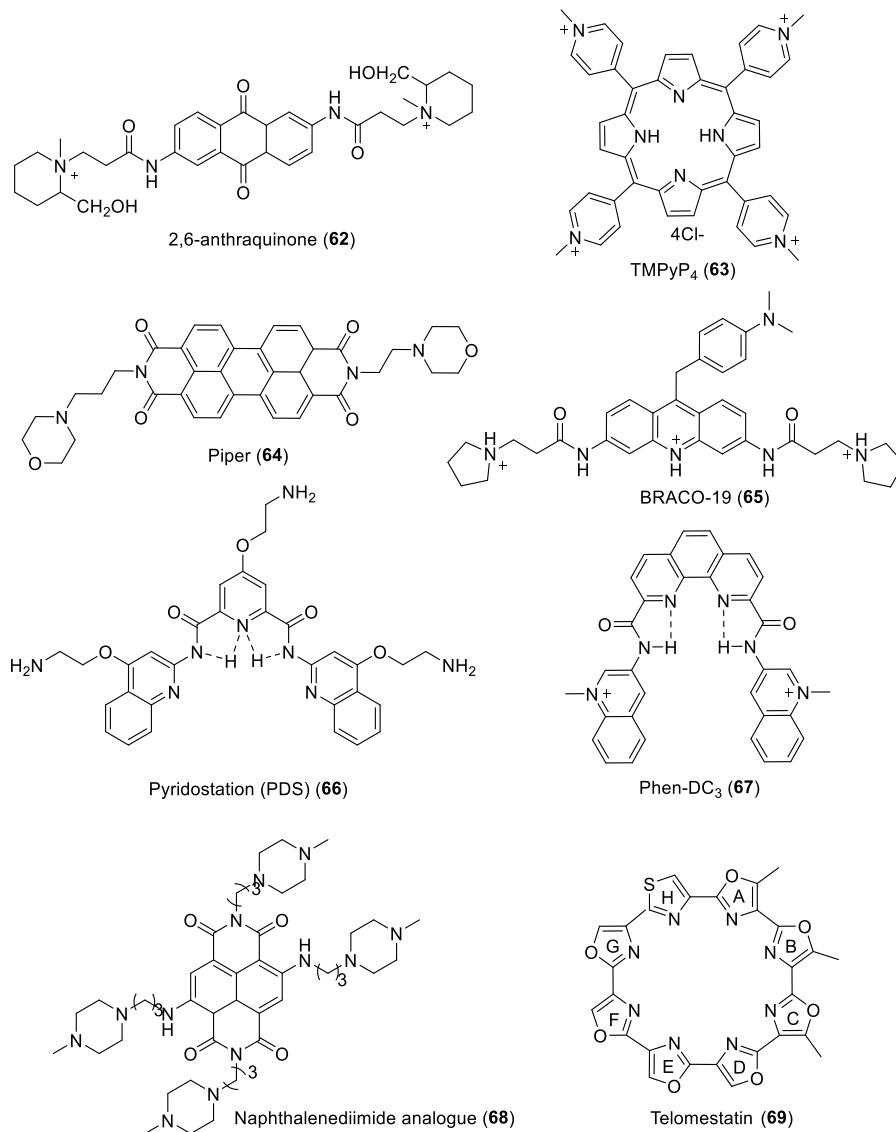
Figure 1.46. Ligands targeting hybrid G-quadruplexes.

In summary, all the G-quadruplex ligands developed so far share several common features.

1). Positively charged side chains increase the binding affinity by interacting with negatively charged G-quadruplex backbones. However, the presence of multiple charges can reduce selectivity over duplex DNA

2). An extended aromatic surface is required for enhanced G-quadruplex stabilization.

3). The presence of positively charged nitrogen either may enhance interactions of the ligand with the cavity center of G-quadruplexes.

4). The ligand with biaryl or flexible linkages can readily fit in the grooves of G-quadruplexes.

5). Increase in hydrophobicity may increase the binding affinity to the G-quadruplex to a certain extent.

6). The G-quadruplex structures have water ribbons in the grooves. The binding interaction between the basic side chain and negatively charged backbone can occur through the intermediacy of water bridges.

In light of these findings, my research focuses on developing potent ligands that bind to G-quadruplexes via a dual recognition mode (extended aromatic surface and side chains). Two classes of aromatic molecules, nucleobases and 1,10-phenanthroline, are used. Neomycin and polyamines will be used as cationic side chains. The synthesis and biochemical evaluations of these molecules will be discussed in the following chapters.

- In chapter 2, four novel nucleobase-neomycin conjugates are synthesized, and their binding to triplex DNA and telomeric G-quadruplex DNA is evaluated. Also, the anti-bacterial effect of these conjugates is determined.

- In chapter 3, several 1,10-phenanthroline-neomycin conjugates are synthesized and characterized. Their binding to telomeric G-quadruplex DNA and duplex DNA are studied using various biophysical methods. In addition, the inhibition of telomerase activity by these compounds is determined using the TRAP assay. In chapter 4, a series of 2-clip-phen (a 1,10-phenanthroline dimer) derivatives with amino side chains are synthesized and characterized. The binding of these compounds with three different G-quadruplexes is studied using various biophysical methods.

CHAPTER 2: NUCLEOBASE-MODIFIED NEOMYCIN TO TARGET A-FORM OF

NUCLEIC ACID

2.1. Introduction

Aminoglycosides are potent antibiotics that have widely used in bactericidal infection. The mechanism of action of most aminoglycosides is binding to RNA targets and creating fissures in the outer membrane of bacterial cells. For instance, neomycin, an aminoglycoside, binds to the 16S ribosomal A-site, which causes misreading of the genetic code and inhibition of translocation[146,147], leading to protein synthesis disruption and cell death. They are effective against aerobic, gram-negative bacteria and certain gram-positive bacteria. Aminoglycosides are ototoxic and have acute nephrotoxicity associated with them. They accumulate in kidney cells with a long half-life and damage kidneys over time[148]. A recent report showed that a Ru(II) arene-neomycin conjugate has moderate antiproliferative activity against MCF-7 cells (human breast adenocarcinoma) at low micromolar concentrations with low selectivity over normal cells[149].

Recent studies revealed that neomycin has a binding preference for the A-type nucleic acid structures like triplex DNA, duplex RNA, and DNA-RNA hybrid[150]. The specificity of neomycin to these structures is due to the flexible glycosidic linkages and structured amino groups, which are positively charged under physiological conditions[151]. Over the past years, several neomycin conjugates have been synthesized to modulate its binding to several nucleic acid structures. The C5"-position of the ring III is a common site for tethering another moiety because the chemical modification at this site has a minimal effect on the binding of neomycin to its substrates. Using this approach, Tor group synthesized an acridine-neomycin conjugate that

binds with high affinity to the HIV-1 Rev-Response Element[152]. Several intercalator-neomycin

conjugates have been developed by tethering an aromatic moiety at the C5"- position, showing

enhanced binding to the triplex DNA over the unmodified neomycin[153]. In addition, a perylene-

neomycin conjugate has been found to increase the binding affinity of the intercalator (perylene)

to G-quadruplex DNA[154]. In all the cases mentioned here, the enhanced binding of these

conjugates to their nucleic acid substrates result from a synergistic effect, in which neomycin

binds into the groove and the intercalator moiety stacks with aromatic bases of the nucleic acid

structure simultaneously. Neomycin has become a promising scaffold for the development of

ligands that binds unique nucleic acid structures specifically.

On the other side, due to the over and indiscriminate use of aminoglycosides, the

emergence of resistant strains, especially those containing aminoglycosides modifying enzymes

has become a global concern for public health[155]. Modifications of existing aminoglycosides to

circumvent the resistance while keeping their antibiotic potential is one of the useful approaches

to quickly tackle this problem as the therapeutic and toxicity profiles of these aminoglycosides

are already known. Zhang *et al*[156] synthesized a series of 5"-acetylated neomycin derivatives

and found two lead compounds that were effective against methicillin and vancomycin-resistant

bacterial species. In another study, bisubstrate derivatives of neamine (truncated neomycin) and

adenosine were synthesized as the competitive inhibitors of the enzyme aminoglycoside 3'-O-

phosphotransferases[157].

In this chapter, we present the synthesis of four novel nucleobase-neomycin conjugates

containing a triazole linker. The binding of these conjugates with duplex, triplex, and G-

quadruplex DNA was studied. The goal of the study was to explore the size limitation of the

intercalator moiety in an intercalator-neomycin conjugate on its binding with DNA. In addition,

the cytotoxicity of these novel conjugates against HeLa and MCF-7 cancer cells and their

antibacterial effect were determined.

2.2. Experimental Section

2.2.1. Material and Methods

Chemicals for synthesis were purchased from Fisher Scientific (Pittsburgh, PA) and used

without further purification. Reactions were carried out under Argon using dry solvent unless

otherwise noted. IR spectra were collected on a Shimadzu (Pleasanton, CA) IR Prestige-21 FT-

IR spectrophotometer. [1]H spectra were collected on a JEOL (Peabody, MA) ECA 600 MHz FT-

NMR spectrometer. HRMS (ESI) spectra were collected on an Agilent 1290 UHPLC connected

to an Agilent 6230 Time-of-Flight mass spectrometer through an electrospray ionization source

at Chevron, Richmond, California. UV spectra were collected on a Varian Cary 100 Bio UV-Vis

spectrophotometer equipped with a thermoelectrically controlled 6 x 6 cell holder (Walnut

Creek, CA). The MTT read-outs were measured by a TriStar LB 941 multimode microplate

reader. Dulbecco's modified Eagles medium (DMEM) was purchased from Sigma (St. Louis,

MO). Advanced MEM, glutaMAX, penicillin, and streptomycin were purchased from

Invitrogen (Carlsbad, CA). Fetal bovine serum (FBS) was purchased from Tissue Culture

Biologicals (Seal Beach, CA).

2.2.2. Synthesis of Nucleobase-neomycin Conjugates

Compounds 2-4 and 11-14 have been synthesized according to the published

procedure[158–160]. The synthesis scheme is described in Scheme 2.1.

Synthesis of 2. To the solution of neomycin B (1.0g, 1.1 mmol) in water (7ml),

triethylamine (2ml, 14.43 mmol) was added, and the reaction mixture was stirred for 10 mins at

rt. Di-tert-butyl dicarbonate (2.1g, 9.756 mmol) in DMF (20ml) was added dropwise into the

flask over a period of 10 mins. The reaction mixture was stirred at 60 °C for 5 h. The completion of the reaction was monitored with TLC (R_f = 0.414, CH_2Cl_2: CH_3OH = 9:1 (v/v), silica gel 60 F_{254}, EMD Millipore]. The residue was precipitated with water (200 ml) and filtered. Flash chromatography of the residue [silica gel, $CH_3COOC_2H_5$, R_f = 0.19] yielded compound **2** (700 mg and 52.4 %) as a white solid. ^{1}H-NMR (600 MHz, CD$_3$OD) δ 1H-NMR (600 MHz, CDCl$_3$) δ 6.18 - 4.57 (m, 7H), 4.54 - 4.03 (m, 4H), 4.00 - 2.89 (m, 16H), 1.65 - 1.05 (m, 54H).

Synthesis of 3. In a 50-ml round-bottomed flask, 2,4,6-triisopropylbenzene sulphonyl chloride (7g, 23 mmol) and **2** (1g, 0.83mmol) were dissolved in anhydrous pyridine (20ml). The reaction mixture is stirred at rt for 48 h under N$_2$ and then concentrated under vacuum. The crude compound was in two stages. In stage one, all the unreacted 2,4,6-triisopropylbenzene sulphonyl chloride was removed using CH_2Cl_2 as eluent. The column was then flushed with CH_3OH:CH_2Cl_2 (2:8 v/v), [R_f = 0.487 in a 9:1 (v/v) CH_2Cl_2: CH_3OH = 9:1 (v/v), silica gel 60 F_{254}, EMD Millipore]. Residue was concentrated and flash chromatography of the residue [silica gel, CH_3OH:CH_2Cl_2 (2 - 5 % v/v)] yielded compound **3** (971.5 mg, 79 %) as white solid. ^{1}H - NMR (600 MHz, CD$_3$OD) δ 7.27 (s, 2H), 5.42 (s, 1H), 5.16 (s, 1H), 4.85 (s, 1H), 4.01 - 4.42 (m 7H), 3.86 (s, 1H), 3.65 - 3.79 (m, 3H), 3.31 - 3.63 (m, 9H), 3.09 - 3.25 (m, 1H), 2.85 - 3.00 (m, 2H), 2.67 - 2.79 (m, 2H), 1.90 - 2.00 (m, 1H), 1.32 - 1.57 (m, 54H), 1.15-1.29 (m, 18H).

Synthesis of 4. To the solution of 3 (330mg, 0.22 mmol) in DMF (5ml), was added NaN$_3$ (110 mg, 1.687 mmol). The reaction mixture was stirred at 90 °C for 24h. The reaction completion was monitored with TLC [R_f = 0.26, CH_3OH:CH_2Cl_2 = 0.625:9.375 (v/v), silica gel 60 F_{254}, EMD Millipore]. The reaction mixture was concentrated under vacuum. Flash chromatography of the residue [silica gel, CH_3OH:CH_2Cl_2 = 0.625:9.375 (v/v)] yielded

compound **4** (223mg, 82 %) as a white solid. The yield is 223 mg, 82 % as a white solid. ^{1}H-NMR (600 MHz, CD$_3$OD) δ 5.46 (s, 1H), 5.15 (s, 1H), 4.91 (s, 1H), 4.23 - 4.39 (m, 2H), 3.99 - 4.09 (m, 1H), 3.67 - 3.96 (m, 5H), 3.42 - 3.63 (m, 8H), 3.10 - 3.21 (m, 1H), 1.89 - 2.00 (m, 1H), 1.39 - 1.53 (m, 54H).

Synthesis of 5a. Adenine-9-prop-2ynyl (**11**) (12.5 mg, 0.072 mmol) was added to an anhydrous DMF solution (6 mL) of sodium ascorbate (13.7 mg, 0.069 mmol), copper (II) sulfate (2.0 mg, 0.008 mmol), and compound **4** (71.0 mg, 0.058 mmol). The reaction mixture was stirred at room temperature overnight. The completion of the reaction was monitored by TLC under UV [Rf= 0.23, CH$_3$OH:CH$_2$Cl$_2$ 0.8:9.2 (v/v), silica gel 60 F$_{254}$, EMD Millipore]. The reaction mixture was concentrated under vacuum. Flash chromatography of the residue [(silica gel, CH$_3$OH:CH$_2$Cl$_2$ 1:9 (v/v)] yielded compound **5a** (67.5 mg, 82.0%) as a pale white solid. IR (KBr) (cm^{-1}) 3343, 2976, 2938, 1644, 1649, 1631, 1367, 1248, 1165, 1043. ^{1}H NMR (CD$_3$OD, 600 MHz) δ 8.32 (s, 1H), 8.22 - 8.29 (m, 2H), 6.47 - 6.83 (m, 2H), 5.58 - 5.70 (m, 1H). 5.33 - 5.44 (bs, 1H), 5.01-5.11 (m, 1H), 4.97 (s, 1H), 4.78 - 4.86 (m, 1H), 4.66 - 4.76 (m, 1H), 4.07 - 4.41 (m, 3H), 3.98 (t, 1H, J = 5.8 Hz), 3.90 (s, 1H), 3.66 - 3.83 (m, 3H), 3.50 - 3.65 (m, 2H), 3.41 - 3.49 (m, 2H), 2.9 8- 3.08 (m, 1H), 1.90 - 2.01 (m, 1H), 1.15 - 1.56 (m, 54H). HRMS (ESI) calcd. for C$_{61}$H$_{100}$N$_{14}$O$_{24}$ [M+ H]$^+$ = 1413.7113, found 1413.7090.

Synthesis of 5b. Guanine-9-prop-2ynyl (**12**) (14.0 mg, 0.074 mmol) was added to an anhydrous DMF solution (6 mL) of sodium ascorbate (15.0 mg, 0.075 mmol), copper (II) sulfate (3.0 mg, 0.012 mmol), and compound **4** (46.0 mg, 0.037 mmol). The reaction mixture was stirred at room temperature overnight. The completion of the reaction was monitored by TLC under UV [Rf= 0.275, CH$_3$OH:CH$_2$Cl$_2$:NH$_4$OH 1.25:8.75:0.1 (v/v/v), silica gel 60 F$_{254}$, EMD Millipore]. The reaction mixture was concentrated under vacuum. Flash chromatography of the

residue [(silica gel, $CH_3OH:CH_2Cl_2:NH_4OH$ 1.25:8.75:0.1 v/v/v)] yielded compound **5b** (25.0 mg, 47.3%) as a pale white solid. IR (KBr) (cm^{-1}) 3408, 1689, 1633, 1516, 13676, 1251, 1167, 1043, 856. ^{1}H NMR (CD$_3$OD, 600 MHz) δ 8.36 (s, 1H), 7.83 (s, 1H), 6.62 - 6.84 (m, 1H), 5.56 (d, J = 14.4 Hz, 1H), 5.36 (s, 1H), 5.28 (d, J = 15.6 Hz, 1H), 5.01 (s, 1H), 4.96 (s, 1H), 4.6 - 4.73 (m, 1H), 3.90 - 4.15 (m, 4H), 3.87 (s, 1H), 3.68 - 3.79 (m, 2H), 3.38 - 3.59 (m, 7H), 3.28 - 3.37 (m, 1H), 3.04 (br s, 1H), 1.84 - 1.96 (m, 1H), 1.2 5- 1.49 (m, 54H). HRMS (ESI) calcd. for $C_{61}H_{100}N_{14}O_{25}$ [M+ H]$^+$ =1429.7062, found 1429.7042.

Synthesis of 5c. Cytosine-1-prop-2ynyl (**13**) (24.0 mg, 0.161 mmol) was added to an anhydrous DMF solution (6 mL) of sodium ascorbate (14.6 mg, 0.0738 mmol), copper (II) sulfate (2.0 mg, 0.008 mmol), and compound 4 (64.0 mg, 0.054 mmol). The reaction mixture was stirred at room temperature overnight. The completion of the reaction was monitored by TLC under UV [Rf= 0.435, $CH_3OH:CH_2Cl_2:NH_4OH$ 1.5:8.5:0.05 (v/v/v), silica gel 60 F_{254}, EMD Millipore]. The reaction mixture was concentrated under vacuum. Flash chromatography of the residue [silica gel, $CH_3OH:CH_2Cl_2$ 1.25:8.75 (v/v)] yielded compound **5c** (30.0 mg, 42.0%) as a pale white solid. IR (KBr) (cm^{-1}) 3369, 2978, 2932, 1701, 1647, 1524, 1394, 1367, 1280, 1251, 1169, 1043, 865. ^{1}H NMR (CD$_3$OD, 600 MHz) δ 8.07 (s, 1H), 7.73 (d, J = 7.26 Hz, 1H), 5.85 (d, J = 7.26 Hz, 1H), 5.47 (s, 1H), 5.41 (s, 1H), 5.00 - 5.13 (m, 3H), 4.92 (d, J = 1.56 Hz, 1H), 4.66 (dd, J = 13.02 Hz, 4.38 Hz, 1H), 4.58 (s, 1H), 4.07 - 4.36 (m, 3H), 3.95 (m, 1H), 3.88 (s, 1H), 3.69 - 3.79 (m, 2H), 3.57 - 3.67 (m, 1H), 3.39 - 3.55 (m, 4H), 3.31 - 3.38 (m, 2H), 3.09 (m, 1H), 1.86 - 1.97 (m, 1H), 1.25 - 1.49 (m, 54H). HRMS (ESI) calcd. for $C_{61}H_{100}N_{14}O_{25}$ [M+ H]$^+$ = 1389.7001, found 1389.6990.

Synthesis of 5d. Thymine-1-prop-2ynyl (**14**) (13.7 mg, 0.0836 mmol) was added to an anhydrous DMF solution (6 mL) of sodium ascorbate (15.8 mg, 0.079 mmol), copper (II) sulfate (2.0 mg, 0.008 mmol), and compound 4 (55.0 mg, 0.0442 mmol). The reaction mixture was stirred at room temperature overnight. The completion of the reaction was monitored by TLC under UV [Rf= 0.2125, $CH_3OH:CH_2Cl_2$ 0.75:9.25 (v/v), silica gel 60 F_{254}, EMD Millipore]. The reaction mixture was concentrated under vacuum. Flash chromatography of the residue [silica gel, $CH_3OH:CH_2Cl_2$ 0.75:9.25 (v/v)] yielded compound **5d** (60.0 mg, 96.7%) as a pale white solid. IR (KBr) (cm^{-1}) 3371, 2978, 2934, 1690,1514, 1458, 1392, 1367, 1267, 1250, 1167, 1043, 862, 781. ^{1}H NMR (CD$_3$OD, 600 MHz) δ 8.10 (s, 1H), 7.56 (s, 1H), 6.55 - 6.67 (m, 1H), 6.37 - 6.53 (m, 1H), 5.47 (s, 1H), 5.41 (s, 1H), 4.89 - 5.14 (m, 3H), 4.54 - 4.68 (m, 1H), 4.04 - 4.47 (m, 3H), 3.95 (t, J = 6.6 Hz, 1H), 3.70 - 3.80 (m, 2H), 3.60 - 3.69 (m, 1H), 3.36 - 3.59 (m, 3H), 3.08 (t, J = 8.4 Hz 1H), 1.87 - 1.95 (m, 1H), 1.85 (d, J = 1.3 Hz, 3H), 1.25-1.48 (m, 54H). HRMS (ESI) calcd. for $C_{61}H_{101}N_{11}O_{26}$ [M+ H]$^+$ = 1404.6997, found 1404.6971.

General procedure for the synthesis of 6a–6d. Hydrochloric acid in dioxane (4M, 1 mL) and ethanedithiol (3 μL) were added into a dioxane solution (1 mL) of compound **5a** (17.0 mg), 5b (24.0 mg), 5c (12.0 mg), or 5d (24.0 mg) in a 15-mL conical vial. The reaction mixture was swirled for 5 min and became turbid. The deprotected pro- duct was further precipitated by adding ether and hexane (1 mL each). The solid was collected by centrifugation (3000 rpm), decanted, washed with ether (3 × 1 mL) and hexane (3 × 1 mL), and dried overnight under vacuum. The residue was re-dissolved in milli-Q water and lyophilized to give a pale-white fluffy solid.

Compound 6a. Yield: 11.6 mg, 93.5%. IR (KBr) (cm^{-1}) 3396, 3058, 2948, 1685, 1601, 1491, 1144, 1044, 1018, 769, 638. ^{1}H NMR (D$_2$O, 600 MHz) δ 8.32 (s, 1H), 8.30 (s, 1H), 8.17 (s, 1H), 5.94 (d, J = 4.2 Hz, 1H), 5.56 (s, 2H), 5.31 (d, J = 3.6 Hz, 1H), 5.21 (d, J = 1.8 Hz, 1H), 4.76 (dd, J = 15.0 Hz, 5.0 Hz, 1H), 4.68 (m, 1H), 4.45 (m, 2H), 4.23 (m, 1H), 4.12 (t, J = 6.0 Hz, 1H), 4.05 (t, J = 9.6 Hz, 1H), 3.82 -3 .93 (m, 4H), 3.72 (s, 1H), 3.62 (t, J = 9.6 Hz, 1H), 3.21 - 3.51 (m, 9H), 2.38 (m, 1H), 1.82 (m, 1H). HRMS (ESI) calcd. for C$_{31}$H$_{53}$N$_{14}$O$_{12}^+$ [M+ H]$^+$ = 813.3962, found 813.3945.

Compound 6b. Yield: 14.7 mg, 84.0%. IR (KBr) (cm^{-1}) 3383, 2922, 1701, 1637, 1604, 1500, 1490, 1369, 1143, 1047, 1024, 781, 605. ^{1}H NMR (D$_2$O, 600 MHz) δ 8.16 (s, 1H), 8.05 (s, 1H), 5.52 (d, J = 4.2 Hz, 1H), 5.34 (s, 2H), 5.23 (d, J = 4.2 Hz, 1H), 5.19 (d, J = 1.2 Hz, 1H), 4.71 - 4.75 (m, 1H), 4.65 (s, 1H), 4.43 (q, J = 4.2 Hz, 1H), 4.38 (t, J = 4.8 Hz, 1H), 4.20 (t, J = 4.8 Hz, 1H), 4.09 (t, J = 3 Hz, 1H), 3.74 - 3.90 (m, 4H), 3.70 (s, 1H), 3.17 - 3.51 (m, 11H), 2.35 (dt, J = 8.4 Hz, 4.2 Hz, 1H), 1.75 - 1.86 (m, 1H). HRMS (ESI) calcd. for C$_{31}$H$_{53}$N$_{14}$O$_{13}^+$ [M+ H]$^+$ = 829.3911, found 829.3909.

Compound 6c. Yield: 6.7 mg, 77.0%. IR (KBr) (cm^{-1}) 3392, 1676, 1497, 1144, 1053, 1018, 797. ^{1}H NMR (D$_2$O, 600 MHz) δ 8.08 (s, 1H), 7.83 (d, J = 7.8 Hz, 1H), 6.04 (d, J = 7.2 Hz, 1H), 5.93 (d, J = 3.0 Hz, 1H), 5.30 (d, J = 2.4 Hz, 1H), 5.22 (s, 1H), 5.02 (s, 2H), 4.55 - 4.84 (m, 2H), 4.45 (s, 2H), 4.24 (t, J = 3.6 Hz, 1H), 4.12 (s, 1H), 4.05 (t, J = 9.6 Hz, 1H), 3.77 - 3.99 (m, 4H), 3.73 (s, 1H), 3.60 (t, J = 10.2 Hz, 1H), 3.19 - 3.55 (m, 9H), 2.38 (m, 1H), 1.83 (q, J = 12.0 Hz, 1H). HRMS (ESI) calcd. for C$_{30}$H$_{53}$N$_{14}$O$_{12}^+$ [M+ H]$^+$ = 789.3850, found 789.3859.

Compound 6d. Yield: 14.7 mg, 84.0%. IR (KBr) (cm^{-1}) 3387, 2918, 1701, 1601, 1244, 1220, 1143, 1049, 1026, 773. ^{1}H NMR (D$_2$O, 600 MHz) δ 8.03 (s, 1H), 7.51 (d, J = 1.2 Hz, 1H), 5.82 (d, J = 3.6 Hz, 1H), 5.28 (d, J = 3.6 Hz, 1H), 5.21 (d, J = 1.2 Hz, 1H), 4.94 (d, J = 2.4 Hz,

2H), 4.75 (dd, J = 15.0 Hz, 3.0 Hz, 1H), 4.56 - 4.72 (m, 1H), 4.39 - 4.48 (m, 2H), 4.20 - 4.25 (m, 1H), 4.11 (t, J = 3.0 Hz, 1H), 3.84-3.96 (m, 3H), 3.74 - 3.83 (m, 2H), 3.69 - 3.73 (m, 1H), 3.54 (t, J = 10.2 Hz, 1H), 3.47 - 3.51 (m, 1H), 3.20 - 3.46 (m, 8H), 2.32 (m, 1H), 1.70 - 1.82 (m, 4H). HRMS (ESI) calcd. for $C_{31}H_{54}N_{11}O_{14}^+$ $[M+ H]^+$ = 804.3846, found 804.3850.

2.2.3. Measurements of High-resolution Mass Spectra

Samples were introduced by flow injection analysis (FIA) in positive-ion mode. The mobile phase was 90% H_2O and 10% MeOH with 0.1% formic acid. The nebulizer pressure was 50 psi with 12 L/min of nitrogen (drying gas) at 325 °C for desolvation. The capillary voltage was set at 3500 V, the fragmentor voltage was set at 150 V, the skimmer voltage was set at 65 V, and the octopole voltage was set at 750 V. Spectra were collected from 300–1500 Da at a rate of 2 spectra/sec with 4962 transients/spectrum.

2.2.4. Preparation of Triplex DNA for Thermal Denaturation

The oligonucleotides dA_{22} (1 μM) and dT_{22} (2 μM) were mixed with a solution (1 mL, pH 7.0) of sodium cacodylate (10 mM), EDTA (1 mM), and sodium chloride (200 mM) in the absence or presence of a ligand (10 μM). The mixtures were heated at 90 °C for 5 min, slowly cooled at 25 °C, and incubated at 4 °C overnight before use.

2.2.5. Preparation of G-quadruplex DNA for Thermal Denaturation

The G-quadruplex sequence 5′-AGG(TTAGGG)$_3$T(1 μM) was mixed with a solution (1 mL, pH 7.0) of lithium cacodylate (10 mM), EDTA (1 mM), and potassium chloride (50 mM) in the absence or presence of a ligand (10 μM). The mixtures were heated at 90 °C for 5 min, slowly cooled at 25 °C, and incubated at 4 °C overnight before use.

2.2.6. Thermal Denaturation of Triplex DNA or G- quadruplex DNA

UV melting spectra were recorded at 260 nm for triplex DNA and 295 nm for G-quadruplex DNA as a function of increasing temperature (4-80 °C, heating rate: 0.5 °C/min). The melting temperatures (T_m) were determined using the first derivative method.

2.2.7. The MTT Assay

All the experiments were carried out in triplicate. MCF-7 (human breast adenocarcinoma) and HeLa cells were maintained in advanced DMEM medium supplemented with 5% fetal bovine serum (FBS), L-glutamine, and anti-mycotic (antibiotic) at 37 °C in a humid atmosphere containing 5% CO_2. Neomycin and nucleobase-neomycin conjugates (**6a**–**6b**) were pre-dissolved in sterilized water. The cells (3000 cells/well, 200 µL) were seeded in a 96-well microtiter plate and incubated for 24 h. The supernatant in each well was removed, followed by adding neomycin or conjugates in FBS-free media (200 µL) to reach a final concentration of 500 µM. After incubation at 37 °C for 24 h, an MTT solution (20 µL, 0.5 mg/mL) was added to each well and incubated again for 4 h. The formazan product was dissolved in 150 µL of DMSO, and the absorbance of each well was measured at 590 nm using a plate reader. Untreated cells and media without cells were used as controls. To calculate the viability (%), the following equation was used.

$$\text{Viability \%} = \frac{\text{A treated} - (\text{A media})}{(\text{A untreated}) - (\text{A media})} \times 100 \text{ \%}$$

$A_{treated}$: The absorbance of the solution containing treated cells. A_{media}: The absorbance of the media. $A_{untreated}$: The absorbance of the solution containing untreated cells.

2.2.8. Statistical Analysis

Minitab software was used to determine statistical significance. Two-sample student's t-test was performed to show statistically significant ($p < 0.05$) and insignificant ($p > 0.05$) data.

2.2.9. Kirby-Bauer Test Protocol (Disk Diffusion Zone Determination)

Neomycin or conjugate (50 nmol) impregnated paper discs were placed onto the surface of Müller Hinton agar plates (70191 Fluka) that were pre-seeded with bacteria (Bacillus subtilis, Escherichia coli, Proteus vulgaris, and Serratia marcescens) and incubated aerobically at 37 °C for 24 h. The diameters of inhibition zones were measured using a ruler to the nearest 1 mm.

2.2.10. Determination of Minimum Inhibition Concentrations (MICs)

Bacterial species were grown overnight in the Müller Hin- ton broth at 37 °C. A dilution of the overnight culture was made in broth to an approximate concentration of 106 cfu/ mL. A Müller Hinton broth (100 µL) solution containing the desired antibacterial reagent (neomycin or compound **6a**–**6d**) was inoculated with a freshly diluted bacterial suspension (5 µL) in a 96-well plate and incubated at 37 °C overnight. The OD_{600} value in each well was recorded to determine the bacterial growth. The OD_{600} value of the broth was used as a reference to determine the complete inhibition of bacterial growth, and the OD_{600} value of the bacterial solution without any reagent was used as a positive control.

<div align="center">

2.3. Results and Discussion

</div>

2.3.1. Synthesis of Nucleobase-neomycin Conjugates via Click Chemistry

Synthesis of nucleobases with a propargyl group (**11**-**14** containing adenine, guanine, cytosine, and thymine, respectively) was reported previously[159,160]. The C5"-azido neomycin precursor (4) were synthesized separately for the click reaction based on a previously published procedure[152]. The modification at the C5"-position ring III of neomycin does not alter its conformation for binding with nucleic acids[154,158]. All the six amino groups of neomycin (**1**) were protected using di-tert-butyl decarbonate to yield compound **2**. The C5"-hydroxyl group of **2** was converted into a good leaving group using 2,4,6- triisopropylbenzenesulfonyl chloride.

The use of this bulky reagent enabled us to selectively protect the C5" primary hydroxy of ring III of neomycin with a Boc group. Replacement of 2,4,6- triisopropylbenzenesulfonylate in **3** with azide via an S_N2 mechnism yielded compound **4**. Compound **4** and nucleobase precursors (**11**-**14**) were finally coupled in the presence of Cu (I) via a click reaction to yield the compounds **5a to 5d** containing a 1,4 substituted[161] 1,2,3-triazole linker. The boc-protected amino groups were protected using 3M HCl/dioxane to produce final nucleobase-neomycin derivatives **6a** to **6b** (**6a**: adenine-neomycin; **6b**: guanine-neomycin; **6c**: cytosine-neomycin; **6d**: thymine-neomycin) with a yield ranging from 77.0-93.5%. All the target compounds were characterized by IR, ^{1}H NMR, and high-resolution ESI mass spectrometry.
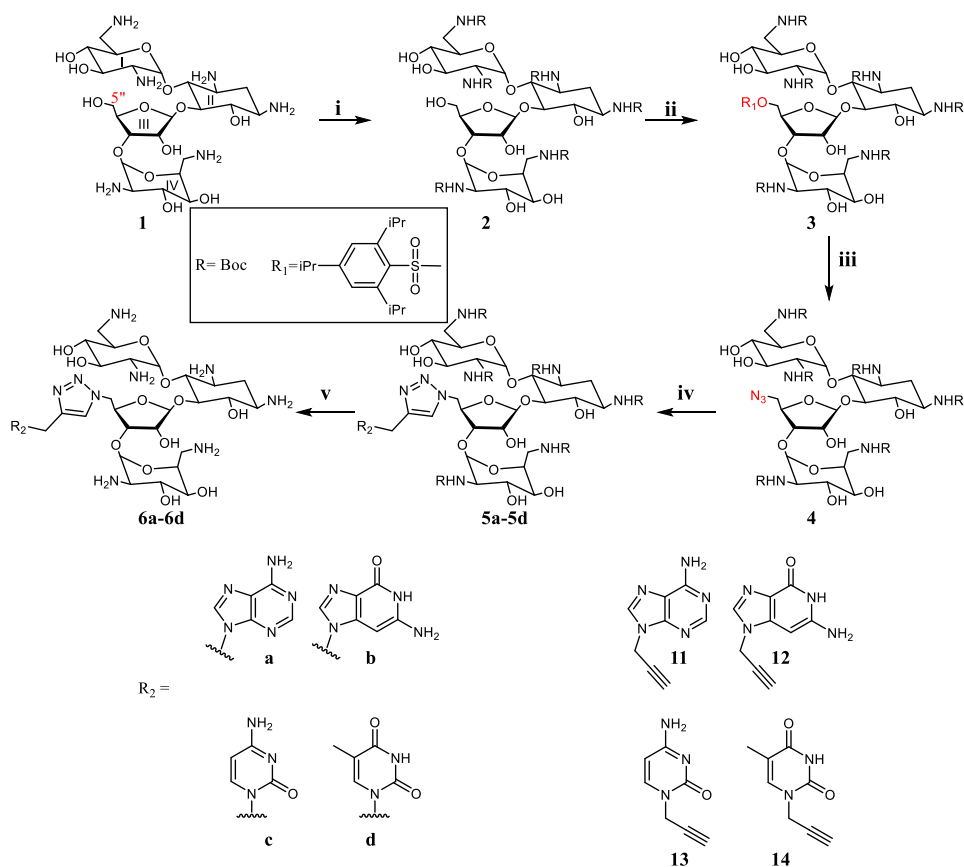
Scheme 2.1. Synthesis of nucleobase-neomycin conjugates (**6a-6d**), reagents and conditions: (i) (Boc)$_2$O, DMF-H$_2$O, 60 °C; (ii) 2,4,6-triisopropylbenzenesulfonyl chloride, pyridine, r.t.; (iii) NaN$_3$, DMF, 90 °C; (iv) **11–14**, sodium ascorbate, CuSO$_4$, DMF, r.t.; (v) 4M HCl in Dioxane, r.t.

2.3.2. **Nucleobase-neomycin Conjugates Stabilizes DNA-triplex Significantly**

The stabilization of neomycin-nucleobase conjugates on triplex DNA was ascertained using thermal denaturation studies monitored by UV at 260 nm. The observed change (increase) in absorbance as a function of temperature is plotted, as shown in (**Figure 2.1a**). The melting curve of an oligonucleotide DNA triplex (dA$_{22}$•2dT$_{22}$) is biphasic, including two distinct transitions. The first transition (increase in hyperchromicity) represents the dissociation of the triplex into a DNA duplex (dA$_{22}$•dT$_{22}$) and a single strand (dT$_{22}$). In the absence of the ligand, this transition is shown at 8.7 °C (T$_{m3\rightarrow2}$) using the published protocol[162]. The second

hyperchromic event ($T_{m2\rightarrow1}$) at 54.1 ^{o}C corresponds to the dissociation of the duplex DNA

(dA$_{22}$•dT$_{22}$) to two single-stranded random coils (dA$_{22}$ and dT$_{22}$). In order to compare the effect

of various ligands on triplex stabilization, we chose a fixed concentration, 10 µM, to conduct the

experiments. As shown in **figure 2.1b**, and Table 2.1, neomycin significantly stabilizes the

DNA triplex ($T_{m3\rightarrow2}$ = 39.3 ^{o}C), an increase in melting temperature of 30.6 ^{o}C ($\Delta T_{m3\rightarrow2}$) while

having no stabilization effect on the DNA duplex, which is consistent with the previously

published reports.
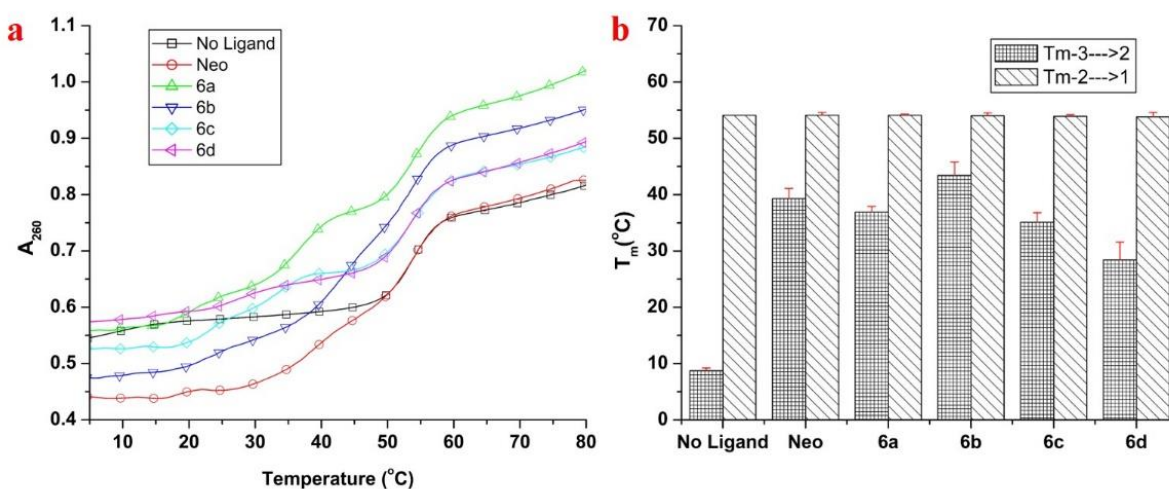
Figure 2.1. a) UV melting of a 22 mer T.A.T triplex (1 µM) in the absence and presence of
neomycin (neo) and compound **6a-6d** (10 µM). **b)** Bar graph showing the melting temperature of
(Continued) triplex to duplex ($T_{m3\rightarrow2}$) and duplex to random coil ($T_{m2\rightarrow1}$). Buffer conditions: 10
mM sodium cacodylate, 200mM NaCl, 1mM EDTA at pH 7.0.

        Neomycin snuggly fits into the Watson-Hoogsteen groove of the neomycin due to

charge/shape complementarity to the binding site[151]. Similar results were obtained with the

nucleobase-neomycin conjugates. All the compounds showed an increase in the triplex DNA

melting temperature ($T_{m3\rightarrow2}$) with a negligible effect on the duplex melting temperature ($T_{m2\rightarrow1}$).

Based on increase in the $\Delta T_{m3\rightarrow2}$, the order of triplex stabilization effect of these conjugates is guanine-neomycin (**6b**) > neomycin (**neo**) > adenine-neomycin (**6a**) > cytosine-neomycin (**6c**) > thymine-neomycin (**6d**).

Table 2.1
Melting Temperatures ($T_{m3\rightarrow2}$) of Triplex DNA to Duplex DNA and ($T_{m2\rightarrow1}$) Duplex to Single Strand Random Coil. 22 mer T.A.T triplex (1µM) in the Absence and Presence of Neomycin (neo) and Compounds 6a-6d (10uM).

Compound	$T_{(m3\rightarrow2)}$	$\Delta T_{(m3\rightarrow2)}$	$T_{(m2\rightarrow1)}$	$\Delta T_{(m2\rightarrow1)}$
No Ligand	8.7 ± 0.5	n/a	54.1 ± 0	n/a
Neo	39.3 ± 1.8	30.6	54.1 ± 0.5	0
6a	36.9 ± 1.0	28.2	54.1 ± 0.2	0
6b	43.4 ± 2.4	34.7	54.0 ± 0.5	-0.1
6c	35.1 ± 1.7	26.4	53.9 ± 0.3	-0.2
6d	28.7 ± 3.2	19.7	53.8 ± 0.8	-0.3

Interestingly, three conjugates (thymine-neomycin, adenine-neomycin, and cytosine-neomycin) could not stabilize the triplex DNA as efficiently as neomycin. The increment in $T_{m3\rightarrow2}$ values by these three conjugates was a few degrees lower than that by neomycin. This observation could result from the increased non-favorable steric crowding at the binding site of neomycin due to the conjugation of a nucleobase. It is known that the binding pocket of neomycin can be affected by the surrounding structural geometry[163]. On the other hand, compound **6b** (guanine-neomycin) increased the $T_{m3\rightarrow2}$ melting temperature 4.1 °C more than neomycin. The order for stacking surface area of nucleobases is guanine (G) > adenine (A) > cytosine (C) > thymine (T). Adenine, cytosine, and thymine in the conjugates have a smaller aromatic surface, which may not be sufficient for stacking. These three nucleobases could become a steric hindrance factor in the binding of neomycin to triplex DNA. By contrast, the guanine moiety of **6b** has the largest surface area that could efficiently intercalate with the base

triplets while neomycin occupies the groove of the triplex DNA. However, the observed

stabilization of triplex DNA by **6b** is relatively less efficient than that of other neomycin-

conjugates with a much larger surface area of the intercalator moiety such as pyrene-neomycin[164]

and BQQ-neomycin[153]. As a comparison, at 4 μM concentration, a pyrene-neomycin conjugate

has ΔT_m of 26 °C for triplex melting. BBQ-neomycin conjugate binds the triplex DNA so tightly

that the triplex DNA directly dissociates into random coils without the biphasic transitions. The

other difference between **6b** and these compounds is the linker. The linker of **6b** includes a

triazole ring. Both pyrene-neomycin and BQQ-neomycin conjugates contain an aliphatic linker

that is more flexible than the triazole ring. This study implicates that the nature of the linker in a

conjugate could be an essential factor of binding in addition to the aromatic stacking surface.

The stabilization of a telomeric G-quadruplex by these four conjugates was determined

using thermal denaturation studies monitored by UV at 290 nm in the presence and absence of a

conjugate. In a typical experiment, CD signals at 290 nm were monitored as a function of

temperature ranging from 20-90 °C. G-quadruplex dissociates into random coils at higher

temperatures with concomitant hyperchromicity at 290 nm. The temperature at which 50% of

the G-quadruplex dissociates into random coils is called melting temperature (T_m). As indicated

in **figure 2.2**, none of the conjugates could increase the dissociation temperature of the telomeric

G-quadruplex. In K^+, the G-quadruplex DNA used in the experiments adopted the hybrid type

conformation[165].
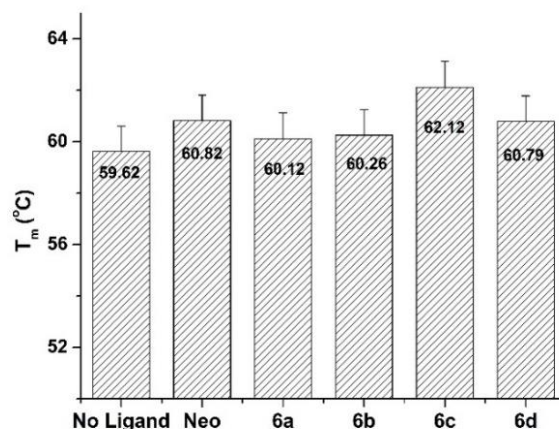
Figure 2.2. Bar graph showing G-quadruplex DNA (1µM) melting temperature in the presence of neomycin (10 µM) and compounds **6a**-**6d**. Buffer conditions: 10mM LiCaco, 50mM KCl, 1mM EDTA, pH 7.0.

The surface area of a G-tetrad is much larger than that of a base triplet. Stacking with a guanine nucleobase with a G-tetrad via π-π interactions may not be strong enough to obtain a noticeable increase in the melting temperature. In addition, only one or two conjugate molecules are able to end stack with a G-quadruplex, while more molecules can interact with a triplex DNA. Arya *et al* reported a similar result on a perylene-neomycin[154], which binds efficiently with G-quadruplex DNA, and a larger thermal stabilization effect was observed. The large stacking surface of perylene of the conjugate is believed to be a crucial factor for these observations. It is noteworthy that neomycin does not thermally stabilize G-quadruplex DNA. At most, it is a relatively weaker G-quadruplex groove binder[166]. In addition, the triazole ring in the linker might be too rigid for the two moieties to interact synergistically with G-quadruplex DNA. Taken together, we can conclude that the aromatic surface area of guanine could be the minimum for binding of the triplex DNA with an intercalator-neomycin conjugate via a

synergistic effect. A larger aromatic surface is required if one designs an intercalator-neomycin

conjugate for targeting G-quadruplex DNA.

2.3.3. Nucleobase-neomycin Conjugates Are More Cytotoxic Than Neomycin

The cytotoxicity of the conjugates was checked for two cancer cell lines, MCF-7 and

HeLa, using the MTT assay. This assay measures a cellular metabolic activity that reflects cell

viability. As a general procedure, cells are incubated with a chemical reagent, 3-(4,5-

dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT). NAD(P)H-dependent

oxidoreductase in live cells' mitochondria reduces MTT to produce insoluble, purple-colored

formazan (**Figure 2.3**).
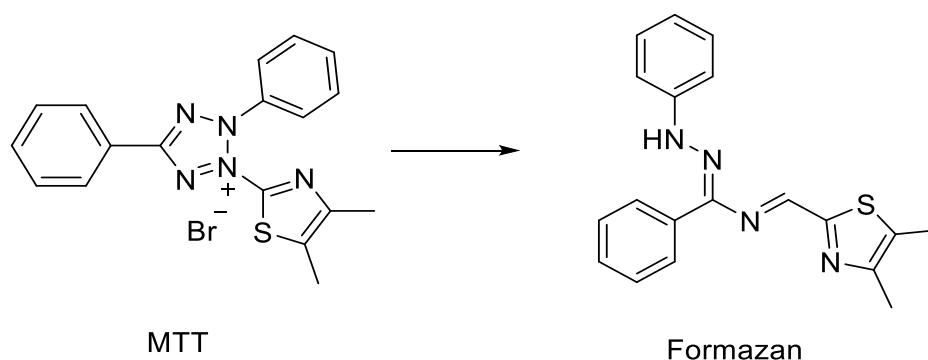
MTT   Formazan

Figure 2.3. The conversion of MTT to formazan with mitochondrial reductase.

The color of the latter can be measured using colorimetry at 500-600 nm. Neomycin has

relatively low toxicity towards MCF-7 cells ($IC_{50} > 250$ µM) as previously reported[149]; therefore,

we chose a concentration of 500 µM for all the compounds as the initial study. As shown in

Figure 2.4, the cell viability values of neomycin toward MCF-7 and HeLa cells were 94.6 ± 12.5

and 96.4 ± 1.5 %, respectively, showing expected low toxicity. All nucleobase-neomycin

conjugates showed better cytotoxicity against the two cells in a statistically significant manner.

All the conjugates (except **6d**) were more cytotoxic towards MCF-7 cells than HeLa cells. The

order of increasing toxicity towards MCF-7 cells and HeLa cell lines was **neo** < **6d** < **6a** < **6c** <

6b and **neo** < **6a** < **6d** < **6c** < **6b**, respectively.
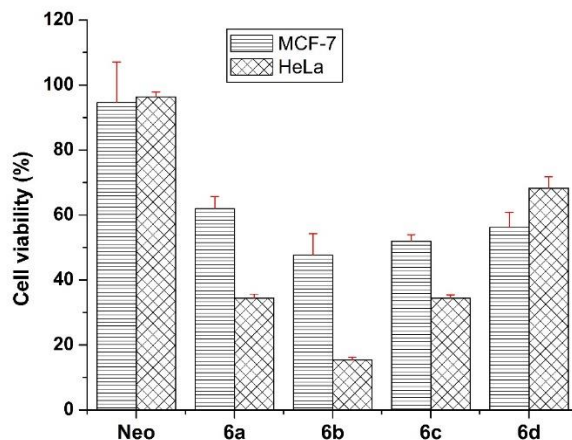
Figure 2.4. Cytotoxic effects of neomycin and compounds **6a**-**6d** (500 µM) in MCF-7 and HeLa
cell lines. Cell viability was determined by the MTT assay.

The guanine-neomycin conjugate (**6b**) was the most cytotoxic compound against both

cell lines. The cell viability of HeLa cells in the presence of 500 µM **6b** was 15.4 ± 0.8 %,

which is 5-fold better than neomycin and at least 2-fold better than the other conjugates in the

series. The cell viability of MCF-7 cells in the presence of **6b** under identical conditions was

47.6 ± 6.7 %. The cytotoxicity range of the conjugates towards MCF-7 cells (47.6 - 61.9%) was

smaller than that of HeLa cells (15.4-68.3%). This disparity might arise from different cellular

uptake of the conjugates by the two cell lines. **6b** contains the largest stacking moiety and

showed the best cytotoxic effect against the cancer cells lines. It is possible that the larger

guanine moiety can have more interactions with certain substrates, but the clear correlation

between the stacking moiety and cytotoxicity can not be drawn. The observed toxicity could also result from the binding of these conjugates to the cellular RNA[150,152] as neomycin is known to bind to the A form of the nucleic acids. We did not conduct concentration-dependent studies as IC_{50} values of these conjugates would be higher than the 250 µM, which is considered inactive[149]. Because nucleobases are not cytotoxic, the increased cytotoxicity of the conjugates should result from the structural modification at the C5" position of neomycin. Therefore, tethering a moiety to the C5"-position of the neomycin can be used to develop novel neomycin-based cytotoxic agents. Grau-Campistany and coworkers[149] have recently reported a Ru(II)-neomycin conjugate that showed enhanced cytotoxicity towards two cancer cells used in our studies.

2.3.4. Nucleobase-neomycin Conjugates Showed a Weaker Antibacterial Effect Than Neomycin

The antibacterial effect of nucleobase neomycin conjugates was assessed on both aminoglycoside resistant and non-resistant bacterial strains. Four antibiotic susceptible strains including *Serratia marcescens* (*S. marcescens*, Gram-negative), Escherichia coli (*E. coli*, Gram-negative), *Bacillus subtilis* (*B. subtilis*, Gram-positive), and *Proteus vulgaris* (*P. vulgaris*, Gram-negative) were first tested against 50 nmol of nucleobase-neomycin conjugates using the Kirby-Bauer (Agar-disk-diffusion) test[167]. Neomycin with the same concentration was used as a control. The size of an inhibition zone on the Agar disk reflects the inhibitory potency of a compound on bacterial growth. The larger the inhibition zone, the stronger the inhibition effect of the compound is. The diameter of an inhibition zone can be measured with a ruler (**Figure 2.5**). The size of the inhibition zone (cm) for neomycin was 1.5 for S. marcescens, 2.2 for *E. coli,* 1.9 for *B. subtilis*, and 2.0 for *P. vulgaris*. The median values of inhibition zones by four

conjugates were calculated for each bacterial strain. The values (cm) are 1.2 for *S. marcescens*, 1.*3 E. coli*, 1.6 for *B. subtilis*, and 0.7 for *P. vulgaris*. (**Figure 2.5**)
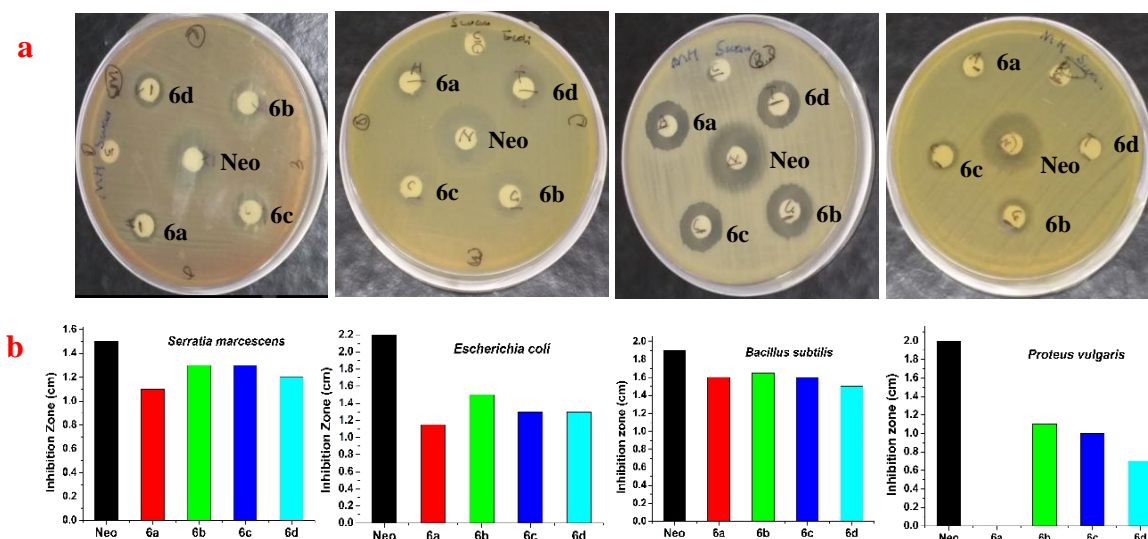
Figure 2.5. Petri-dish images of the Kirby-Bauer test. **a)** The zone of inhibition for compounds **6a**-**6d** (50 nmol) is a light circle around the disk. Neomycin is used as a positive control. **b)**. Bar graph depicting the relative size (cm) of the inhibition zone for each compound against a bacterial species.

By subtracting these median values from the corresponding values of neomycin, we interpreted the effect of conjugation on the antibacterial activity of neomycin. Neomycin showed the best antibacterial effect against all the four bacterial strains as compared to nucleobase-neomycin conjugates. Conjugation of a nucleobase to the C5"- position of neomycin reduced the antibacterial effect of neomycin. It is possible that modifications at the C5"-position could interfere with the binding of neomycin to its RNA target site. The observed reduction is bacterial strain-dependent. Conjugation of nucleobases to neomycin has less reduction effect on *S. marcescens* (Gram-negative) and *B. subtilis* (Gram-positive) than *E. Coli* (Gram-negative) and *P. vulgaris* (Gram-negative). Such a difference could result from a different inhibitory

mechanism or a different mechanism of cellular penetration between the Gram-positive and Gram-negative bacteria species. Amongst all the conjugates, guanine-neomycin (**6b**) is the most potent against all four bacterial species. Adenine-neomycin (**6a**) was the least potent against all the Gram-negative bacterial species. It is noteworthy that **6a** did not show any measurable effect on the inhibition of *P. vulgaris* at the concentration we used.

The minimum inhibitory concentration (MIC) of the conjugates against the four bacterial species was determined on 96-well microtiter plates. The absorbance of the well at 600 nm (OD_{600}) in the absence of bacteria was taken as a control, representing the complete inhibition effect of bacterial growth. As shown in **figure 2.6**, the order of bacterial species based on decreasing inhibitory effect by all compounds was *B. subtilis > E. coli > S. marcescens > P. vulgaris*.
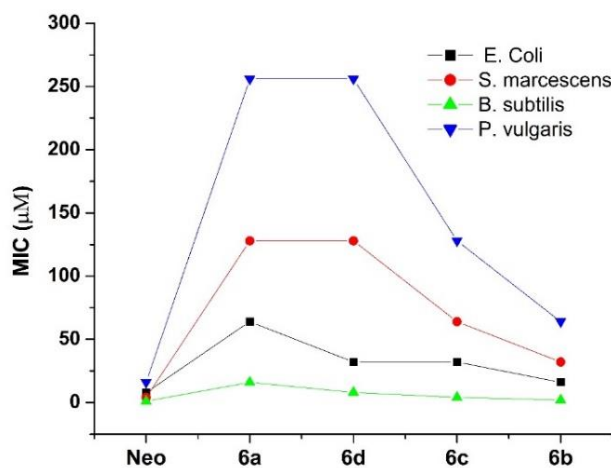
Figure 2.6. Minimum inhibition concentration of the compounds **6a**-**6d** against four bacterial species (*S. marcescens* (Gram positive), *E. coli* (Gram negative), *B. subtilis* (Gram negative) and *P. vulgaris* (Gram negative).

All the conjugates inhibited the growth of Gram-positive *B. subtilis* better than that of the other Gram-negative bacteria (**Figure 2.6**). **6b** is the most potent inhibitor amongst four conjugates against bacteria species used in the study. The MIC value of **6b** against Gram-positive *B. subtilis* was 2 µM, which is about twice higher than the value for neomycin. By contrast, **6b** had a MIC value of 32 µM against Gram-negative *S. marcescens*, which is about 8 times higher than the value for neomycin. The reduced activity against Gram-negative species could be attributed to the factors related to:

1. Bacterial anatomy- whether Gram-positive or negative.

2. Altered drug structure/conformational restraints.

Gram-negative bacteria have an additional outer membrane (OM) composed of a bilayer including glycolipids lipopolysaccharides (LPS), glycerol phospholipids, and periplasm, which provide protection against a wide range of antibiotics[168]. The hydrophobic drugs can diffuse through the outer membrane, but the intake of hydrophilic drugs like aminoglycosides is maintained through special protein gated channels called porins in the outer membrane[169]. The addition of nucleobase moiety may have altered the polarity of the molecule, which could result in the altered transport of conjugates into the cell to affect the target site. This is supported by the fact that both neomycin and its conjugates showed comparative inhibition of gram-positive bacteria species *B. subtilis*, as evident from the MIC values and Kirby-Bauer test. In addition, covalent tethering of nucleobase moiety at the C5" may introduce steric interactions at the binding site. A previous study suggested that the tethering of a moiety at the C5" position of neomycin could affect its interaction with the G1491 base at the binding site of 16S rRNA(**Figure 2.7a** and **Figure 2.7b**), which may lower the affinity of the conjugate to the target site with concomitant loss in antibacterial activity[170].
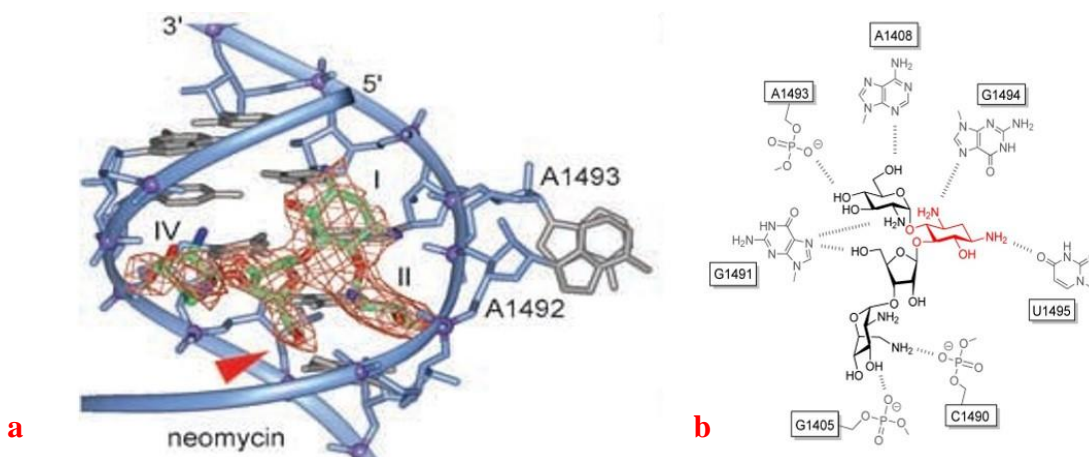
Figure 2.7. a) Neomycin bound in the A site of 16S rRNA[170] (point of nucleobase attachment is labelled in red).   b) Critical contacts between paromomycin (differ from neomycin in one amine group in ring II) and residues at the binding site[171] (right).

Moreover, tethering a moiety at the C5" position of neomycin could reduce the conformational freedom of neomycin[170]. It is known that the ring I and II of neomycin are involved in making conserved (similar among different aminoglycosides of this class)[172] critical connections (with A1408, G1494, U1495 at the decoding A site of the 16S ribosomes[170,171] (**Figure 2.7**). Subtle changes in conformation may reduce the efficiency of neomycin binding to its target.

The different responses of the Gram-negative bacterial species to the conjugates could be due to different permeability/efflux and possible off-targeting. However, it is difficult to pinpoint the exact reason due to the limitation of the study. In addition, no direct correlation was found between the antibacterial activity and the size of conjugated aromatic moiety although the polarity of nucleobases could play a role.

All four nucleobase-neomycin conjugates were also analyzed for their antibacterial activities against an aminoglycoside-resistant *E. coli* strain. This *E. coli* strain is transfected with

an aminoglycoside 3'-O-phosphotransferase (APHs) Ia gene. APHs transfer a phosphate group to the 3' or 5" hydroxyl group of the aminoglycoside molecule. Addition of a phosphate group changes the electronic property of the aminoglycoside and dramatically alters its ability to bind to the target rRNA[171] (**Figure 2.8**), increasing to the survival of the bacteria. This strain is known to have a resistance profile, including kanamycin, neomycin, and a few other aminoglycosides.
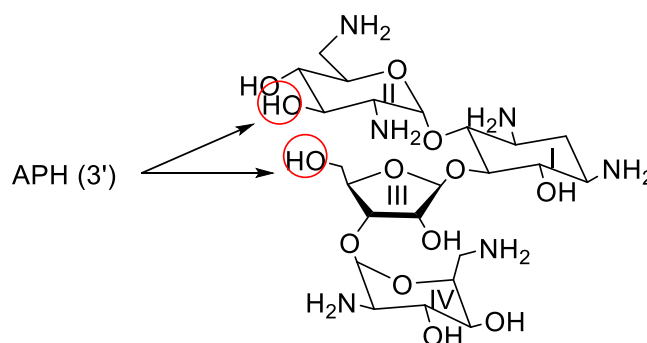
Figure 2.8. APH (3') I modifies neomycin at 3'OH (ring II) or at 5" OH at the ring III labelled with red circle.

Tethering a nucleobase moiety to the 5" position of neomycin could change the antibacterial profile of neomycin to the resistant *E. Coli* strain. It will be interesting to see if nucleobase-neomycin conjugates can act as better antibiotics against aminoglycoside-resistant strains. In our study, the MIC value of neomycin for this E. coli strain was 64 µM, which is 8 times higher than that of the non-resistant *E. Coli* strain, showing the presence of drug resistance against neomycin. The effect of resistance to kanamycin and four conjugates were much stronger than that to neomycin. No meaningful MIC value was observed for kanamycin at the concentration up to 512 µM. Amongst all four conjugates, only **6b** showed a detectable MIC

value of 512 µM, which is 32 times higher than that (16 µM) toward the non-transfected species. These results suggest that none of nucleobase-neomycin conjugates can overcome the bacterial resistance because all the MIC values are much higher than that of neomycin.

2.4. Conclusion

We have successfully completed the synthesis and characterization of four novel nucleobase-neomycin conjugates. The thermal denaturation studies showed that the size of the intercalator moiety has a large impact on the thermal stabilization of DNA triplex and G-quadruplex DNA. The aromatic surface size of guanine could be a minimal requirement for the synergistic binding effect on DNA triplets, but it is not sufficiently large for stacking with G-tetrad. The modifications at the C5" position of the neomycin is an efficient approach to enhance the cytotoxicity of the aminoglycoside (neomycin) against HeLa and MCF-7 cancer cells. The antibacterial activity of the conjugates behaves differently between the Gram-positive and Gram-negative bacteria. It is possible that addition of a nucleobase to the C5" position of neomycin may affect the binding of neomycin to its rRNA, but this modification may not affect the function of 3"-O-phosphotransferase. The study presented here has provided important insights into developing aminoglycoside-based compounds with better DNA binding affinities, altered cytotoxicities, and better antibacterial activities.

CHAPTER 3: PHENANTHROLINE-NEOMYCIN CONJUGATES, A POTENT G4 LIGANDS

3.1 Introduction

Binding of a drug to its substrate with high potency and selectivity is one of the requirements to prevent off-targeting and achieve a desired therapeutic effect with low doses if other pharmacokinetic parameters of the drug are favorable. Possible side effects associated with the drug could be reduced as well.

As explained previously, G-quadruplexes are important regulatory elements in the cellular processes. The Nobel laureate Aaron Klug had a famous quote over 30 years ago, "If G-quadruplexes form so readily in vitro, Nature will have found a way of using them in vivo". Recently, RNA G_2-quadruplexes have been found to regulate the polyamine biosynthesis in the cells in a negative feedback loop manner[173]. Many recent reports have suggested the role of G-quadruplex in the maintenance of cellular processes. Targeting G-quadruplex is a viable approach to disrupt the processes involved in pathological conditions.

G-quadruplexes are macro structures, and all of them share a common feature, a planar G tetrad. In general, the G-quadruplex ligands contain a planar aromatic moiety(s) to stack on, and the G-tetrads through π- π interactions and consequently stabilize the corresponding G-quadruplex. This feature is crucial to the binding selectivity for G-quadruplex over duplex DNA[80,174]. In addition, most of the G-quadruplex ligands have a permanent positive charge or a basic moiety that is charged under physiological conditions. The positively charged groups increase the binding affinity of the ligands to G-quadruplex DNA by electrostatic interactions with the negatively charged phosphate backbones. The relative number and disposition of the side chains have a profound effect on the binding affinity of the ligand to the G-quadruplexes.

Ligands with multiple positive charges have higher binding affinities to G-quadruplex DNA[125,142], for example, trisubstituted acridine (BRACO-19) containing four positive charges is a much stronger G-quadruplex binder than a di-substituted acridine derivative[130] containing three positive charges (**Figure 3.1**). However, there is one caveat. Positively charged groups, if not judiciously placed on the molecule, could decrease the binding selectivity for G-quadruplexes over duple DNA[28,107]. For example, TmPyP4, a potent, selective G-quadruplex ligand, has an extended aromatic surface for favorable π- π stacking interaction with the G-quadruplex but has a poor selectivity over duplex DNA[175] due to the presence of four positive charges. Replacing one charge with a neutral moiety has been shown to improve selectivity for G-quadruplex over duplex DNA[84]. Our group has also previously reported a potent G-quadruplex binding ligand (a thiazole orange-spermine conjugate) that inhibits human telomerase activity at sub-micromolar concentrations[176]. It binds tightly to telomeric G-quadruplex DNA but with limited selectivity for G-quadruplex over duplex DNA, probably due to its highly charged polyamine side chain.

BRACO-19
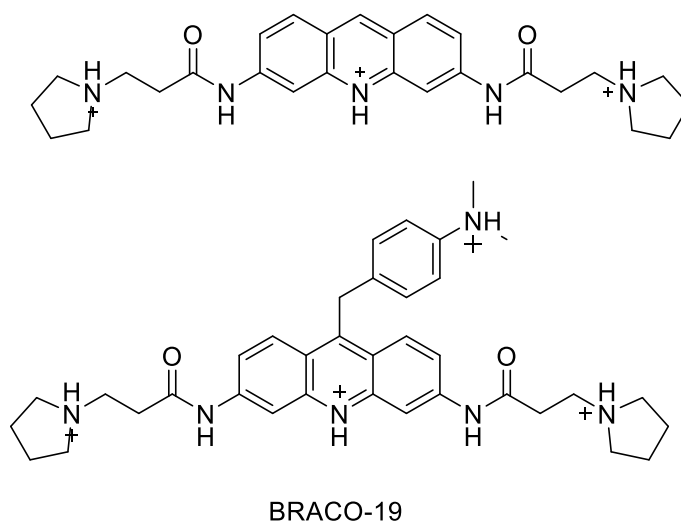
Figure 3.1. Structure of disubstituted acridine (top) and trisubstituted acridine (bottom).

On the other hand, a judicious disposition of positively charged side chains in a ligand can enhance its binding affinity without compromising the selectivity[79,92,105]. By positioning positively charged groups in a structured fashion, one can reduce non-specific binding between the ligands and DNA backbones. In search of such side chains, we came across the aminoglycoside (neomycin) scaffold. Neomycin has six amino groups on four rings spatially separated by the saccharide linkages between the rings **(Figure 3.2)**. The amino groups are positively charged or partially positively charged under physiological conditions[177]. Previous reports have suggested that neomycin binds preferentially to the A form of the nucleic acids[150] such as RNA and triplex DNA with a minimal affinity to duplex DNA even at significantly high concentrations[178]. In addition, neomycin binds to G-quadruplex DNA as a groove binder[166] with a relatively low binding affinity. The apparent binding selectivity of neomycin towards A-form nucleic acids should result from its unique spatial arrangement of multiple positive charges. We believe that by utilizing neomycin as the structured amino-containing side chain, we can develop novel and stronger G-quadruplex binding ligands.
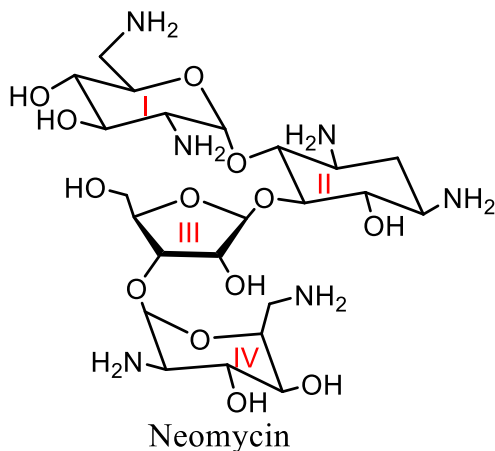
Figure 3.2. Structure of neomycin.

Previous studies from our lab reported several 5-substituted 1,10-phenanthroline derivatives as potent G-quadruplex ligands (**Figure 3.3**). In particular, derivatives with an amino side chain (piperidine, dimethylamine) showed enhanced binding affinity to G-quadruplex DNA[179]. The 1,10-phenanthroline moiety most likely binds with the end G-tetrad via π-π stacking interactions. Results from molecular modeling[180] and solution NMR spectroscopy by other groups have shown that 5 and 6 positions of 1,10 phenanthroline are located close to the grooves of G-quadruplexes when it stacks at the end G-tetrad of a G-quadruplex[138]. Therefore, the side chains at the 5 position of 1,10-phenanthroline derivatives should readily fit into the groove of G-quadruplex, enhancing the binding affinity of the ligand via a dual recognition mode.
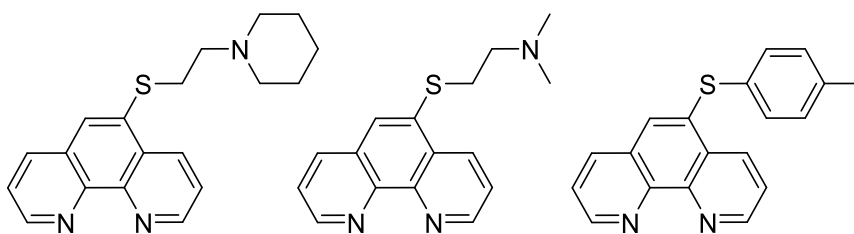
Figure 3.3. 5-substituted phenanthroline derivatives with piperidine, dimethylamine and 4-methylphenyl side chains.

Encouraged by these findings, we decided to develop novel ligands by tethering neomycin to the 5 position of 1,10-phenanthroline as a structured amino group side chain. The tethering position on neomycin was the 5" position of the ring III. Previous studies suggested that modification at this position has minimal influence on the binding affinity of neomycin to secondary nucleic acid structures[152].

In the present work, two phenanthroline-neomycin conjugates were synthesized, which differ in the length of the linker between the groove binder moiety (neomycin) and end stacking moiety (1,10-phenanthroline) (**Figure 3.4**). Their binding affinity and selectivity to G-quadruplex were determined with biophysical methods, including fluorescence intercalator displacement assay (FID), isothermal titration calorimetry (ITC), and thermal denaturation monitored by circular dichroism (CD) or UV. The binding stoichiometry was determined by ESI mass spectrometry, and plausible binding modes were estimated using molecular docking. In vitro telomerase inhibition of the ligands was investigated using a modified Telomerase Repeated Amplification Protocol (TRAP) assay.
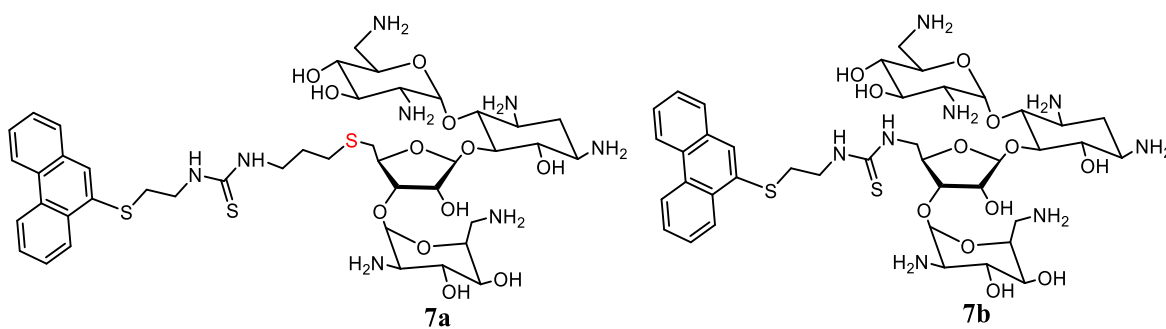
Figure 3.4. Structure of the proposed compounds, with longer linker between 1,10-phenanthroline and neomycin (**7a**), with longer linker between 1,10-phenanthroline and neomycin (**7b**).

3.2. Experimental Section

3.2.1. Materials and General Methods

All the chemicals for synthesis were purchased from Sigma-Aldrich or Fisher Scientific and used without further purification. DNA oligonucleotides were synthesized on an ABI 392 DNA/RNA synthesizer using standard phosphoramidite protocols and purified using polyacrylamide gel electrophoresis. All reagents and solid-support columns for DNA synthesis

were purchased from Glen Research. The concentrations of DNA solutions were determined by

UV spectroscopy, using the molar extinction coefficients (in units of mol of strand/L^{-1} cm^{-1})

obtained from OligoAnalyzer (www.idtdna.com). ε_{260} = 237,000 for **G1** [5'-

AGGG(TTAGGG)$_3$T-3']; ε_{260} = 259,100 for **S1** (5'-

ATGAAGGACGTAACCGGCTCTGAACG-3'); ε_{260} = 167,400 for **S2** (5'-

CGTTCAGAGCCGGTTACGTCCTTCAT-3'); ε_{260} = 238,300 for TS primer (5'-

AATCCGTCGAGCAGAGTT-3'); ε_{260} = 261,000 for RP (5'-

GCGCGGCTTACCCTTACCCTTACCCTAACC-3'); ε_{260} = 151,800 for NT (5'-

ATCGCTTCTCGGCCTTTT-3'); ε_{260} = 370,100 for TSNT (5'-

AATCCGTCGAGCAGAGTTAAAAGGCCGAGAAGCGAT-3'). ^{1}H NMR and ^{13}C NMR

spectra were collected on a JEOL ECA 600 MHz FT-NMR spectrometer. Mass spectra of the

synthesized compounds were analyzed on an Agilent 1290 UHPLC coupled to an Agilent 6230

Time-of-Flight mass spectrometer at Chevron, Richmond, California. Mass spectra of ligand-

DNA complexes were collected on a Thermo Fisher Orbitrap Fusion Tribrid Mass Spectrometer.

UV spectra were recorded on a Varian Cary 100 Bio UV-Vis spectrophotometer equipped with a

thermoelectrically controlled 6×6 cell holder. Circular dichroism spectra were recorded on a

JASCO J-810 spectropolarimeter using a quartz cuvette with a 1 mm or 1 cm optical path length.

Isothermal microcalorimetric measurements were performed on a TA Instruments Affinity ITC

LV. T4 polynucleotide kinase was obtained from New England Biolabs. [γ-^{32}P]-ATP was

purchased from MP Biochemicals. Quantification of 5' ^{32}P-labeled oligonucleotides was carried

out using a Storm 860 phosphorimager and ImageQuant 5.1 software. Cell medium and

supplements (fetal bovine serum, L-glutamine, antimycotic) were acquired from Invitrogen.

HeLa cells were maintained in advanced DMEM/F12 medium supplemented with 5% fetal

bovine serum, L-glutamine, and antimycotic at 37 °C in a humid atmosphere containing 5% CO_2.

DNA labeling was performed by incubating $[\gamma\text{-}^{32}P]$-ATP (30 μCi) and T4 polynucleotide kinase

(20 units) in the presence of an oligonucleotide (10 pmol) at 37°C for 30 min. Unreacted $[\gamma\text{-}^{32}P]$-

ATP was removed using a MicroSpin G-25 column.

3.2.2. Synthesis of Phenanthroline-neomycin Conjugates

The final compounds **7a** and **7b** are synthesized as described in the scheme 3.1 and 3.2.

The synthesis of compounds **2**-**4** was described in chapter 2.

Synthesis of 5a-L. In a 50-ml round-bottomed flask, freshly cut sodium metal (0.648 g,

28 mmol) was added to anhydrous ethanol (20 ml) and stirred at 0 °C for 30 mins. 2-

aminoethanthiol hydrochloride (1.518 g, 13.4 mmol) was added to the flask, and the reaction

mixture was stirred for 5 mins at rt. Compound **3** (460 mg, 0.31 mmol) dissolved in dry ethanol

(10 ml) was added to the flask using a syringe. The reaction mixture was stirred at rt for 5 h.

The competition of the reaction was monitored with TLC [R_f = 0.43 (CH_3OH:CH_2Cl_2:NH_4OH =

10:90:0.5 (v/v/v), silica gel 60 F_{254}, EMD Millipore]. The reaction mixture was concentrated

under vacuum. Flash chromatography of the residue [silica gel, CH_3OH:CH_2Cl_2:NH_4OH =

5:95:0.5 (v/v/v)] yielded compound **5aL** (324mg, 82.3 %) as a white solid.

Synthesis of 5b-S. In the hydrogenation flask, compound **4** (160 mg, 0.129 mmol) was

dissolved in 15ml of anhydrous ethanol. 40mg of the Pd/C catalyst was added to the flask, and

the reaction mixture was shaken overnight under 40 Psi H_2 pressure. The completion of the

reaction was monitored for the disappearance of compound **4** using TLC. The mixture is filtered

through celite, and the residue is concentrated under vacuum to yield (130mg, 83 %) compound

5bS as a white solid.

Synthesis of 9. In a 25-ml round-bottomed flask, β-aminoethanthiol hydrochloride (68 mg, 0.6 mmol) was dissolved in freshly prepared sodium ethoxide in ethanol (0.1M, 4 ml). 1,10-Phenanthroline epoxide **8** (100 mg, 0.510 mmol) dissolved in anhydrous ethanol (4 ml) was charged into the flask with a syringe. The reaction mixture was stirred at 50 °C for 10 h, and the residue was concentrated under vacuum. [R_f =0.22, $CH_3OH:NH_4OH$ = 20:1 (v/v), silica gel 60 F_{254}, EMD Millipore]. Flash chromatography of the residue [silica gel, $CH_3OH:NH_4OH$ (20:1)] yielded **9** (110 mg, 84.6 %) as a slightly yellow oily liquid.

Synthesis of 9NCS. In a 25-mL round-bottomed flask, compound **9** (80.0 mg, 0.314 mmol) and DMAP (catalytic amount) were dissolved in DMF (10 ml) under N_2. A DMF solution (1 ml) of TCDP (109 mg, 0.471 mmol) was added dropwise into the flask using a syringe, and the resulting mixture was stirred overnight at room temperature. The completion of the reaction was monitored by TLC [R_f = 0.52, $CH_3OH:CH_2Cl_2:NH_4OH$ = 8:92:0.5 (v/v), silica gel 60 F_{254}, EMD Millipore] under UV. The reaction mixture was concentrated under vacuum. Flash chromatography of the residue [silica gel, $CH_3OH:CH_2Cl_2$ = 5:95 (v/v)] yielded compound **9NCS** as a pale yellow oil (52.0 mg, 55.5%). IR (KBr, cm^{-1}) 3,173, 3,018, 2,924, 2,187, 2,122, 2086, 1674, 1652, 1,590, 1,560, 1,502, 1,418, 1,344, 1,292, 1,284, 1,240, 1,142, 1,103, 1,079, 942, 874, 824. ^1H NMR (600 MHz, CD_3OD): δ 9.1 (dd, J = 4.2, 1.8 Hz, 1H), 9.03 (dd, J = 4.8, 1.8 Hz, 1H), 8.89 (dd, J = 7.8, 4.2 Hz, 1H), 8.38 (dd, J = 7.8, 4.2 Hz, 1H), 8.11 (s 1H), 7.82 (q, J = 4.2 Hz, 1H), 7.73 (q, J = 4.2 Hz, 1H), 3.79 (t, J = 6 Hz, 2H), 3.40 (t, J = 6.6 Hz, 2H). ^{13}C NMR (150 MHz, CD_3OD) δ 151.16, 151.10, 146.91, 146.14, 137.35, 135.20, 133.55, 132.34, 130.36, 129.92, 125.13, 124.79, 45.48, 35.21. HRMS (ESI) calcd for $C_{15}H_{13}N_3S_2$ $[M + H]^+$: 298.0473, found 298.0459.

Synthesis of 6a-L. In a 25-ml round-bottomed flask, **5a-L** (50 mg, 0.0385 mmol) was dissolved in anhydrous pyridine (5 ml). An anhydrous pyridine solution (1 ml) of compound **9NCS** (9 mg, 0.0302 mmol) was added dropwise into the flask with a syringe. The reaction mixture was stirred overnight at room temperature under N_2 and concentrated under vacuum. [R_f = 0.24, $CH_3OH:CH_2Cl_2:NH_4OH$ = 8:92:1 (v/v/v), silica gel 60 F_{254}, EMD Millipore]. Flash chromatography of the residue [silica gel, $CH_3OH:CH_2Cl_2:NH_4OH$ = 8:92:1 (v/v/v)] yielded compound **6a-L** (26.5 mg, 56 %). IR (KBr, cm^{-1}) 3430, 2976, 2930, 1691, 1515, 1423, 1367, 1248, 1168, 1038. ^{1}H NMR (600 MHz, CD$_3$OD):): δ 9.07 (dd, J = 4.2, 1.8 Hz, 1H), 8.97 (dd, J = 4.2, 1.8 Hz, 1H), 8.76 (dd, J = 8.4, 1.8 Hz, 1H), 8.36 (dd, J = 7.8, 1.8 Hz, 1H), 8.25 (s, 1H), 7.78 (q, J = 4.2 Hz, 1H), 7.70 (q, J = 4.2 Hz, 1H), 6.41 - 6.75 (m, 2H), 5.40 (s, 1H), 5.12 (s, 1H), 4.91 (s, 1H), 4.13 - 4.37 (m, 2H), 3.99 - 4.11 (m, 1H), 3.63 - 3.98 (m, 6H), 3.36 - 3.62 (m, 6H), 2.67 - 2.99 (m, 4H), 1.87 - 2.00 (m, 1H), 1.12 - 1.54 (m, 54H). HRMS (ESI) calcd for $C_{68}H_{107}N_{10}O_{24}S_2$ [M + H]$^+$: 1571.6935, found 1571.6856.

Synthesis of 6b-S. In a 25-ml round-bottomed flask, **5b-S** (46.8 mg, 0.0385 mmol) was dissolved in anhydrous pyridine (5 ml). An anhydrous pyridine solution (1 ml) of compound **9NCS** (10.4 mg, 0.0350 mmol) was added dropwise into the flask with a syringe. The reaction mixture was stirred overnight at room temperature under N_2 and concentrated under vacuum. [R_f = 0.30, $CH_3OH:CH_2Cl_2:NH_4OH$ = 8:92:0.8 (v/v/v) silica gel 60 F_{254}, EMD Millipore], Flash chromatography of the residue [silica gel, $CH_3OH:CH_2Cl_2:NH_4OH$ = 8:92:0.8 (v/v/v)] yielded compound **6b-S** (23.7 mg, 40.7%). IR (KBr, cm^{-1}) 3,340, 2,976, 2,920, 2,851, 1685, 1521, 1,384, 1,252, 1,174, 1,031. ^{1}H NMR (600 MHz, CD$_3$OD): ^{1}H NMR (600 MHz, CD$_3$OD): δ 9.11 (dd, J = 4.2, 1.8 Hz, 1H), 9.00 (dd, J = 4.2, 1.8 Hz, 1H), 8.83 (dd, J = 8.4, 1.8 Hz, 1H), 8.41 (dd, J = 7.8, 1.8 Hz, 1H), 8.35 (s, 1H), 7.82 (q, J = 4.2 Hz, 1H), 7.76 (q, J

= 4.2 Hz, 1H), 6.59 - 6.94 (m, 2H), 5.51 (s, 1H), 5.03 (s, 1H), 4.24 - 4.55 (m, 2H), 3.68 - 4.21 (m, 8H), 3.69 - 3.79 (m, 2H), 3.45 - 3.64 (m, 6H), 3.06 - 3.22 (m, 3H), 1.88 - 1.99 (m, 1H), 1.27 - 1.57 (m, 54H). HRMS (ESI) calcd for $C_{68}H_{107}N_{10}O_{24}S_2$ [M + H]$^+$: 1511.6901, found 1511.6894.

Synthesis of 7a and 7b. 5-10 mg of **6a-L** (10 mg) or **6b-S** (6mg) is dissolved in trifluoroacetic acid (2 ml) in a small RBF and stirred for 15 mins. The reaction mixture was concentrated under vacuum. The residual was dissolved in water, and the aqueous solution is washed with DCM (3 x 3 ml). Lyophilization of the aqueous layer yielded compound **7a** (92.4 %) and **7b** (78.9 %).

7a. IR (KBr pellet, cm^{-1}) 3436, 2919, 2847, 1623, 1503, 1412, 1043. ^{1}H NMR (600 MHz, D$_2$O) : δ 9.03 (d, J = 4.2 Hz, 1H), 8.84 - 8.99 (m, 2H), 8.69 (d, J = 7.8 Hz, 1H), 8.07 (s, 1H), 7.99 (dd, J = 8.4, 4.8 Hz, 1H), 7.93 (dd, J = 8.4, 4.8 Hz, 1H), 5.89 (d, J = 4.2 Hz, 1H), 5.24 (d, J = 3 Hz, 1H), 5.12 (s, 1H), 4.56 - 4.69 (m, 2H), 4.10 - 4.36 (m, 4H), 4.05 (t, J = 3 Hz, 1H), 3.93 (t, J = 9.6 Hz, 1H), 3.85 (t, J = 9.6 Hz, 1H), 3.60 - 3.81 (m, 4H), 3.53 (t, J = 9.6 Hz, 1H,), 3.15 - 3.49 (m, 12H), 3.09 (q, J = 7.2 Hz, 1H), 2.91 - 3.01 (m, 1H), 2.65 (dd, J = 15, 7.8 Hz, 1H), 2.42 - 4.62 (m, 2H), 2.33 (dt, J = 12.6, 4.2 Hz, 1H), 1.74 (q, J = 12.6 Hz, 1H). HRMS (ESI) calcd for $C_{40}H_{64}N_{10}O_{12}S_3$ [M+2H]$^{2+}$: 486.1910, found 486.1933.

7b. IR (KBr pellet, cm^{-1}) 3431, 2920, 2850, 1624, 1500, 1384, 1040. ^{1}H NMR (600 MHz, D$_2$O): δ 9.09 (d, J = 4.3 Hz, 1H), 8.89 - 9.03 (m, 2H), 8.76 (d, J = 8.4 Hz, 1H), 8.18 (s, 1H), 8.02 (dd, J = 8.4, 5.4 Hz, 1H), 7.97 (dd, J = 8.4, 4.8 Hz, 1H), 5.90 (d, J = 3 Hz, 1H), 5.27 (d, J = 3 Hz, 1H), 5.12 (s, 1H), 4.65 (s, 2H), 4.31 (t, J = 4.8 Hz, 1H), 4.02 - 4.26 (m, 3H), 3.97 (t, J = 10.2 Hz, 1H), 3.50 - 3.92 (m, 7H), 3.05 – 3.50 (m, 11H), 2.35 (dt, J = 8.4, 2.4 Hz, 1H), 1.75 (q, J = 13.2 Hz). HRMS (ESI) calcd for $C_{38}H_{60}N_{10}O_{12}S_2$ [M+H]$^+$: 911.3755, found 911.3754. Yield was 5.00 mg, 78.9 %.

3.2.3. MS Analysis Procedures of Synthesized Compounds

The samples were diluted in methanol to approximately 1 M concentration. A sample solution (1 µL) was introduced by flow injection analysis (FIA) without a chromatographic column except for compound **7a/7b** where 2 µL was injected. The solvent system was 1:1 methanol:water with 0.1% formic acid. Ions were delivered to the mass spectrometer with an electrospray ionization (ESI) source. The nebulizer pressure was 50 psig. at 325 °C and nitrogen drying gas flow was 12 mL/min. The capillary voltage was 3000 V for all samples. For compound **6** (Boc protected), the fragmentor voltage was 150 V. Data were collected in the m/z range of 100-3200. Spectra were collected at a rate of 2 spectra/s allowing for 4903 transients/spectrum. For compound **7**, the fragmentor voltage was 70 V. Data were collected in the m/z range of 500-2000. Spectra were acquired at a rate of 1 spectra/s allowing for 9978 transients/spectrum.

3.2.4. Preparation of G-quadruplex DNA and Duplex DNA for CD and UV Experiments

The G-quadruplex DNA strand (**G1**, 100 µM) was dissolved in a mixture (100 µL) of lithium cacodylate buffer (10 mM, pH 7.0) and KCl (50 mM) and incubated at 90 °C for 10 min, slowly cooled to 25°C, and incubated at 4 °C overnight. Similarly, duplex **D1** (complementary strands of S1 and S2, 100 µM each) was dissolved in a mixture of lithium cacodylate buffer (10 mM, pH 7.0) and KCl (150 mM) and incubated at 90 °C for 10 min, slowly cooled to 25°C, and incubated at 4 °C overnight. All solutions were stored at 4 °C for further use.

3.2.5. Preparation of G-quadruplex DNA for ESI Mass Spectrometry Experiments

G1 (100 µM) in 150 mM NH_4OAc solution was heated to 95 °C for 5 min followed by cooling slowly to room temperature and incubated at 4°C overnight.

3.2.6. Thermal Denaturation Analysis Monitored by CD

The mixture solutions (200 µL) of **G1** (5 µM) in lithium cacodylate buffer (10 mM, pH 7.0) and KCl (50 mM) in the absence or presence of a ligand (Phen, neomycin, and **7a/7b**, 10 µM) were prepared. The CD spectra were recorded at 290 nm as a function of increasing temperature (25-85 ˚C) at a heating rate of 0.5 ˚C/min. Melting profiles were converted into a folded faction based on a published procedure, and the melting temperatures were determined at the 0.5 folded fraction.

3.2.7. CD Titration

Aliquots of a stock solution containing compound **7a/7b** were gradually added into a solution (2 mL) of **G1** (5 µM) in lithium cacodylate (10 mM, pH 7.0) and 50 mM KCl at 20 °C. The final molar ratios of **7** to **G1** were varied from 0 to 5. After each addition, the solution was gently mixed and incubated for 5 min for equilibrium before collecting the spectrum. The overall volume change during the titration was kept less than 1% of the initial volume. CD spectra were recorded as a function of wavelength (200-400 mm, 100 nm/min scan speed, 0.5 nm data pitch) using a 1 cm pathlength quartz cuvette. Each spectrum was averaged over three scans.

3.2.8. Thermal Denaturation Analysis Monitored by UV

The mixture solutions (1 mL) of **D1** (1 µM) in lithium cacodylate buffer (10 mM, pH 7.0) and KCl (150 mM) in the absence or presence of a ligand (Phen, neomycin, and **7a/7b**, 10 µM) were prepared. The UV melting curves were recorded at 260 nm as a function of temperature (20-85 °C, heating rate: 0.5 °C/min). Melting profiles were converted into a folded faction, and the melting temperatures were determined at the 0.5 folded fraction.

3.2.9. Fluorescent Intercalator Displacement Assay

A reaction solution containing 0.25 µM DNA (G-quadruplex (**G1**) or duplex (**D1**) in lithium cacodylate (10 mM, pH 7.0) and KCl (50 mM) was incubated with 0.5 µM thiazole orange (TO) for 30 min in the dark. Aliquots of a ligand were added to this mixture solution and equilibrated for 5 min in the dark before measurement. The fluorescence spectra (Ex: 501 nm, slit width: 5.0 nm, scan speed 120 nm/min) were recorded from 510-650 nm after each addition. The percent fluorescence displacement (%F) was calculated at λ_{max} by using the formula %FI. displacement=$100-[(FA/FA_0)\times100]$, where FA_0 is the initial fluorescence intensity in the absence of a ligand, and FA is the fluorescence intensity upon addition of a ligand.

3.2.10. Isothermal Titration Calorimetry (ITC)

In a typical experiment, 2 µL aliquots of compound **7b** (200 µM) were injected from a 264 µL rotating syringe (125 rpm) into an isothermal sample cell containing 185 µL of a telomeric G-quadruplex (**G1**) (20 µM/strand) or a duplex DNA (**D1**) solution (10 µM/strand) at 25 °C. The corresponding control experiment was carried out by injecting 2 µL aliquots of compound **7b** (200 µM) into a solution of buffer alone. The duration of each injection was 4 s, and the delay between injections was 120 s. The initial delay prior to the first injection was 100 s. A heat burst curve (microcalories per second vs. seconds) was generated from each injection. The area under each curve was calculated using the NanoAnalyzer software (Version 3.8.0) to yield a measure of the heat associated with that injection. The heat of ligand binding to DNA associated with that injection was obtained by subtracting the heat associated with each ligand-buffer injection from the corresponding heat associated with each ligand-DNA injection.

3.2.11. Preparation of Cell Lysate for TRAP Assay

Approximately 1 x 10^7 HeLa cells were suspended in CHAPS buffer (1 mL) and incubated on ice for 30 min. The supernatant was collected after spinning the sample at 12,000 x g for 10 min and stored at -80°C.

3.2.12. TRAP Assay

The 5′-[^{32}P] end-labeled TS primer (0.1 µg) was incubated in a reaction mixture (50 µL) [buffer: Tris-HCl (20 mM, pH 8.3), $MgCl_2$ (1.5 mM), KCl (63 mM), 0.05% Tween 20, EGTA (1 mM)] with cell lysate (1 µL), RP (0.1 µg, reverse primer for telomerase extended products), NT (0.1 µg, primer for polymerase as internal standard), TSNT (0.01 amol, template for NT primer), 2 units of Taq polymerase, dNTP mix (50 µM), and ligand (3.125-100 µM) at 30°C for 30 min and 94 °C for 2 min to deactivate telomerase. The reaction mixture was then incubated using a 3-step PCR method (94 °C for 20 s, 59 °C for 30 s, and 72 °C for 30 s) with a total of 25 cycles. The residues were concentrated and resuspended in formamide loading buffer (7 µL). Analytical separation was carried out using 10% denaturing polyacrylamide gel [5% crosslink, 45% urea w/w].

3.2.13. Docking Procedure

The geometry optimization of compound **7b** was carried out at the hybrid density functional B3LYP level with 6-31G(d) basis set using the Gaussian'09 program package. The Natural bond order (NBO) analysis was carried out to calculate the atomic charges on the optimized geometries. **G1** structure for docking simulation was modeled after the solution structure of the complex of platinum (II) based ligand with human telomeric G-quadruplex (PDB id: 5Z80) whose structure was obtained from RCSB protein data bank. In the complex structure, the bound ligand molecule was removed. The resulting template has two more A residues at the

5'-end and one additional A at the 3'-end in comparison to **G1**. In the NMR structure, Adenine-25 was mutated to thymine, and residues adenine-1, 2, and 26 were truncated to yield the structure of **G1** for docking studies. The flexible ligand docking experiments were carried out using the DOCK 6.9 program. Analyses of docked structures and docked image creations were carried out using the UCSF Chimera program.

3.2.14. ESI-MS Analysis of Complex Formation Between G1 and 7a/7b

G1 (20 µM) and compound **7a/7b** at various concentrations (0, 20, 40, 60, and 80 µM) were mixed as 50% methanol-water solutions. The mixed solutions were injected into a Thermo Fisher Orbitrap Fusion Tribrid Mass Spectrometer at a rate of 5 µL/min. The ion spray voltage was set to -2,250 V and ion transfer tube temperature to 300 °C. Nitrogen sheath and auxiliary gas flows were 45 and 0 arbitrary units, respectively. MS spectra were recorded in the m/z range from 1400 to 1800 with an average of 150 scans. Data were collected and analyzed using the Xcalibur 4.0 software.

3.3. Results and Discussion

3.3.1. Synthesis of Phenanthroline-neomycin Conjugates

As shown in the synthesis scheme 3.1 and 3.2, neomycin precursors were synthesized based on the published procedures[158,164]. The C-5'' position of the ribofuranose ring of neomycin was used for conjugation as this site has a minimal effect on the structure of the neomycin and its ability to bind to nucleic acids[154,158]. The six amino groups of neomycin were protected with di-tert-butyl dicarbonate in 1:1 DMF-H_2O to yield compound **2**. Only the 5'' position of ribofuranose ring of neomycin has a primary hydroxyl group, which was selectively tosylated using 2,4,6-triisopropylbenzenesulfonyl chloride (TIPS-Cl) in dry pyridine at room temperature to yield **3**. The TIPS group was replaced with sodium azide via an S_N2 reaction

mechanism in DMF at elevated temperature to yield compound **4**, which was further converted

to **5b-S** via catalytic hydrogenation utilizing Pd/C as the catalyst. The other analogue **5a** was

synthesized by reacting compound **3** with β-mercaptoethanamine hydrochloride using sodium

ethanol in ethanol at room temperature. Compound **9** was synthesized by reacting 5,6-epoxy-

5,6-dihydro-[1,10]phenanthroline **8** with β-mercaptoethanamine hydrochloride using sodium

ethanol in ethanol[181]. Reaction of compound **9** with 1,1'-thiocarbonyldi-2(1H)-pyridone

(TCDP) in DMF yielded compound **9NCS**. Compound **9NCS** was coupled with compound **5a-L**

and **5b-S** through a thiourea linkage to yield a boc-protected phenanthroline neomycin

conjugates **6a-L** and **6b-S**, respectively. Deprotection of amine groups of **6a-L** and **6b-S** was

accomplished successfully with trifluoroacetic acid to yield final compounds **7a** and **7b**.
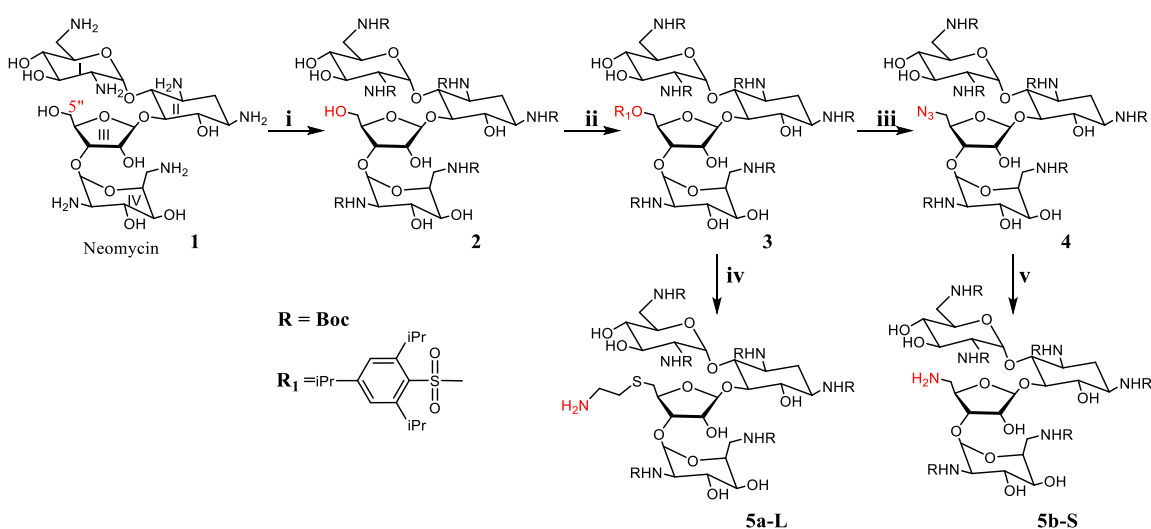
Scheme 3.1. Synthesis of neomycin amine at C5" of ring III. Reagents and conditions: (i) (Boc)$_2$O, DMF-H$_2$O, 60 °C; (ii) 2,4,6-triisopropylbenznenesulfonyl chloride, pyridine, r.t.; (iii) NaN$_3$, DMF, 90 °C; (iv) Sodium ethoxide, ethanol, β-mercaptoethanamine; (68%) (v) Pd/C, H$_2$, r.t.
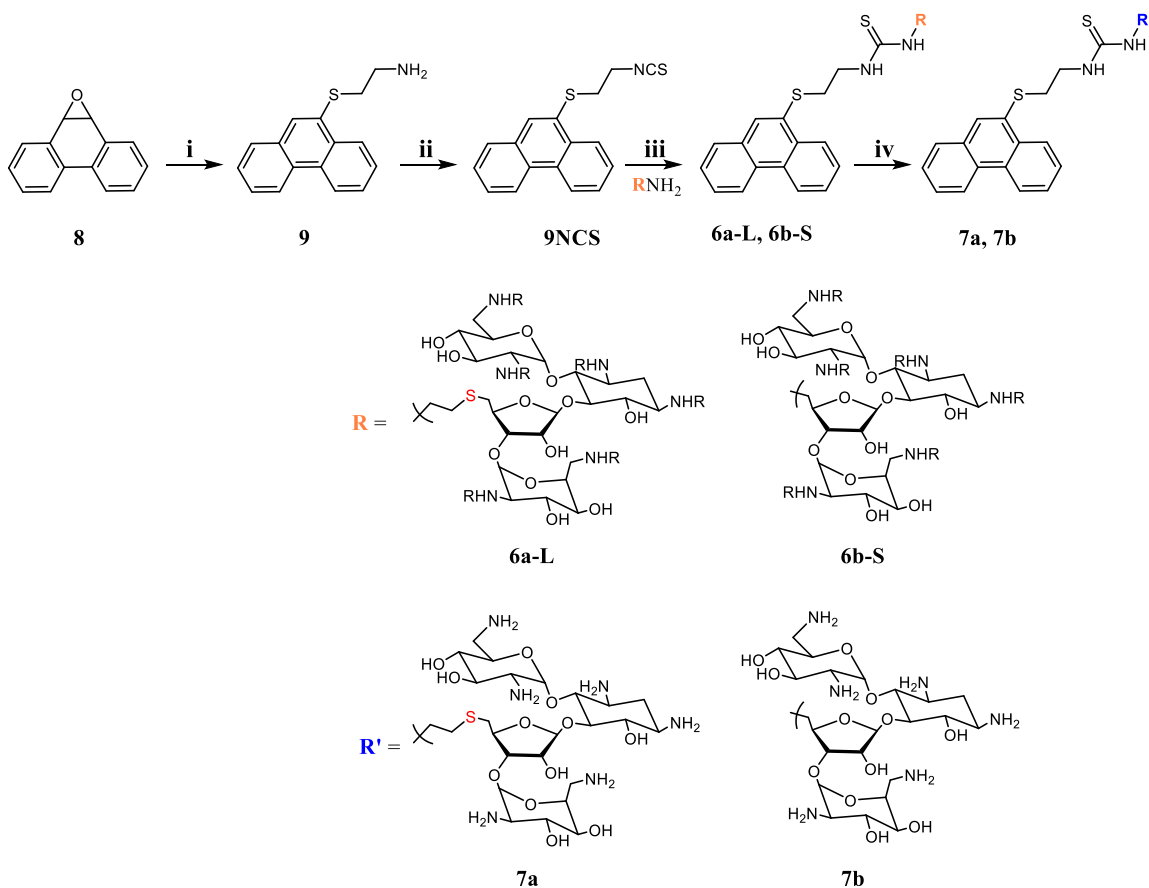
Scheme 3.2. Synthesis of phenanthroline derivatives **9NCS**, **7a** and **7b**. Reagents and conditions: i-iv: (i) Sodium ethoxide, ethanol, β-mercaptoethanamine. (ii) TCDP, DMF, r.t., overnight stir (iii) dry pyridine (iv) TFA, 15 mins.

3.3.2. Phen-neo Conjugates Bind With High Affinity to G-quadruplex (G1) DNA

The binding of a telomeric G-quadruplex (**G1**) or a duplex (**D1**) with phenanthroline-neomycin conjugates **7a** and **7b**, phenanthroline (Phen), and neomycin was first assessed using the Fluorescence Intercalator Displacement (FID) Assay[182]. This assay is based on the phenomenon that binding of thiazole orange (TO) with nucleic acids can drastically increase its quantum yield of dye from 2×10^{-4} to 0.1, about 1000 times. Free TO in solution does not fluoresce. In a typical FID experiment, TO is first incubated with DNA to turn on fluorescence. Addition of a ligand of interest into the TO-DNA solution can gradually displace TO out of

DNA, resulting in the loss of fluorescence. The concentration of the ligand required to displace 50% of pre-bound TO from DNA is defined as DC_{50}. Low DC_{50} values represent the high binding affinity of the ligand towards DNA. For G-quadruplex and duplex DNA, this value is designated as $^{G4}DC_{50}$ and $^{ds}DC_{50}$, respectively. It is noteworthy that TO is non-specific DNA binding dye; therefore, it does not selectively bind to different types of nucleic acids[183]. The FID assay has been used for the detection of ligands binding to non-canonical structures in the genomic DNA[184] in conjunction with other biophysical methods[182]. Over the years, it has been successfully used to determine potent and selective G4 ligands that inhibit human telomerase activity[176].

Initial studies were carried out by incubating **7a** or **7b** with Ni^{2+} ion because coordination of phenanthroline with metal ions can drastically enhance the stacking surface of the ligand. Our results did not show enhanced binding of **7a** and **7b** to **G1** compared with that of **7a** and **7b** alone (**Figure 3.5b**). Neomycin is known to coordinate with metal ions[185]. It is possible that Ni^{2+} could not effectively coordinate with the Phen moiety of **7a** and **7b** due to the interference of neomycin at the concentration range used. Therefore, due to this uncertainty, we only report the binding of **7a** and **7b** (without Ni^{2+}) with G-quadruplex DNA in the following section. As shown in **figure 3.5a**, the corresponding rank order of the $^{G4}DC_{50}$ values was **7b** < **7a** < neomycin << Phen. The $^{G4}DC_{50}$ value for **7b** (0.9 µM) was impressively low, 50-fold less than that of neomycin (49.8 µM) and at least 500-fold less than that of phenanthroline alone (> 500 µM). On the other hand, **7a**, a longer linker analog of **7b**, had a $^{G4}DC_{50}$ value of 1.32 µM. The low $^{G4}DC_{50}$ values for **7b** showed a stronger displacement of TO from **G1** by **7b**, suggesting that the structured amino-containing binding motif (neomycin) can significantly enhance the binding with G-quadruplex DNA. Using the same conditions, the $^{ds}DC^{50}$ values for these compounds

were also measured with a DNA duplex (**D1**). The $^{ds}DC^{50}$ values for **7a**, **7b,** and neomycin are 2.78, 3.4 µM, and 5.7 µM, respectively. The relative selectivity of G-quadruplex over duplex DNA was determined by the ratio of $^{ds}DC_{50}/^{G4}DC_{50}$. The rank order for such selectivity was **7b** (3.8) > **7a** (2.0) > neomycin (0.11). It is clear that compound **7b** binds more preferentially to G4 than duplex DNA as compared to **7a**. Interestingly, neomycin has a smaller $^{ds}DC^{50}$ value (5.7 µM) than its $^{G4}DC_{50}$ value (49.8 µM). Phen binds to both duplex and G-quadruplex DNA very weakly based on the FID data. Both of the $^{G4}DC_{50}$ and $^{ds}DC^{50}$ values for Phen are over 500 µM; therefore, its selectivity of G4 over duplex DNA is negligible. An FID experiment of a mixture of neomycin and Phen (1:1 molar ratio) was also carried out. The $^{G4}DC_{50}$ value for the mixture is 9.0 µM, much smaller than those of neomycin (49.8 µM) and Phen (> 500 µM), indicating that the simultaneous binding of neomycin and Phen with G-quadruplex DNA has a synergistic effect. As a comparison, this value is 10-fold greater than that of **7b** (0.9 µM) and 6.8-fold greater than that of **7a**. Hence, conjugation of neomycin with Phen via covalent bonds provides a more significant enhancement of binding with G-quadruplex DNA than merely mixing two ligands. It is noteworthy that **7b** shows more significantly enhanced binding with G-quadruplex DNA than duplex DNA. The $^{ds}DC^{50}$ value (4.5 µM) of the mixture of neomycin and Phen is similar to the values of neomycin (5.7 µM) and **7** (3.4 µM). The binding of duplex DNA by **7b** could mainly result from the non-specific binding by electrostatic interactions of cationic charges of the neomycin moiety.
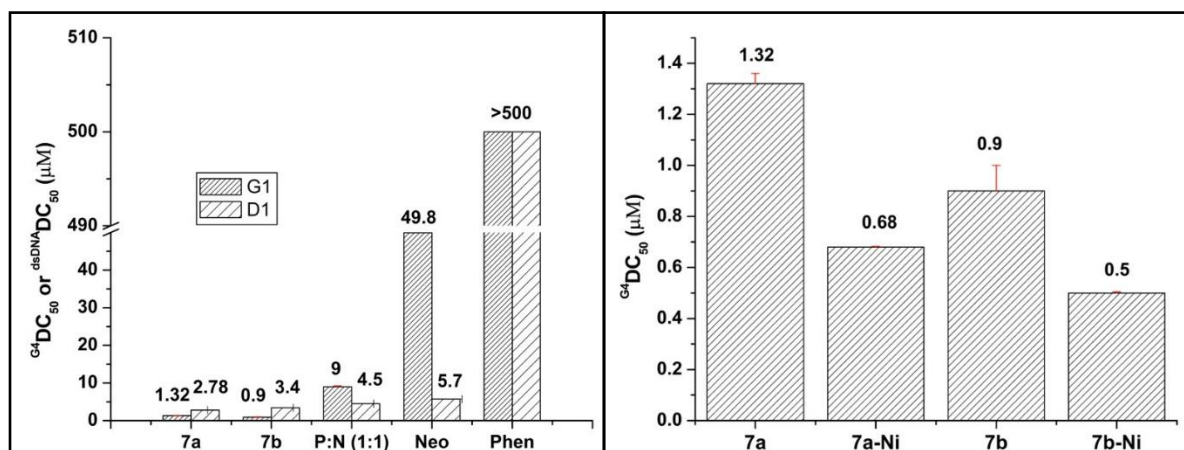
Figure 3.5. a) $^{ds}DC_{50}$ and $^{G4}DC_{50}$ values for compounds **7a** and **7b**, neomycin (neo), Phen, a mixture of Phen and neomycin (1:1 molar ratio). b) the $^{G4}DC_{50}$ values of **7a**, **7a-Ni**, **7b,** and **7b-Ni** from the FID assay. **7a-Ni** and **7b-Ni** represent a mixture of Ni^{2+} with **7a** and **7b** in a 1:2 ratio, repectively. Buffer: 10 mM lithium cacodylate, 50 mM KCl, pH 7.0.

3.3.3. Compound 7b Binds and Stabilizes Telomeric Hybrid G-quadruplex Without Producing Any Change in Conformation of the G-quadruplex

Circular dichroism spectroscopy is a useful technique to study the conformations of biomolecules. For instance, it has been extensively used to study the conformation[186] and handedness of the G-quadruplexes[187]. Human telomeric G-quadruplex can adopt a variety of conformations depending upon the metal ion[165]. In the presence of K^+ ions, a hybrid topology[188] is observed, which typically shows a maximum (+) around 290 nm, a shoulder (+) centered around 260 nm, a weak band (+) with a maximum around 250 nm, and a band (-) with a maximum around 240 nm. Whereas in Na^+ it adopts a basket shape topology with a maximum band (+) around 290 nm and a maximum (-) band at 260 nm. The change in intensity of the bands along with their red or blue shift on ligand binding can provide useful information about the process of ligand binding.

Thermal denaturation monitored by CD is used to study the stabilizing influence of a ligand on G-quadruplex DNA[189]. The ability of **7b** to stabilize **G1** was confirmed using thermal denaturation of **G1** monitored by circular dichroism at 290 nm. As shown in **figure 3.6a** and 2b, the representative folded fractions of **G1** in the absence and presence of a ligand (10 μM) are plotted as a function of temperature. The melting temperature (T_m) of **G1** was measured to be 61.0 ± 0.6 °C. In the presence of neomycin or Phen at 10 μM, the T_m values were 60.6 ± 0.3 °C and 60.8 ± 0.5 °C, respectively. Both neomycin and Phen could not stabilize **G1** under this condition. Phen did not increase the T_m value of **G1**, which is consistent with the previous studies[179]. On the other hand, **7a** destabilizes **G1** by 4 °C. In the presence of **7b** at 10 μM, a shorter linker analog of **7a**, the T_m value was 66.6 ± 0.3 °C, a 5.6 °C increment as compared with that of **G1**. Further increase in the concentration of **7b** did not increase the T_m of **G1** significantly (**Figure 3.6c**). When incubating **G1** with neomycin (10 μM) and Phen (10 μM), no increment of T_m (61.0 ± 0.8 °C) was observed. Because both neomycin and phenanthroline alone were ineffective in stabilizing the **G1**, therefore, the stabilization effect of **7b** toward **G1** must be resulting from a synergistic enhancement of binding affinity by conjugating neomycin and phenanthroline together. Similar results were obtained from thermal denaturation monitored by UV at 295 nm (**Figure 3.6d**).
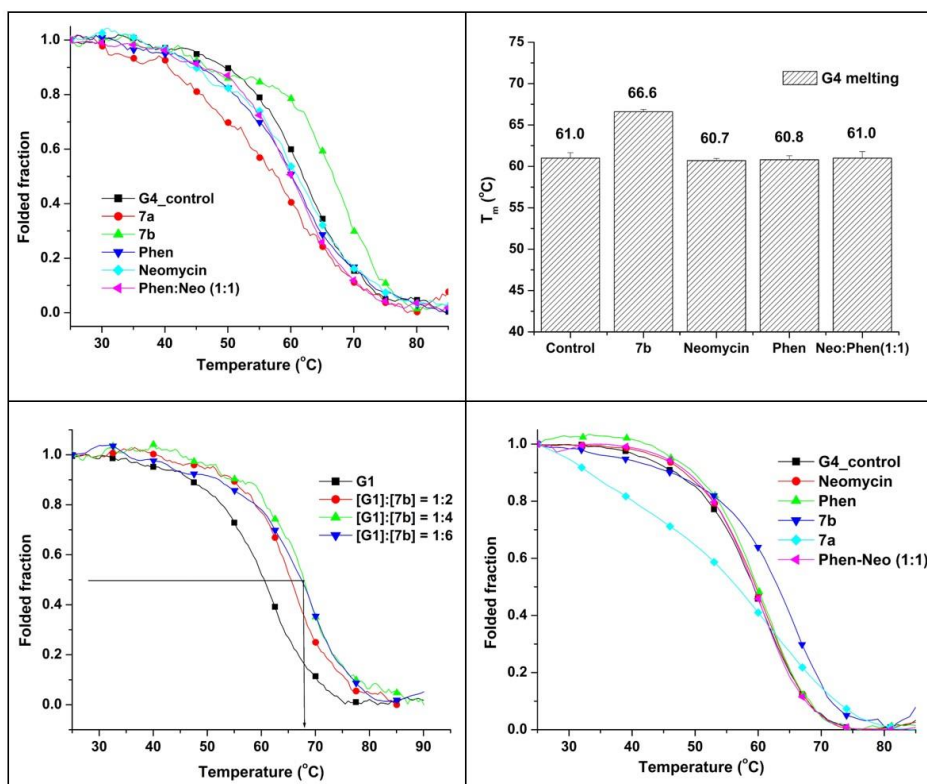
Figure 3.6. a) CD thermal denaturation of **G1** in the presence of phen-neo conjugates, phenanthroline, neomycin, and phen:neo (1:1). b) T_m of **G1** in the presence of the ligands. c) Effect of molar concentration of **7b** on T_m of **G1**. d) UV thermal denaturation of **G1** in the presence of phen-neo conjugates, phenanthroline, neomycin, and phen:neo (1:1).

The difference in thermal stabilization of **G1** by **7a** and **7b** prompted us to investigate their interactions with CD titration experiments. As shown in **Figure 3.7b**, **G1** in K^+ adopts a hybrid-type topology with a maximum (+) around 292 nm, a shoulder (+) centered around 270 nm, a weak band (+) with a maximum around 250 nm, and a band (-) with a maximum around 235 nm, which is consistent with previous reports.

In general, titration of **G1** with **7b** did not lead to an overall topological change of **G1**. A hybrid conformation was conserved even at the 5:1 ligand to DNA ratio. When the ligand to DNA molar ratio was below 3, a gradual increase in the intensity of the band (-) at 235 nm, the

band (+) at 250 nm, and the band (+) around 292 nm was observed, respectively, suggesting that a further stabilization of the folded hybrid conformation. When the ligand to DNA molar ratio was above 3, a slight decrease in the intensities of these three bands occurred. A redshift of the band (+) at 293 up to 300 nm was clearly observed. It is known that the lateral or diagonal loops of G-quadruplexes greatly affect the shape of their CD spectra. At high ligand to DNA molar ratios, it is plausible that the neomycin moiety of **7** could interact with the loops of **G1** via non-specific interactions, causing the observed shift in band maxima (**Figure 3.7b**).

By contrast, titration of **7a** with **G1** showed a distinct difference (**Figure 3.7a**). At the ligand to DNA ratio of 0.5, an increase in the intensity of band (+) around 292 nm along with the loss of shoulder peak at 270 nm and a slight dip of the peak at 260 nm was observed. At the ligand to DNA molar ratio of 3, the intensity of the band around 292 nm increased along with the redshift of the band, and the intensity of the peak (-) at 260 nm also deepened, suggesting the topology of **G1** was more close to the antiparallel G-quadruplex [the maximum (+) around 292 nm and a band (-) with a maximum around 260 nm]. When the ligand to DNA molar ratio was above 3, a decrease in the intensity of the band (+) around 300 nm and concomitant increase in the maximum (-) around 260 nm was observed probably due to the non-specific interactions between the ligand and DNA at higher concentrations. The transient increase in the antiparallel topology character observed in the CD titrations might be due to the unfolding pathway of the **G1**.
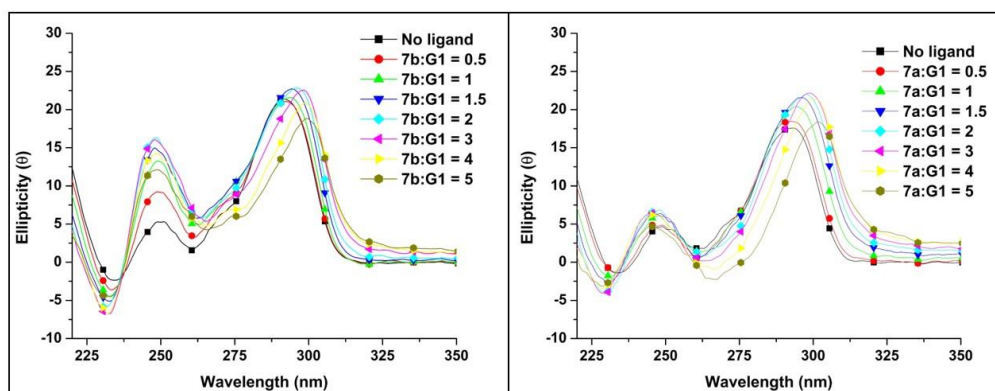
Figure 3.7. a) CD titrations of compound **7a** against **G1**. b) CD titrations of compound **7b** against **G1**.

Based on kinetic studies of folding and unfolding pathways of the human telomeric G-quadruplexes, Chaires *et al*[190] proposed the following folding/unfolding mechanism for htelo G-quadruplexes (**Figure 3.8**). The unfolding of the G-quadruplex was proposed to be sequential[190], which can be used to explain the T_m difference observed for **7a** and **7b** in our studies.

Non-specific binding of **7a** with random coil destabilizes the established equilibrium in solution. More hybrid G-quadruplex unfolds sequentially through a triplex like intermediate, an antiparallel intermediate, and finally a hairpin-like intermediate. The observation of the antiparallel signature by CD at the ligand to DNA molar ratio of 3 and higher and the observed destabilization of ($\Delta T_m = -4$ °C) could fit for this explanation. It is noteworthy that we observed turbidity in solution when titrating **7a** into **G1** at the ligand to DNA molar ratio more than 3. Turbidity could come from the precipitation of DNA because of the non-specific binding of **7a** with **G1.** In the FID experiments, the unfolding of **G1** by **7a** could release TO of **G1**, leading to a concomitant decrease in fluorescence.
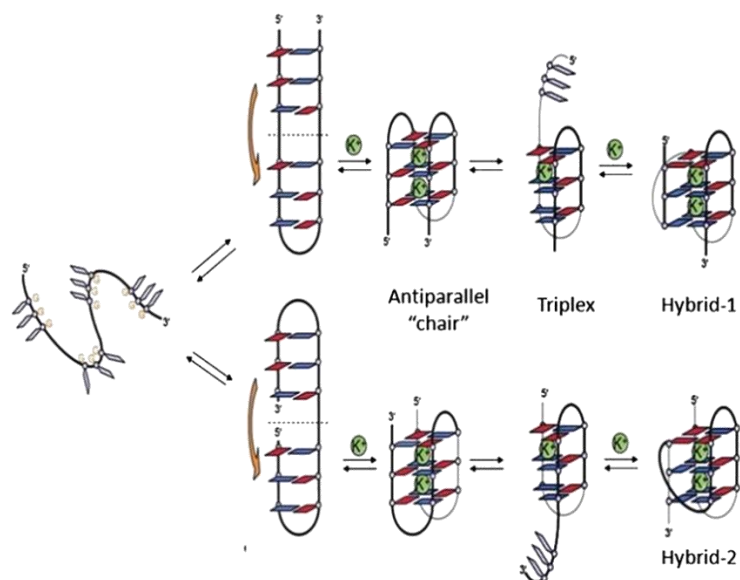
Figure 3.8. Folding and unfolding pathway of the htelo G-quadruplexes[190].

The stabilization effect of neomycin, **7a** and **7b**, and Phen on a duplex DNA (**D1**) was investigated using thermal denaturation monitored by UV (**Figure 3.9**). None of them showed a noticeable stabilization effect on the melting of **D1**, suggesting a weak binding of phen-neo conjugates with duplex DNA.
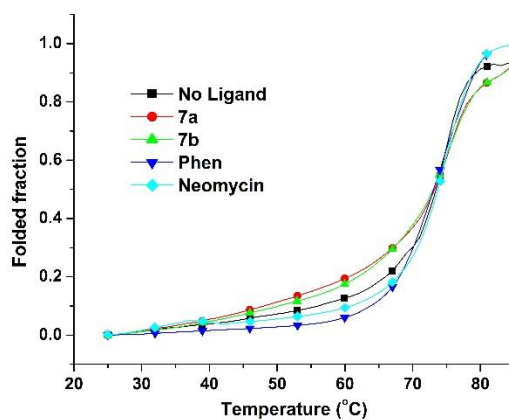
Figure 3.9. UV thermal denaturation of dsDNA in the presence phen-neo conjugates, phenanthroline and neomycin.

3.3.4. Compound 7b Binds to G1 in a 1:1 Binding Stoichiometry

The electrospray-ionization mass spectrometry (ESI-MS) is a highly sensitive analytical technique for the analysis of small molecule-DNA interactions. The use of ESI-MS for the G-quadruplex structure analyses[136], relative binding affinity calculations[82], and binding stoichiometry[176] is well documented. Compared with other analytical techniques like NMR and x-ray crystallography, which are also used for G-quadruplex-ligand binding interactions, ESI-MS requires simple sample preparation or less sample consumption due to its high sensitivity. In addition, studies have documented that G-quadruplexes maintained their solution-phase structure in the gas phase[191,192].

Because of the possible ion suppression effect of non-volatile alkali metals, ammonium acetate was commonly used for studying G-quadruplex DNA in ESI-MS. Ammonium ions assist in the folding of G-quadruplex just as K^+ ions do. As shown in **figure 3.10**, **G1** in NH_4^+ adopts a hybrid topology determined by CD spectroscopy. Because NH_4^+ ions are present between the adjacent G-tetrads; therefore, the number of NH_4^+ associated with the G-quadruplex structure indicates the number of G-tetrads in the topology. For example, the observation of two NH_4^+ ions in a G-quadruplex complex suggests that it contains three G-tetrads.[136,193]
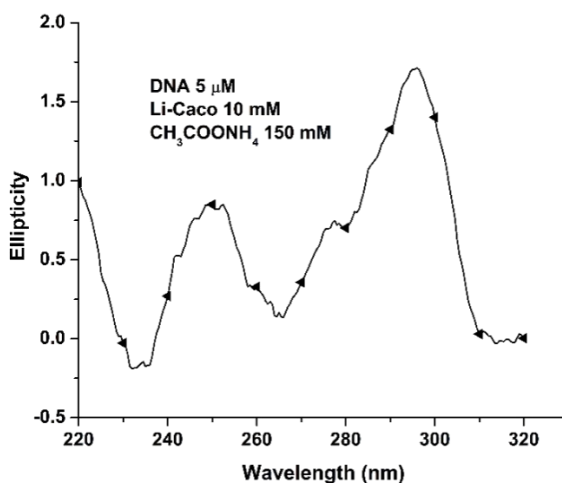
Figure 3.10. CD spectrum of the human telomere DNA (5 µM) annealed in 150 mM ammonium acetate and 10 mM LiCaco buffer.

In our study, the complexation of a ligand (**7a**, **7b**, or neomycin) with **G1** in NH_4^+ was investigated with electrospray ionization (ESI) mass spectrometry. All the spectra shown here are normalized relative to the most abundant ion peak (base peak) in the spectrum. Mixtures of a ligand (**7a** or **7b**) and **G1** at a molar ratio ranging from 1:1 to 1:4 in 150 mM NH_4^+ were prepared and analyzed. A solution of **G1** under the same condition in the absence of a ligand was used as a control. A mass spectrum representing a mixture of **G1** and **7b** ([**G1**]:[**7**] = 1:2) is shown in **figure 3.11b**. The formation of the G-quadruplex structure was confirmed by the peak at m/z 1459.85, representing the ion $[G1+2NH_4^+-7H^+]^{5-}$. The inclusion of two NH_4^+ ions is a distinct indication of the presence of three G-tetrads in the G-quadruplex. Compound **7b** was found to form a 1:1 complex with **G1**, indicated by the ion $[G1+2NH_4^+ +7b-7H^+]^{5-}$ at m/z 1642.13. The relative abundance of the complex is higher than those of free intact G-quadruplex when the molar ratio of **7b** to **G1** is larger than 1 suggesting the strong tenancy of **7b** to bind with the G-

quadruplex. It is noteworthy that varying the molar ratios of **7b** and **G1** did not lead to different

binding stoichiometries (**Figure 3.11c** and **3.11d**)
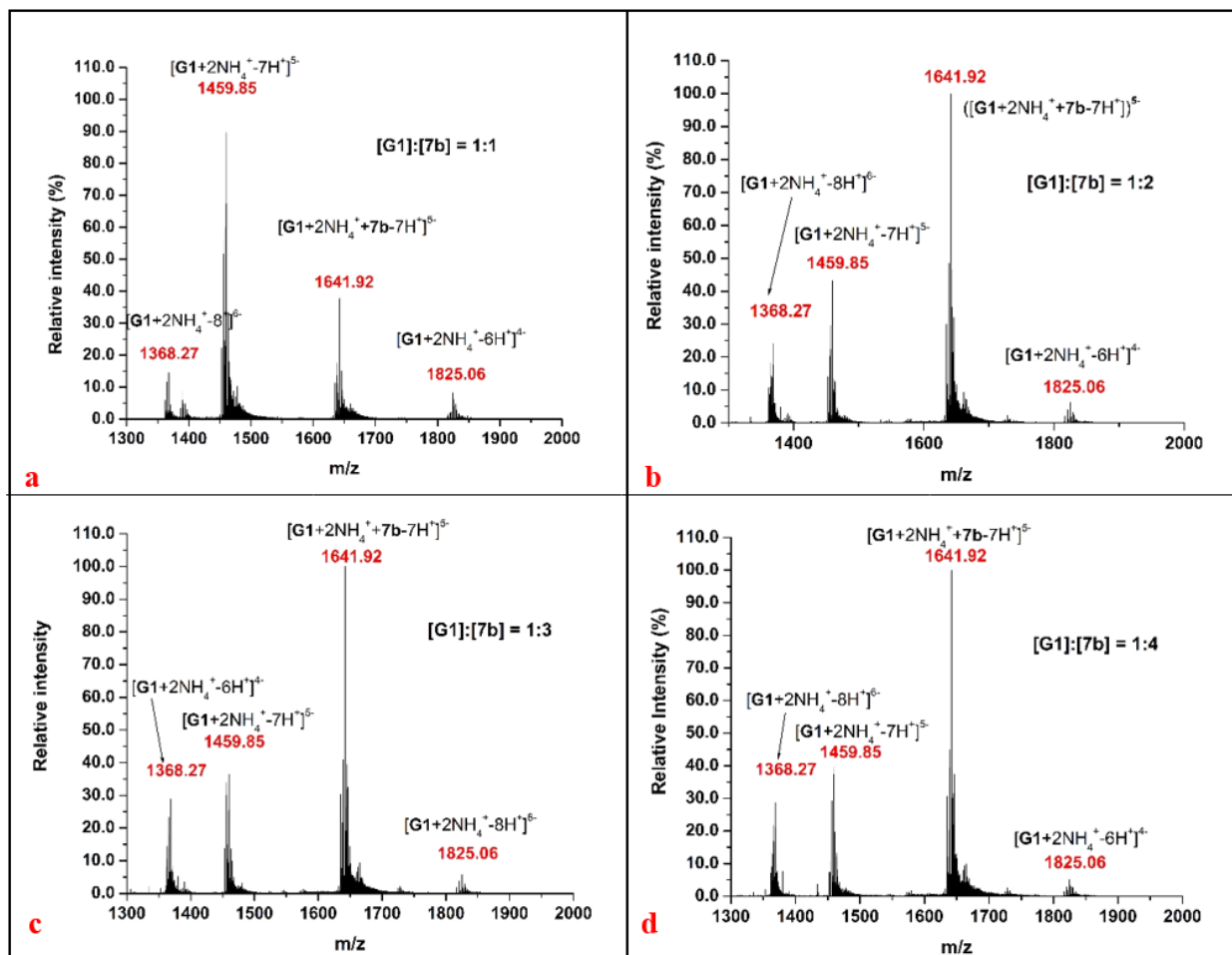
Figure 3.11. Mass spectrum of a mixture solution of **G1** and **7b** at different molar ratio. The peak with the maximum intensity was set as the base peak (100%).

Interestingly, mixtures of **7a** and **G1** did not show any complexation in ESI mass

spectrometry, even with the molar ratio of **7a** to **G1** at 4 (**Figure 3.12b**). It is noteworthy that

precipitation was observed when the ratio of **7a** to **G1** was above 3. Failing to observe the

complexation of **7a** with **G1** suggested a relatively weak interaction between the two, which is

consistent with our thermal stabilization studies. Because of this, in the following sections, we

will only pursue further biochemical and biophysical characterization of **7b** with telomeric G-quadruplex DNA.
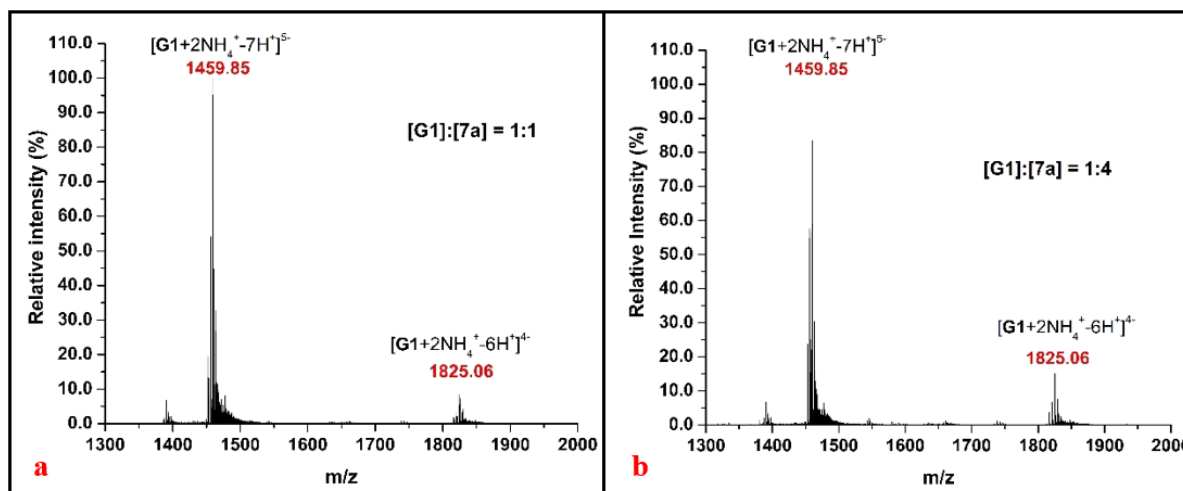
Figure 3.12. Mass spectrum of a mixture solution of **G1** and **7a** at [**7a**]:[**G1**] = 1 (left) and at [**7a**]:[**G1**] = 4 (right). The peak with the maximum intensity was set as the base peak (100%).

3.3.5. 7b Binds With a Much Higher Binding Affinity to G1 as Compared to Neomycin and Phenanthroline

Isothermal titration calorimetry (ITC) was used for the analysis of the thermodynamics of binding. ITC is a gold standard for measuring binding affinity, and it is the only technique that measures the binding enthalpy directly. It allows for the dissection of the free energy of binding into the relative contribution of the entropy and enthalpy components, providing indispensable information about the kinds of forces involved in binding and ligand optimization[194]. A complete thermodynamic characterization provides more information on the nature of driving force for the binding process if they are enthalpic or entropic in origin.

In a typical ITC experiment, a series of aliquots of a ligand are injected into a sample cell containing the binding substrate. The heat released or gained during each injection is integrated and plotted against the molar ratio of ligand to substrate. The resulting thermogram is then fitted with a binding model that can describe the binding event. The fitting provides crucial parameters, including the number of binding sites, enthalpy, entropy, and free energy of the binding event (**Figure 3.13**)[194].
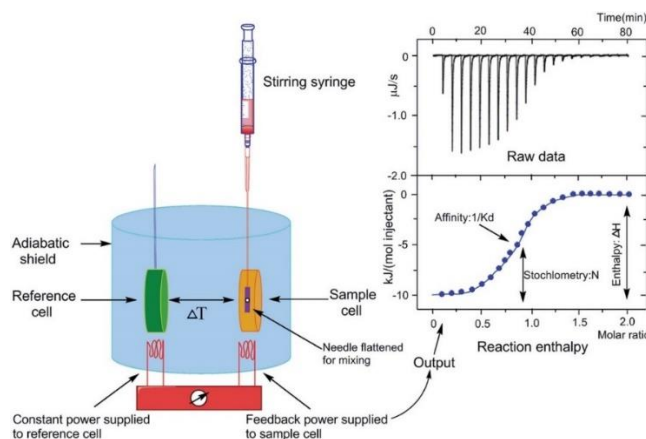
Figure 3.13. Basic principle of isothermal titration calorimetry[194].

In the present study, ligands (**7b** and neomycin) were analyzed for their thermodynamics of binding to a telomeric G-quadruplex (**G1**) and a 26-mer duplex (**D1**). **Figure 3.14** shows the representative ITC profiles (upper panels) resulting from a series of injections of **7b** (200 µM) into a solution (20 µM) of **G1** (A) and **D1** (B) in K$^+$ at 25 °C. Each injection of the ligand to the **G1**/**D1** is accompanied by the heat burst from the binding event and heat of dilution. The area under the thermogram peak is calculated with Nano Analyze software (version 3.8.0) and plotted against the molar ratio of ligand to DNA in the calorimetry cell. Calorimetric data from injections of **7b** into K$^+$ phosphate buffer under the same experimental conditions is used as

blank for the heat of dilution (black peaks in the thermograms). The blank data was deducted on

the peak by peak basis from the heat of interaction between ligand **7b** and **G1/D1**.

The dotted lines (**Figure 3.14**, lower panels) show the resulting corrected injection heats

(by subtracting the blank data for corrections of dilution effects) plotted as a function of the

[**7b**]/[DNA] molar ratio.
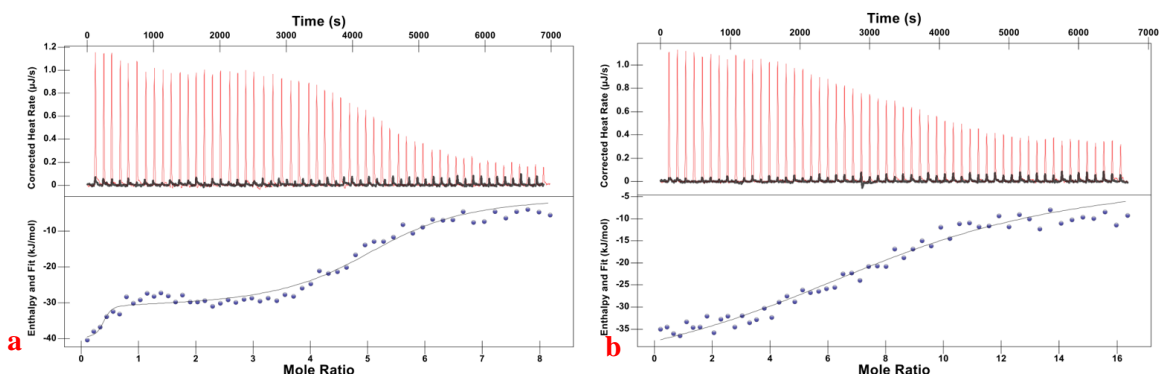
Figure 3.14. upper panel: ITC profiles of **G1** (20 µM) (A) and **D1** (10 µM) (B) titrated with **7b**
(200 µM) in a 10 mM potassium phosphate buffer with a total of 150 mM K$^+$, pH 7.0 at 25 °C.
The black peaks represent the heat of dilution. lower panel: The dotted line represents the
integrated corrected injection heats plotted as a function of the [**7b**]/[**G1**] (A) and [**7b**]/[**D1**] (B)
molar ratio. The smooth solid line reflects the calculated fit of the data using the binding models
(multiple-sites for G4 and independent for duplex) from the software NanoAnalyze. Upward
peaks represent an exothermic event.

Two slopes in the graph (**Figure 3.14a**, lower panel) are indicative of two primary

binding events (there could multiple-binding sites with equal binding affinities). The data best

fit a multiple-sites model with a 90% confidence level and result in the two association constants

(K$_a$s), $(8.9\pm2.4)\times10^8$ M^{-1} and $(3.5\pm0.5)\times10^5$ M^{-1}. Based on the K$_a$ values, the first binding event

is a tight binding with a sharp transition, and the second binding event shows a gradual slope

towards the binding saturation of **G1** by **7b**, suggesting a relatively weak binding. Neomycin is

positively charged under the conditions used in the study, so it is reasonable to expect weak non-specific electrostatic interactions between the charged molecules. A similar low-affinity second binding event at higher concentrations was also found in the study of G-quadruplex ligands such as BRACO-19[195] and Se2SAP[196] by other groups, which have permanently charged side chains. Binding enthalpies for the two events were determined to be ΔH_1 [-(37.3±1.6) kJ/mol] and ΔH_2 [-(32.3±0.5) kJ/mol], indicative of highly enthalpically favored binding for both events. The first binding event is entropy favorable with a TΔS value of +13.8 kJ/mol, which means the binding process is accompanied by favorable release of small molecules (e.g., water, ions). Just like duplex DNA, G-quadruplexes also have organized water molecules (hydration of spine) in the grooves[197]. The neomycin moiety of **7b** binds into the grooves of **G1** could release water molecules and/or metal ions. According to a survey of thermodynamic signatures for ligand-DNA binding mode, groove binding is predominantly entropically driven, while intercalation is enthalpically driven. Our observation suggests that the tight-binding event could involve both groove-binding and π-π stacking, implying a synergetic binding effect between Phen and neomycin. The second binding event is slightly entropy unfavorable with a TΔS value of -0.7 kJ/mol, which might be coming from non-specific interactions with the loop nucleotides or backbones without the release of relative equivalents of the water or ions. The presence of the first strong binding event was highlighted by a separate ITC titration experiment in which a series of injections of **7b** (200 μM) was titrated into a solution (40 μM) of **G1** to reach a final mole ratio of approximately 3 (**Figure 3.15a**).

The overall stoichiometry for binding of **7b** to **G1** in the K$^+$ solution was determined to be approximately five ligands per G-quadruplex DNA. The sizeable binding stoichiometry also suggests the second (weak) binding event may be associated with non-specific binding. It is

noted that only a 1:1 binding stoichiometry was obtained from the ESI/MS experiments, as

mentioned above. This discrepancy could result from the stability variations of ligand-DNA

complexes in the solution phase and the gas phase of experimental conditions; therefore, the

binding of **7b** to weaker binding sites (non-specific binding) might not be observed in the

ESI/MS experiments. The binding selectivity of **7b** for **G1** was analyzed by performing another

experiment at 25 °C with **D1**. The binding isotherm of **7b** with **D1** (**Figure 3.14b**, lower panel)

best fits with an independent (single-site) model with a 90% confidence level and result in an

association constant of $(5.9\pm0.3)\times10^4$ M^{-1} as evident from the gradual slope. It is evident that the

binding of **7b** with **D1** is much weaker than both binding events of **7b** with **G1** when comparing

K_a values. The results obtained here are consistent with the aforementioned biophysical studies,

in which compound **7b** binds preferentially to telomeric G4 over duplex DNA. As a comparison,

the titration of neomycin to **G1** at 25 °C was performed under the same condition. The binding

of neomycin to **G1** showed a much lower association constant of $(6.1\pm0.2)\times10^3$ M^{-1} (**Figure

3.15b**). Given that Phen does not bind to G4s and weak binding of neomycin with **G1**, the

strong binding of **7b** with **G1** must come from covalently tethering neomycin and phenanthroline

together.
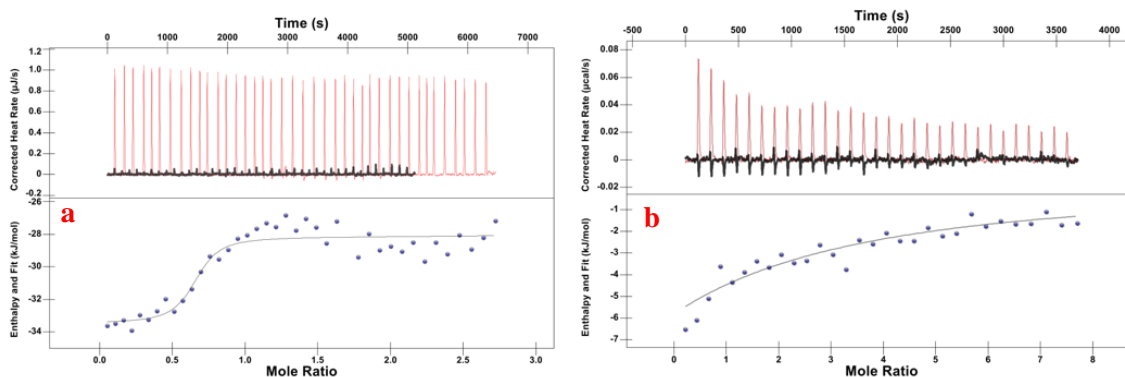
Figure 3.15. upper panel: ITC profiles of **G1** (40 µM) titrated with **7b** (200 µM) (A) and **G1** (20 µM) titrated with neomycin (400 µM) (B) in a 10 mM potassium phosphate buffer with a total of 150 mM K⁺, pH 7.0 at 25 °C. The black peaks represent the heat of dilution. lower panel: The dotted line represents the integrated corrected injection heats plotted as a function of the [**7b**]/[**G1**] (A) and [Neomycin]/[**G1**] (B) molar ratio. The smooth solid line reflects the calculated fit of the data using the binding models (Independent model for **7b** and multiple sites for neomycin) from the software NanoAnalyze. Upward peaks represent an exothermic event.

3.3.6. 7b Binds to G1 With a Dual Recognition Mode

Plausible binding modes of **7b** with human telomeric G-quadruplex DNA was investigated using the DOCK 6.9 program. Because the G4 structure of the exact sequence used this work is not available in the data bank, we modeled **G1** structure using the NMR solution structure (PDB id: 5Z80) of the complex formed by a platinum ligand and a human telomeric G-quadruplex as the initial point. It is noted that the G-quadruplex-forming regions of **G1** and 5Z80 are the same. Upon manual removal of the platinum ligand and modification of a few residues at the 5' and 3'-positions, the resulting G4 structure has the same sequence as **G1** and is a hybrid-type topology. The simple docking results show that compound **7b** can bind in the original binding site of the hybrid G-quadruplex, where the platinum ligand was found in the solution NMR structure. The most energetically favorable binding mode (**Figure 3.16a**) of **7b** has the phenanthroline ring stack on top of the base pair in the G-quadruplex through π-π stacking and the neomycin moiety in the groove of the G-quadruplex. Further analysis shows π-

π stacking between the phenanthroline ring of **7b** and residues of adenine-7 and guanine-8 (**Figure 3.16b**) and hydrogen bonding interactions of the neomycin moiety with guanine-21 amino group in the groove with the primary amine group on idopyranose ring of neomycin. Similarly, one of the amine groups in the glucopyranose ring of neomycin (circled) interacts with the guanine in the upper G-tetrad. In addition, neomycin was found to have hydrogen bonding interactions with the phosphate backbone of residues 19-21 (**Figure 3.16c**). The proposed binding mode is cohesive with the synergistic effect observed from biophysical studies. Using docking simulation, we also found several binding modes that do not interfere with each other. A representative binding mode of **G1** with three molecules of **7b** is presented in (**Figure 3.16d**). These binding modes could represent plausible non-specific interactions between **G1** and **7b** due to the electrostatic interactions of the neomycin moiety with DNA backbones.
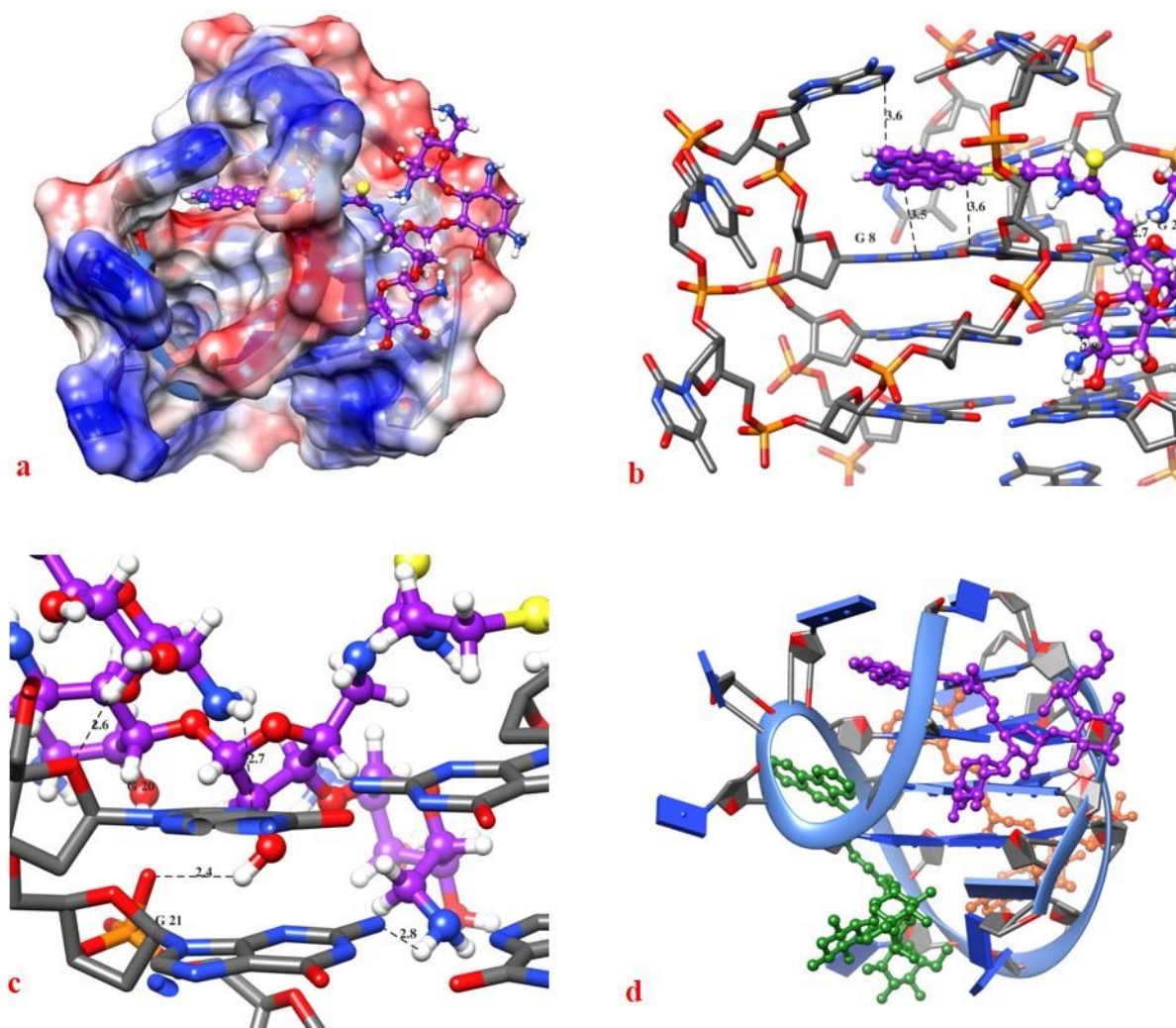
Figure 3.16. a) Docked structure of compound **7b** with a human telomeric G-quadruplex (**G1**). G-quadruplex is represented as surface colored by electrostatic potential, and the ligand is represented as ball and stick. b) π-π stacking interactions of the phenanthroline moiety with the top G-tetrad of **G1**. c) Binding interactions of the neomycin moiety with the groove of **G1**. The distances are in Å. d) A plausible binding mode of three molecules of **7b** at higher concentration to **G1**.3.3.6. **7b** inhibits the telomerase activity at low micromolar concentrations.

Upon confirming the favorable interactions between **7b** and the telomeric G4, we investigated their inhibition of human telomerase activity in vitro using a modified (Telomerase Repeat Amplification Protocol) TRAP assay.

3.3.7. 7b Inhibits Telomerase at Low Micromolar Concentration as Determined by TRAP Assay

TRAP assay was originally developed by Kim *et al*[198] to study the relationship between cancer cell immortality and telomerase concentration (above a certain threshold). It includes three steps (**Figure 3.17**):

> 1. Extension: Telomeric repeats are added to the radiolabeled telomerase substrate (TS Primer). In humans, a 6-mer repeat 5'-TTAGGG-3' is added by the telomerase during each round of addition.

> 2. Amplification: Extension of extended DNA by PCR with reverse primer RP and forward primer TS.

> 3. Detection of telomerase products: The presence and absence of telomerase activity is analyzed by gel electrophoresis. Typically, a ladder-like pattern reflects the activity of telomerase because it adds different units of 5'-TTAGGG-3' during the process.

The compound may inhibit the DNA polymerase during the amplification step, which would give false-positive results. We used an internal standard to distinguish if a compound inhibits telomerase activity (through G-quadruplex stabilization) and not DNA polymerase (directly or by binding to duplex) from amplifying the telomerase extended product because either event will give a positive result.

An internal standard TSNT, a non-G-rich oligonucleotide primer, is used in the assay, which is amplified with its own reverse primer NT during the amplification step. With the addition of the internal standard, the results can be interpreted in the following ways. First, if both the ladder pattern and internal standard are observed, then the ligand has no effect on both the telomerase and polymerase activity. Second, if the telomerase extended products are reduced or disappeared and the product of the internal standard is still present, the compound inhibits telomerase activity while having no effect on the polymerase activity. Therefore, the use of the

internal standard helps in predicting ligand selectivity for G-quadruplex (telomerase) over duplex

DNA (polymerase).

1. Elongation of the primer

TS Primer
5'-AATCCGTCGAGCAGAGTT-3' + dNTPs

Different extended products

Telomerase

5'-TTAGGG-3'

2. PCR amplification

ACX primer

Gel electrophoresis
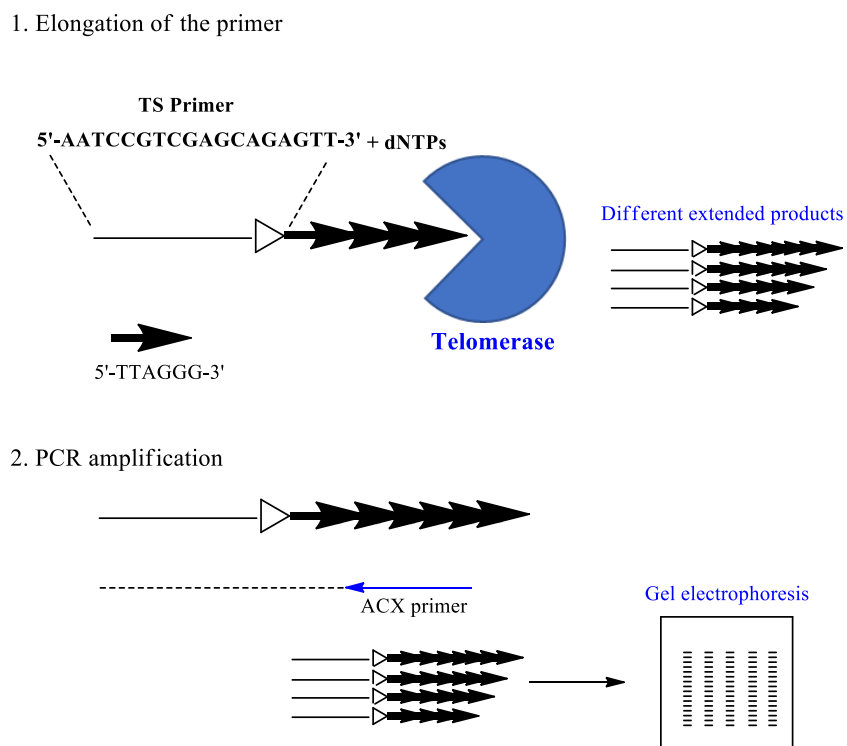
Figure 3.17. Schematic presentation of a TRAP assay.

Two control compounds, BIBR (50 µM) and TmPyP4 (100 µM), were employed in the

assay. BIBR[199] is a telomerase translocation inhibitor which inhibits the telomerase activity

without interacting with DNA. TMPyP4 is a known G-quadruplex DNA binding ligand[82,175].

As shown in **figure 3.18**, Lane 1 is the negative control Lane with no telomerase. Only

the product of the internal standard (IS) band was observed at the bottom. Lane 2 is a positive

control Lane with telomeres in the absence of any ligand. The latter pattern indicates the

telomerase activity, and the product of the internal standard (IS) at the bottom shows the

polymerase activity. Lane 3 represents the reaction with BIBR. The extension product bands were negligible, while the band for the internal standard was clearly observed, suggesting that BIBR inhibited telomerase but not polymerase in the reaction. In contrast to BIBR, TMPyP4, a potent non-selective G-quadruplex ligand, fully inhibited the activities of both telomerase and polymerase. No telomerase extension and PCR product bands (IS band disappeared) could be observed (Lane 4). The large stacking surface of TMPyP4 could intercalate into duplex DNA at this high concentration (50 μM) and thus inhibit the polymerase activity. These results were consistent with previously reported data. Compound **7b** (Lanes 5-8) selectively inhibited the telomerase activity while it has no effect on the polymerase activity in the concentration range of 1.6 to 12.5 µM. The inhibition of telomerase activity by **7b** could be unambiguously observed at 3.1 µM. Such inhibition is clearly concentration-dependent. At 12.5 µM, the activity of telomerase is greatly inhibited while the polymerase activity remains unaffected (IS band observed). Lane 9-11 represent the reactions in the presence of a mixture of phenanthroline and neomycin (1:1 molar ratio). Even at 25 µM, the mixture could not inhibit the telomerase activity because no extension products were observed. This observation further suggests a synergistic enhancement of the G-quadruplex mediated telomerase inhibition by phenanthroline-neomycin conjugation.
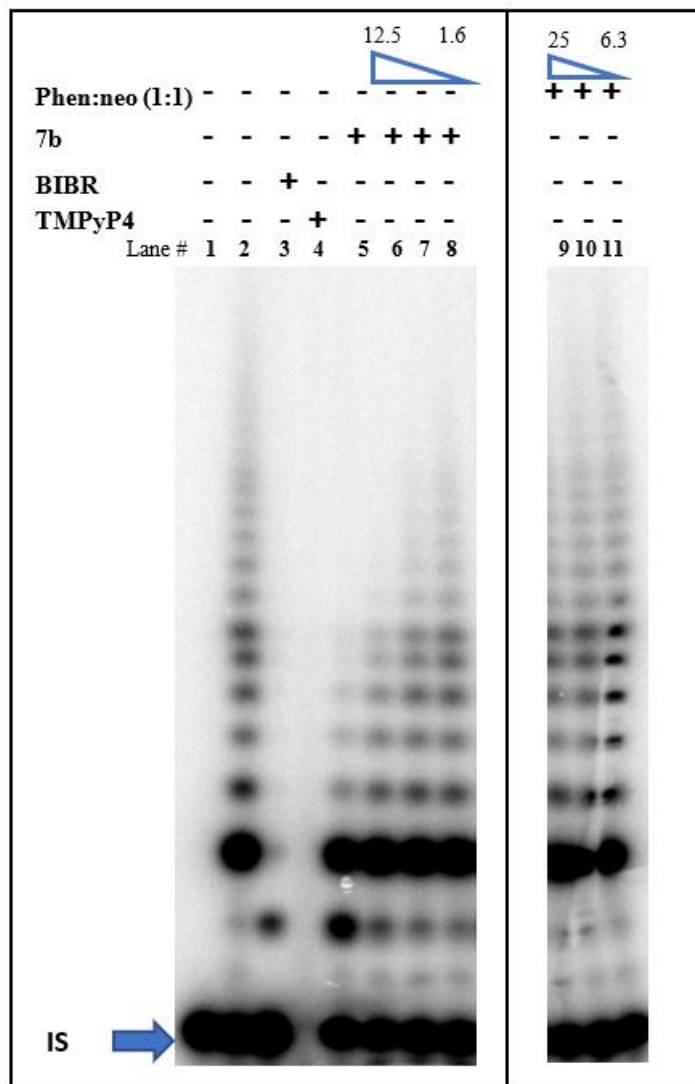
Figure 3.18. ^{32}P-TRAP reactions in the presence of **7b** (left) at concentrations of 12.5, 6.25, 3.125 and 1.56 µM in Lanes from 5 to 8, respectively. ^{32}P-TRAP reactions in the presence of phen:neo (1:1) (right) at concentrations of 25, 12.5, and 6.25 µM in Lanes from 9 to 11, respectively. Reactions in the presence of BIBR (50 µM) and TMPyP4 (50 µM) were carried out as controls. IS: PCR product as the internal control.

3.4. Conclusion

In summary, the data here demonstrate that structured amino-containing side chains such as neomycin can be used as a suitable binding motif for 5-substituted phenanthroline based G4

ligands and telomerase inhibitors. Phenanthroline-neomycin conjugates can bind to telomeric G-quadruplex selectively over duplex DNA. The binding of conjugates to telomeric G-quadruplex DNA depends on the linker length between two moieties. Phenanthroline-neomycin conjugates bind to G-quadruplex in a synergistic fashion, in which phenanthroline stacks with the end G-tetrad while neomycin binds into the groove. Phenanthroline-neomycin conjugates can inhibit telomerase activity at low micromolar concentrations. The research presented in this chapter suggests that a structured amino group side chain can be used as a motif for designing selective G-quadruplex ligands and telomerase inhibitors.

CHAPTER 4: 2-CLIP-PHEN DERIVATIVES AS G4 LIGANDS TARGETING

ANTIPARALLEL G-QUADRUPLEXES

4.1. Introduction

G-quadruplexes are dynamic DNA secondary structures, which have received much attention due to their direct involvement in the regulation of proto-oncogenes. The G-quadruplexes are formed in a stretch of G-rich strands, and many G-quadruplex conformations are in a dynamic equilibrium[36,37]. The interplay of these conformations may control gene expression in various ways; however, the understanding of such relationships is still in infancy. Efforts are usually spent on targeting the most stable conformation under physiological conditions studied by solution-phase NMR[63,101,200], molecular dynamics simulations, along with a variety of biochemical and biophysical techniques[201]. Most of the ligands have good selectivity for G-quadruplex over duplex DNA[195]. Examples of selective ligands toward G-quadruplexes with different conformations are scarce. Hundreds of thousands of putative G-quadruplexes forming sequences are present in the human genome. Without efficient ligands to target different G-quadruplexes, off-targeting side effects[202] could occur in therapeutic applications. There are more than 1000 G-quadruplex ligands in the G4 database. Most of them have extended aromatic surface to end-stack a G-quadruplex and contain cationic side chains to enhance the binding with electron-rich G-quadruplex grooves.[92,105] As mentioned in the previous chapters, extended aromatic rings could increase the selectivity of a ligand for G-quadruplex over duplex DNA, while cationic side chains often decrease the ligand's selectivity to some extent. Two drugs CX-5461[203] and CX-3543[204] (Quarfloxin) have entered clinical trials, which can induce selective lethality in DNA repair-deficient cancer cells by stabilizing G-

quadruplexes in vivo[203]. Both drugs cannot selectively target G-quadruplex conformations as found in vitro studies. In fact, most of the G-quadruplex binding ligands developed so far have limited conformational selectivity[61,87,175,205].

In our effort to develop more potent G-quadruplex ligands, we came across a novel class of molecules that show good selectivity for antiparallel G-quadruplexes over parallel and hybrid G-quadruplexes. In chapter 3, I reported phenanthroline-neomycin conjugates as G4 ligands. Previous studies in our group showed when 5-substituted 1,10-phenanthroline derivatives coordinate with Ni^{2+} ions, the resulting complexes have enhanced binding affinity towards G-quadruplexes[179]. The enhancement comes from forming a larger aromatic surface by coordinating two phenanthroline molecules with a Ni^{2+} ion. In light of this finding, we initially tried to complex phenanthroline-neomycin conjugates (Chapter 3) with Ni^{2+} to develop more potent G4 ligands; however, our attempt was unsuccessful because the interference of neomycin with metal ions made the study complicated. So, we sought to covalently link two 1,10-phenanthroline moieties with a serinol bridge that can mimic the phenanthroline-Ni^{2+} planar structure (**Figure 4.1**). The proposed molecule known as 2-Clip-phen was previously developed by Pitie *et al*[206]. 2-Clip-phen coordinated with Cu^{2+} has been used as an artificial endonuclease, which cleaves deoxyribose units from the minor grooves of DNA[207–209]. Efforts have been made to improve the DNA binding affinity by conjugating 2-Clip-phen with spermine[208]. 2-Clip-phen was tethered with acridine[210] and poly N-methylpyrrole[209] to increase its cellular uptake and DNA cleavage efficiency. 2-Clip-phen has also been linked to cis-platin, and the resulting complex showed dual functionality – DNA cleavage and DNA crosslinking[26,211] **(Figure 4.2)**. Ozalp-Yaman *et al*[212] synthesized bis-Clip phen platinum compounds for better DNA cleavage

and cytotoxicity. Recently, conjugation of clip-phen with a triplex-forming oligonucleotide (TFO) for site-specific DNA cleavage has reported[213].

From the structure viewpoint, 2-Clip-phen has a very large extended aromatic surface that in theory, should stack with G-tetrads well. We believe it could be used as a lead to develop G-quadruplex ligands. Surprisingly, to the best of our knowledge, no studies using 2-Clip-phen as a G-quadruplex ligand scaffold have been reported in the literature. Only one similar structure, a bis-phenanthroline derivative with a single-nitrogen-atom linker, was reported by Bianco *et al*[214] for recognition of G-quadruplexes (**Figure 4.3**), which has a weak-to-moderate G-quadruplex binding ability. Given the ability of 2-Clip-phen to cleave DNA, we believe that it can be used as a promising scaffold for developing G-quadruplex binding ligands and for G-quadruplex specific cleaving agents (future direction).
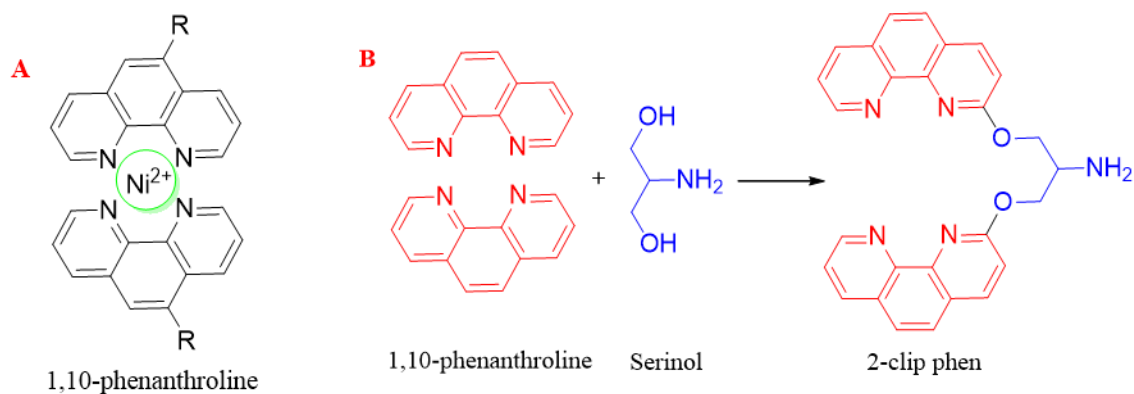

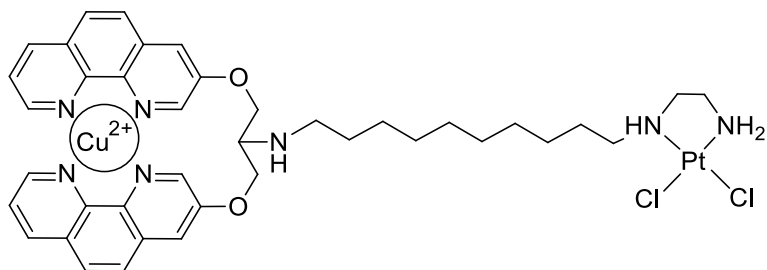*Figure 4.1.* A) 5-substituted phenanthrolines B) 2-Clip-phen design.

Figure 4.2. The complex of the Clip-phen with cis-platin.

$M = Ni^{2+}, Cu^{2+}$
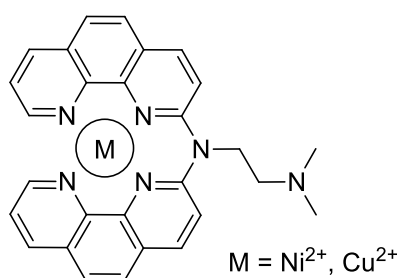
Figure 4.3. Bis-phenanthroline ligand with a three-carbon linker between two 1,10-phenanthroline rings.

Our multiple attempts to synthesize the 2-Clip-phen and neomycin conjugates were unsuccessful. The 2-Clip-phen moiety was degraded under the acidic deprotection condition for Boc groups (Scheme 4.2). Hence, we replaced neomycin with various amino-containing side chains and synthesized a series of conjugates. To our surprise, these conjugates can specifically target antiparallel G-quadruplexes over hybrid and parallel G-quadruplexes. This discovery is exciting, and it allows us to add novel G-quadruplex ligands that can distinguish conformations to the G4 ligand library. The structures of the 2-Clip-phen conjugates also provide future directions to develop novel G4 ligands. The synthesis and biophysical evaluations of these ligands will be presented in Chapter 4.

4.2. Experimental Section

4.2.1. Materials and Methods

All the chemicals for synthesis were purchased from Sigma-Aldrich or Fisher Scientific

and used without further purification. DNA oligonucleotides were synthesized on an ABI 392

DNA/RNA synthesizer using standard phosphoramidite protocols and purified using

polyacrylamide gel electrophoresis. All reagents and solid-support columns for DNA synthesis

were purchased from Glen Research. The concentrations of DNA solutions were determined by

UV spectroscopy, using the molar extinction coefficients (in units of mol of strand/L^{-1} cm^{-1})

obtained from OligoAnalyzer (www.idtdna.com). ε_{260} = 237,000 for **G1** [5'-

AGGG(TTAGGG)$_3$T-3']. ε_{260} = 248200 for **G3** (5'-TGG GGA GGG TGG GGA GGG TGG

GGA-3'). ε_{260} = 259,100 for **S1** (5'-ATGAAGGACGTAACCGGCTCTGAACG-3'); ε_{260} =

167,400 for **S2** (5'- CGTTCAGAGCCGGTTACGTCCTTCAT-3'). ^{1}H NMR and ^{13}C NMR

spectra were collected on a JEOL ECA 600 MHz FT-NMR spectrometer. Mass spectra of the

synthesized compounds were analyzed on an Agilent 1290 UHPLC coupled to an Agilent 6230

Time-of-Flight mass spectrometer. Mass spectra of ligand-DNA complexes were collected on a

Thermo Fisher LTQ Mass Spectrometer. UV spectra were recorded on a Varian Cary 100 Bio

UV-Vis spectrophotometer equipped with a thermoelectrically controlled 6 × 6 cell holder.

Circular dichroism spectra were recorded on a JASCO J-810 spectropolarimeter using a quartz

cuvette with a 1 mm or 1 cm optical path length.

4.2.2. Synthesis of 2-Clip-phen (2-CP) Derivatives

Compounds **2**-**CP**-**1** and **2**-**CP**-**2** were synthesized according to the scheme 4.1. **5a-NCS**

was synthesized using the published procedure[154]. Compound **5b**-**NCS** was synthesized from

5b-**S**, which was synthesized as described in chapter 3. We were not able to deprotect the **2**-**CP**-

1 and **2**-**CP**-**2** to get the final compounds even when we used the mildest conditions possible[215] for deprotection. The synthesis of **2**-**CP**-**1** and **2**-**CP**-**2** is described below (Scheme 4.1 and Scheme 4.2). We also synthesized a series of novel 2-Clip-phen derivatives (**10**-**18**) by decorating 2-Clip-phen scaffold with potent G-quadruplex binding side chains[88,130] (scheme 4.3 and scheme 4.4).

Synthesis of 2-Clip-phen (2-CP). Synthesis of this compound was done following the published procedure[206] with some changes. To the solution of sodium hydride (232 mg, 5.8 mmol) in dry DMF (4ml) at 0°C in two neck flask were added 2-bromo-1,10-phenanthroline (200 mg, 0.77 mmol) and serinol (35mg, 0.386 mmol). The reaction mixture was stirred at 0°C for 24 h and then allowed to warm to room temperature. The crude product was dissolved in 10 ml of DMF and 3.4 ml of ethanol and precipitation was done with 33 ml of water. The precipitates were collected with vacuum filtration and re-dissolved in 18 ml of hot methanol and precipitated again with water. Finally, the precipitates were dissolved in dichloromethane and precipitated with hexane. The compound has R_f value of 0.17, $CH_3OH:CH_2Cl_2:NH_4OH =$ 6:94:0.6 (v/v/v), silica gel 60 F_{254}, EMD Millipore. Yield = 143 mg (32 mmol, 41.4%).^{1}H NMR (600 MHz, CDCl$_3$): δ 9.15 (dd, $J = 4.2$, 1.8 Hz, 2H), 8.21 (dd, $J = 7.8$, 1.8 Hz, 2H), 8.09 (d, $J = 9.0$ Hz, 2H), 7.72 (d, $J = 8.4$ Hz, 2H), 7.57 (m, 2H), 7.16 (d, $J = 9.0$ Hz, 2H), 4.98 (d, $J = 5.4$ Hz, 4H), 3.84 (quint, $J = 5.4$ Hz, 1H). ^{13}C-NMR (151 MHz, CDCl$_3$) δ 162.99, 150.05, 139.18, 136.27, 126.39, 123.86, 122.61, 114.14, 68.45, 50.75.

Synthesis of 5b-NCS. In 20-ml round-bottomed flask, **5b**-**S** (60 mg, 0.493 mmol) was dissolved in 6 ml of dry DMF. To the stirred solution TCDP (14.3 mg, 0.0616 mmol) along with catalytic amount of DMAP was added. The reaction mixture was stirred for 12 h at room temperature under Argon. The completion of the reaction was monitored with TLC [$R_f = 0.3$

(CH$_3$OH:CH$_2$Cl$_2$ = 8:92 (v/v), silica gel 60 F$_{254}$, EMD Millipore]. The residue was concentrated

under vacuum. Flash chromatography of the residue [silica gel, CH$_3$OH:CH$_2$Cl$_2$ = 8:92: (v/v)]

yielded compound **5b-NCS** (47 mg, 76 %) as a white solid. ^{1}H-NMR (600 MHz, CDCl$_3$) δ 6.56-

4.48 (m, 8H), 4.45-2.65 (m, 19H), 1.82-1.29 (m, 54H).

Synthesis of 2-CP1. In 10-ml round-bottomed flask, **5a-NCS** (58 mg, 0.44 mmol) was

dissolved in anhydrous pyridine (5 ml). An anhydrous pyridine solution (1 ml) of **2-CP** (21.6

mg, 0.048 mmol) was added to the flask with a syringe. The reaction mixture was stirred

overnight at room temperature under N$_2$. The reaction progress was monitored with TLC [R$_f$ =

0.27 (CH$_3$OH:CH$_2$Cl$_2$:NH$_4$OH = 8:92:0.8 (v/v/v), silica gel 60 F$_{254}$, EMD Millipore]. The

residue was concentrated under vacuum. Flash chromatography of the residue [silica gel,

CH$_3$OH:CH$_2$Cl$_2$ = 8-10:92-90 (v/v)] yielded compound **2-CP1** (32 mg, 41 %) as a white solid.

IR (KBr pellet, cm^{-1}) 3419, 2968.17, 2936, 1702, 1605, 1498, 1369, 1262, 1155, 1026, 843. ^{1}H

NMR (CD$_3$OD) δ 8.93 (m, 2H), 8.28 (d, 2H, *J* = 7.8 Hz), 7.93 (bs, 2H), 7.35 - 7.77 (m, 6H), 7.00

- 7.24 (m, 2H), 5.43 (bs, 2H), 5.25 (d, 1H, *J* = 9.6 Hz), 5.01 - 5.19 (m, 3H), 4.18 - 4.42 (m, 2H),

4.02 - 4.14 (m, 1H), 3.69 - 4.01 (m, 6H), 3.59-3.68 (s, 2H), 3.41 - 3.59 (m, 6H), 2.72 - 3.01 (bs,

4H), 1.86 - 2.05 (m, 1H), 1.1 - 1.67 (m, 54H). HRMS (ESI) calcd for C$_{83}$H$_{119}$N$_{12}$O$_{26}$S$_2$ [M+H]$^+$:

1763.7800, found 1763.7766.

Synthesis of 2-CP2. In 10-ml round-bottomed flask, **5b-NCS** (47 mg, 0.38 mmol) was

dissolved in anhydrous pyridine (5 ml). An anhydrous pyridine solution (1 ml) of **2-CP** (18.7

mg, 0.042 mmol) was added to the flask with a syringe. The reaction mixture was stirred

overnight at room temperature under N$_2$. The reaction progress was monitored with TLC [R$_f$ =

0.43 (CH$_3$OH:CH$_2$Cl$_2$:NH$_4$OH = 8:92:0.8 (v/v/v), silica gel 60 F$_{254}$, EMD Millipore]. The

residue was concentrated under vacuum. Flash chromatography of the residue [silica gel,

CH$_3$OH:CH$_2$Cl$_2$ = 8:92 (v/v)] yielded compound **2-CP2** (30 mg, 47 %) as a white solid. IR (KBr

pellet, cm^{-1}) 3422, 2987, 2932, 1691, 1507, 1376, 1267, 1180, 1039, 853. ^{1}H NMR (CD$_3$OD) δ

8.89 (dd, 2H, J = 10.8, 3.6 Hz), 8.25 (t, 2H, J = 6.6 Hz), 8.86 (bs, 2H), 7.33-7.68 (m, 6H), 7.12

(q, 2H, J = 9 Hz), 5.78 (bs, 1H), 5.36 (bs, 2H), 4.99-5.25 (m, 3H), 4.60 (s, 1H), 4.28 - 4.51 (m,

2H), 4.19 (bs, 1H), 3.85 - 4.11 (m, 3H), 3.45 - 3.83 (m, 9H), 1.89 - 2.04 (m, 1H), 1.15 - 1.66 (m,

54H). ^{13}C NMR (CD$_3$OD) δ 163.94, 159.04, 159.01, 158.73, 158.62, 158.25, 158.00, 149.94,

149.84, 145.38, 145.35, 144.34, 144.30, 140.05, 138.01, 137.93, 130.41, 130.36, 127.45, 127.30,

126.05, 126.00, 124.40, 123.76, 115.65, 115.62, 113.81, 100.04, 98.65, 80.75, 80.64, 80.43,

80.32, 76.35, 75.81, 74.62, 73.22, 73.16, 73.09, 73.00, 72.41, 71.71, 69.13, 65.87, 65.39, 56.90,

54.57, 53.54, 52.52, 51.24, 49.85, 42.42, 42.02, 35.74, 29.02, 28.92, 28.84, 28.79, 28.76. HRMS

(ESI) calcd for C$_{81}$H$_{115}$N$_{12}$O$_{26}$S [M+H]$^+$: 1703.7466, found 1703.7832.

General procedure for synthesis of AM1-9. 20-50 mg (1 equivalent) of the amine was

dissolved in 5 ml dry DCM. To the solution was added TCDP (1.5 equivalent), DMAP (catalytic

amount). The reaction mixture was stirred at room temperature for 12 h under N$_2$ atm. Reaction

was monitored with TLC.

Compound AM1. Pale oil (52.35%, 35.8mg), The reaction was monitored with TLC [R$_f$

= 0.33, CH$_2$Cl$_2$: CH$_3$OH = 0.45:9.55 (v/v), silica gel 60 F$_{254}$, EMD Millipore]. Column

chromatography [silica gel, CH$_3$OH:CH$_2$Cl$_2$ = 4.5:95.5 (v/v)]. ^{1}H-NMR (600 MHz, CDCl$_3$) δ

3.63 (t, J = 8Hz, 2H), 2.77 (t, J = 8 Hz, 2H), 2.50-2.62 (m, 4H), 1.74 - 1.83 (m, 4H). ^{13}C-NMR

(151 MHz, CDCl3) δ 131.47, 55.23, 54.17, 44.68, 23.60.

Compound AM2. White oil (52.7%, 35 mg), The reaction was monitored with TLC [R$_f$ =

0.21, CH$_2$Cl$_2$: CH$_3$OH = 0.6:9.4 (v/v), silica gel 60 F$_{254}$, EMD Millipore]. Column

chromatography [silica gel, CH$_3$OH:CH$_2$Cl$_2$ = 6:94 (v/v)]. ^{1}H-NMR (600 MHz, CDCl$_3$) δ 4.00

(t, *J* = 4 Hz, 2H), 2.48 - 2.63 (m, 6H), 1.91 (quint, *J* = 8 Hz, 2H), 1.73-1.81 (m, 4H). ^{13}C-NMR (151 MHz, CDCl$_3$) δ 54.29, 53.04, 43.37, 29.11, 23.60.

Compound AM3. Slightly yellow oil (95%, 19mg), The reaction was monitored with TLC [R$_f$ = 0.35, CH$_2$Cl$_2$: CH$_3$OH = 0.45:9.55 (v/v), silica gel 60 F$_{254}$, EMD Millipore]. Column chromatography [silica gel, CH$_3$OH:CH$_2$Cl$_2$ = 4.5:95.5 (v/v)]. ^{1}H-NMR (600 MHz, CDCl$_3$) δ 3.63 (t, *J* = 8 Hz, 2H), 2.77 (t, *J* = 8 Hz, 2H), 2.47 - 2.63 (m, 4H), 1.64 - 1.86 (m, 4H). ^{13}C-NMR (151 MHz, CDCl$_3$) δ 132.13, 58.10, 54.48, 43.18, 26.00, 24.21.

Compound AM4. Slightly yellow oil (49.2%, 31.9 mg), The reaction was monitored with TLC [R$_f$ = 0.425, CH$_2$Cl$_2$: CH$_3$OH = 0.8:9.2 (v/v), silica gel 60 F$_{254}$, EMD Millipore]. Column chromatography [silica gel, CH$_3$OH:CH$_2$Cl$_2$ = 8:92 (v/v)]. ^{1}H-NMR (600 MHz, CDCl$_3$) δ 3.59 (t, *J* = 8 Hz, 2H), 2.28 - 2.50 (m, 6H), 1.87 (quint, *J* = 8 Hz, 2H), 1.59 (quint, 4H), 1.38 - 1.48 (m, 2H). ^{13}C-NMR (151 MHz, CDCl$_3$) δ 55.77, 54.74, 43.44, 27.40, 25.97, 24.42.

Compound AM5. Slightly yellow oil (51.5%, 34mg), The reaction was monitored with TLC [R$_f$ = 0.65, CH$_3$OH:CH$_2$Cl$_2$ = 0.3:9.7 (v/v), silica gel 60 F$_{254}$, EMD Millipore]. Column chromatography [silica gel, CH$_3$OH:CH$_2$Cl$_2$ = 3:97 (v/v)] yielded the desired compound. ^{1}H-NMR (600 MHz, CDCl$_3$) δ 3.71 (t, *J* = 4.5 Hz, 4H), 3.58 (t, *J* = 6.2 Hz, 2H), 2.66 (t, *J* = 6.2 Hz, 2H), 2.50 (t, *J* = 4.1 Hz, 4H). ^{13}C NMR (151 MHz, CDCl$_3$) δ 133.03, 67.02, 57.73, 53.42, 42.97.

Compound AM6. Slightly yellow oil (78.1 %, 50.4mg) The reaction was monitored with TLC [R$_f$ = 0.36, CH$_3$OH:CH$_2$Cl$_2$ = 0.4:9.6 (v/v), silica gel 60 F$_{254}$, EMD Millipore]. Column chromatography [silica gel, CH$_3$OH:CH$_2$Cl$_2$ = 4:96 (v/v)] yielded the desired compound. ^{1}H-NMR (600 MHz, CDCl$_3$) δ 3.68 (t, *J* = 4.5 Hz, 4H), 3.59 (t, *J* = 6.5 Hz, 2H), 2.43 (t, *J* = 6.9 Hz, 6H), 1.82 - 1.86 (m, 2H). ^{13}C-NMR (151 MHz, CDCl$_3$) δ 130.15, 67.03, 55.25, 53.73, 43.10, 26.94.

Compound AM7. Slightly yellow oil (72.5 %, 47 mg). The reaction was monitored with TLC [R_f = 0.146, CH_3OH: $(C_2H_5)_2O$ = 1:9 (v/v), silica gel 60 F_{254}, EMD Millipore]. Column chromatography [silica gel, CH_3OH: $(C_2H_5)_2O$ = 4:96 (v/v)] yielded the desired compound. ^{1}H-NMR (600 MHz, $CDCl_3$) δ 3.56 (t, *J* = 6.2 Hz, 2H), 2.65 (t, *J* = 6.2 Hz, 2H), 2.31 - 2.52 (m, 8H), 2.27 (s, 3H). ^{13}C-NMR (151 MHz, $CDCl_3$) δ 132.63, 57.27, 55.09, 52.96, 46.09, 43.11.

Compound AM8. Slightly yellow oil (54.6 %, 34.6 mg). The reaction was monitored with TLC [R_f = 0.65, CH_3OH:CH_2Cl_2 = 0.7:9.3 (v/v), silica gel 60 F_{254}, EMD Millipore]. Column chromatography [silica gel, CH_3OH:CH_2Cl_2 = 7:93 (v/v)] yielded the desired compound. ^{1}H-NMR (600 MHz, $CDCl_3$) δ 3.56 (t, *J* = 6.5 Hz, 2H), 2.26 - 2.43 (m, 13H), 1.80 - 1.84 (m, 2H). ^{13}C-NMR (151 MHz, $CDCl_3$) δ 130.02, 55.16, 54.80, 53.17, 46.09, 43.19, 27.27.

Compound AM9. Pale oil (15mg, 51.93 %). The reaction was monitored with TLC [R_f = 0.37, CH_3OH:CH_2Cl_2:NH_4OH = 5:95 (v/v)], silica gel 60 F_{254}, EMD Millipore]. Column chromatography [silica gel, CH_3OH:CH_2Cl_2 = 5:95 (v/v)] yielded the desired compound. ^{1}H-NMR (600 MHz, $CDCl_3$) δ 3.62 (t, *J* = 6.5 Hz, 2H), 2.63 (t, *J* = 6.2 Hz, 2H), 2.31 (s, 6H). ^{13}C-NMR (151 MHz, $CDCl_3$) δ 58.55, 45.42, 43.46, 131.98.

General procedure for the synthesis of 2-Clip-phen neomycin conjugates (2-CP1 and 2-CP2). To 5-8M solution of the 5a/5b-NCS in dry pyridine, 1.1 molar equivalent of the 2-Clip-phen is added. The reaction mixture is stirred overnight at room temperature. The highly non-polar boc-protected neomycin impart sufficient non-polar character to the molecule that it can be separated from the unreacted 2-Clip-phen without a significant problem using column chromatography.

General procedure for the synthesis of 10-18. To the solution of 2-Clip-phen (10-25 mg in 5 ml of dry pyridine), 4 equivalents of **AM1-9** was added to consume all the 2-Clip-phen. The pyridine is removed under reduced pressure. Crude residue is rinsed with hexane to remove most of the unreacted **AM1-9**. Flash chromatography of the residue yields the white solid as the final compounds (**10-18**).

Compound 10. White solid (59.5%, 16mg). The reaction was monitored with TLC [R_f = 0.25, CH_2Cl_2: CH_3OH = 0.4:9.6 (v/v), silica gel 60 F_{254}, EMD Millipore]. Column chromatography [silica gel, [silica gel, CH_3OH:CH_2Cl_2:NH_4OH = 4:96:0.4 (v/v/v)] yielded the desired compound. IR (KBr, cm^{-1}) 3270, 3065, 2957, 2936, 2877, 2794, 1608, 1592, 1451, 1554, 1451, 1417, 1394, 1344, 1221. ^{1}H -NMR (600 MHz, CDCl$_3$) δ 9.08 - 9.09 (m, 2H), 8.16 - 8.02 (2H), 7.87 (s, 2H), 7.49 - 7.59 (m, 6H), 6.98 (s, 2H), 5.46 (d, J = 4.6 Hz, 1H), 5.22 (s, 2H), 5.02 (s, 2H), 3.78 (s, 2H), 2.37 (d, J = 201.7 Hz, 6H), 1.29 (s, 4H). ^{13}C-NMR (151 MHz, CDCl$_3$) δ 182.91, 162.92, 149.63, 144.98, 143.94, 138.83, 136.17, 129.02, 126.27, 124.76, 123.66, 122.56, 114.16, 65.77, 54.68, 53.63, 43.68, 23.30.

Compound 11. White solid (52.1%, 10.7mg). The reaction was monitored with TLC [R_f = 0.225, CH_3OH:CH_2Cl_2:NH_4OH = 10:90:1 (v/v/v)], silica gel 60 F_{254}, EMD Millipore]. Column chromatography [silica gel, CH_3OH:CH_2Cl_2:NH_4OH = 10:90:1 (v/v/v)] yielded the desired compound. IR (KBr, cm^{-1}) 3451, 3070, 2955, 2877, 1607, 1592, 1503, 1453, 1348, 1264, 1223. ^{1}H -NMR (600 MHz, CDCl$_3$) δ 9.06 (d, J = 2.7 Hz, 2H), 8.16 (d, J = 7.6 Hz, 2H), 7.97 (s, 2H), 7.54 - 7.66 (m, 6H), 7.03 (d, J = 6.9 Hz, 2H), 5.36 (s, 1H), 5.14 (s, 2H), 4.95 (s, 2H), 3.73 (s, 2H), 2.75 - 2.24 (6H), 1.92 - 1.55 (6H). ^{13}C -NMR (151 MHz, CDCl$_3$) δ 162.97, 149.67, 145.01, 143.96, 139.11, 136.42, 129.21, 126.39, 124.95, 123.88, 122.77, 114.26, 66.25, 53.94, 43.32, 29.79, 27.61, 23.44.

Compound 12. White solid (53.2%, 15.3mg), The reaction was monitored with TLC [R$_f$ = 0.29, CH$_3$OH:CH$_2$Cl$_2$:NH$_4$OH = 5:95:0.5 (v/v/v)], silica gel 60 F$_{254}$, EMD Millipore]. Column chromatography [silica gel, CH$_3$OH:CH$_2$Cl$_2$:NH$_4$OH = 5:95:0.5 (v/v/v)] yielded the desired compound. IR (KBr, cm^{-1}) 3428, 3074, 2931, 2851, 1608, 1592, 1503, 1453, 1416, 1349, 1266, 1222. ^{1}H-NMR (600 MHz, CDCl$_3$) δ 9.06 (s, 2H), 8.03 (s, 2H), 7.79 (s, 2H), 7.47 (s, 6H), 6.91 (s, 2H), 5.55 - 5.36 (1H), 5.32 - 4.92 (4H), 4.04 - 3.62 (2H), 2.78 - 1.76 (m, 6H), 1.48 - 0.90 (m, 6H). ^{13}C-NMR (151 MHz, CDCl$_3$) δ 182.86, 162.66, 149.42, 144.71, 143.72, 138.55, 135.96, 128.82, 126.04, 124.53, 123.46, 122.44, 113.90, 65.40, 57.16, 53.89, 41.34, 29.62, 24.89, 23.73.

Compound 13. White solid (93.4 %, 26.4 mg). The reaction was monitored with TLC [R$_f$ = 0.19, CH$_3$OH:CH$_2$Cl$_2$:NH$_4$OH = 6:94:0.6 (v/v/v)], silica gel 60 F$_{254}$, EMD Millipore]. Column chromatography [silica gel, CH$_3$OH:CH$_2$Cl$_2$:NH$_4$OH = 6:94:0.6 (v/v/v)] yielded the desired compound. IR (KBr, cm^{-1}) 3436, 3065, 2933, 2853, 2805, 1608, 1592, 1560, 1546, 1558, 1452, 1417, 1395, 1349. ^{1}H -NMR (600 MHz, CDCl$_3$) δ 9.06 (d, J = 2.7 Hz, 2H), 8.17 (d, J = 7.6 Hz, 2H), 7.97 (d, J = 6.2 Hz, 2H), 7.54 - 7.66 (m, 6H), 7.04 (d, J = 7.6 Hz, 2H), 5.41 (s, 1H), 5.05 (d, J = 129.2 Hz, 4H), 3.69 (s, 2H), 2.21 (s, 6H), 1.32-1.71 (m, 8H). ^{13}C -NMR (151 MHz, CDCl$_3$) δ 162.99, 149.67, 145.04, 143.97, 139.09, 136.40, 129.21, 126.40, 124.95, 123.85, 122.74, 114.27, 66.01, 56.73, 54.33, 43.26, 26.30, 25.55, 24.17.

Compound 14. White solid (93.8%, 13 mg). The reaction was monitored with TLC [R$_f$ = 0.22, CH$_3$OH:CH$_2$Cl$_2$:NH$_4$OH = 4:96:0.4 (v/v/v)], silica gel 60 F$_{254}$, EMD Millipore]. Column chromatography [silica gel, CH$_3$OH:CH$_2$Cl$_2$:NH$_4$OH = 4:96:0.4 (v/v/v)] yielded the desired compound. IR (KBr, cm^{-1}) 3451, 3280, 3063, 2954, 2867, 1608, 1592, 1562, 1503, 1346, 1418, 1222, 1135, 1222. ^{1}H-NMR (600 MHz, CDCl$_3$) δ 9.06 (s, 2H), 7.76 (d, J = 334.7 Hz, 10H), 6.94 (s, 2H), 5.42 (s, 1H), 5.07-5.27 (m, 4H), 3.75 (s, 2H), 2.35 - 3.17 (m, 4H), 1.63 - 2.19 (m, 7H).

^{13}C-NMR (151 MHz, CDCl$_3$) δ 162.84, 149.50, 144.90, 143.85, 138.63, 136.11, 128.93, 126.23, 124.66, 123.58, 122.55, 114.12, 66.46, 65.45, 57.19, 53.52, 53.16.

Compound 15. White solid (62.6%, 20 mg). The reaction was monitored with TLC [R$_f$ = 0.15, CH$_3$OH:CH$_2$Cl$_2$:NH$_4$OH = 4:96:0.4 (v/v/v)], silica gel 60 F$_{254}$, EMD Millipore]. Column chromatography [silica gel, CH$_3$OH:CH$_2$Cl$_2$:NH$_4$OH = 4:96:0.4 (v/v/v)] yielded the desired compound. ^{1}H-NMR (600 MHz, CDCl$_3$) δ 9.06 (d, *J* = 2.7 Hz, 2H), 8.16 (d, *J* = 7.6 Hz, 2H), 7.95 (s, 2H), 7.53-7.65 (m, 6H), 7.02 (s, 2H), 5.44 (s, 1H), 5.17 (s, 2H), 4.95 (s, 2H), 3.72 (s, 1H), 3.49 (s, 4H), 2.60 (s, 1H), 2.14 (s, 6H), 1.65 (s, 2H). ^{13}C-NMR (151 MHz, CDCl$_3$) δ 163.00, 149.60, 145.03, 143.93, 139.06, 136.40, 129.18, 126.42, 124.93, 123.81, 122.70, 114.28, 66.88, 66.08, 56.38, 53.55, 43.10, 26.18.

Compound 16. White solid (72.2%, 24.5 mg). The reaction was monitored with TLC [R$_f$ = 0.22, CH$_3$OH:CH$_2$Cl$_2$:NH$_4$OH = 6:94:0.6 (v/v/v)], silica gel 60 F$_{254}$, EMD Millipore]. Column chromatography [silica gel, CH$_3$OH:CH$_2$Cl$_2$:NH$_4$OH = 6:94:0.6 (v/v/v)] yielded the desired compound. IR (KBr, cm^{-1}) 3436, 3068, 2945, 2815, 1608, 1592, 1563, 1555, 1503, 1453, 1418, 1347. ^{1}H-NMR (600 MHz, CDCl$_3$) δ 9.03 (s, 2H), 7.40 - 8.06 (m, 10H), 6.87 (s, 2H), 5.02 - 5.42 (m, 5H), 3.75 (s, 2H), 1.74 - 2.54 (m, 13H). ^{13}C-NMR (151 MHz, CDCl$_3$) δ 182.47, 162.59, 149.21, 144.64, 143.65, 138.28, 135.83, 128.69, 125.97, 124.40, 123.26, 122.37, 113.86, 64.88, 56.34, 54.14, 52.27, 50.62, 45.42, 41.50.

Compound 17. White solid (88.3%, 28.1 mg). The reaction was monitored with TLC [R$_f$ = 0.14, CH$_3$OH:CH$_2$Cl$_2$:NH$_4$OH = 6:94:0.6 (v/v/v)], silica gel 60 F$_{254}$, EMD Millipore]. Column chromatography [silica gel, CH$_3$OH:CH$_2$Cl$_2$:NH$_4$OH = 6:94:0.6 (v/v/v)] yielded the desired compound. IR (KBr, cm^{-1}) 3428, 3066, 2940, 2807, 1608, 1561, 1549, 1504, 1452, 1418, 1394, 1347, 1367, 1222. ^{1}H-NMR (600 MHz, CDCl$_3$) δ 9.06 (d, *J* = 2.9 Hz, 2H), 8.18 (d,

J = 7.4 Hz, 2H), 7.98 (s, 2H), 7.54 - 7.67 (m, 6H), 7.04 (d, *J* = 6.7 Hz, 2H), 5.44 (s, 1H), 5.15 (s, 2H), 4.93 (s, 2H), 3.72 (s, 2H), 2.04 - 2.75 (m, 13H), 1.65 (s, 2H). ^{13}C-NMR (151 MHz, CDCl$_3$)

δ 162.98, 149.87, 144.98, 143.90, 139.06, 136.41, 129.17, 126.40, 124.91, 122.69, 121.38, 114.27, 65.92, 55.99, 54.97, 52.98, 45.97, 43.44, 26.54.

Compound 18. White solid (71.5%, 12mg). The reaction was monitored with TLC [R$_f$ = 0.202, CH$_3$OH:CH$_2$Cl$_2$:NH$_4$OH = 5:95:0.5 (v/v/v)], silica gel 60 F$_{254}$, EMD Millipore]. Column chromatography [silica gel, CH$_3$OH:CH$_2$Cl$_2$:NH$_4$OH = 5:95:0.5 (v/v/v)] yielded the desired compound. ^{1}H-NMR (600 MHz, CDCl$_3$) δ 9.10 (d, J = 2.7 Hz, 2H), 8.14 (d, *J* = 7.6 Hz, 2H), 7.94 (s, 2H), 7.74 - 7.37 (6H), 7.05 (d, *J* = 8.2 Hz, 2H), 5.43 - 5.47 (m, 1H), 5.23 (s, 2H), 5.00 (s, 2H), 3.69 (s, 2H), 2.48 (s, 5H), 2.38 - 2.08 (s, 2H), 1.83 (s, 6H). ^{13}C-NMR (151 MHz, CDCl$_3$) δ 150.26, 136.51, 125.05, 118.77, 112.69, 54.23, 66.17, 66.8, 46.05, 43.95, 44.17.

4.2.3. MS Analysis Procedures of Synthesized Compounds

The samples were diluted in methanol to approximately 1 M concentration. A sample solution (1 µL) was introduced by Direct injection without a chromatographic column. High resolution mass spectra (HRMS) were obtained on a JEOL AccuTOF time-of-flight mass spectrometer (Peabody, MA) coupled with an ESI source or Ion sense DART open-air ionization source (Saugus, MA). The solvent system was 1:1 methanol:water. Ions were delivered to the mass spectrometer with an electrospray ionization (ESI) source. The capillary voltage was set to be 4000V for the all the samples. Data were collected in the m/z range of 200-1100. Spectra were collected at a rate of 2 spectra/s allowing for 4800 transients/spectrum.

4.2.4. Preparation of G-quadruplex DNA and Duplex DNA for CD and UV experiments

Three G-quadruplexes with different conformations were used in the study, including **G1** (htelo, hybrid) 5'-AGGG(TTAGGG)$_3$T-3' in K$^+$, **G2** (htelo, antiparallel) 5'-

AGGG(TTAGGG)$_3$T-3'in Na$^+$, and **G3** (c-MYC, parallel) 5'-

TGGGGAGGGTGGGGAGGGTGGGGA-3' in a mixture of Li$^+$ and K$^+$.

The G-quadruplex DNA strand (**G1**, **G2** or **G3,** 100 μM) was dissolved in a mixture (100

μL) of lithium cacodylate buffer (40 mM, pH 7.1) and a corresponding salt (150 mM KCl for

G1, 150 mM NaCl for **G2**, and a mixture of 99 mM LiCl and 1mM KCl for **G3**), heated at 90 °C

for 10 min, slowly cooled to 25°C, and incubated at 4 °C overnight. Duplex **D1** (complementary

strands of S1 and S2, 100 μM each) was dissolved in a mixture of lithium cacodylate buffer (40

mM, pH 7.1) and KCl (150 mM), heated at 90 °C for 10 min, slowly cooled to 25°C, and

incubated at 4 °C overnight. All solutions were stored at 4 °C for further use.

4.2.5. Preparation of G-quadruplex DNA for ESI Mass Spectrometry Experiments

G1 (100 μM) in 150 mM NH$_4$OAc solution was heated to 95 °C for 5 min followed by

cooling slowly to room temperature and incubated at 4°C overnight.

4.2.6. Thermal Denaturation Analysis Monitored by CD

The mixture solutions (200 μL) of (**G1, G2 or G3**) (5 μM) in lithium cacodylate buffer

(40 mM, pH 7.1) and a corresponding salt (150 mM KCl for **G1**, 150 mM NaCl for **G2**, and a

mixture of 99 mM LiCl and 1mM KCl for **G3**) in the absence or presence of a ligand were

prepared. The CD spectra were recorded at a fixed wavelength (290 nm for **G1** and **G2**, 265 nm

for **G3**) as a function of increasing temperature (25-85 ˚C) at a heating rate of 0.5 ˚C/min.

Melting profiles were converted into a folded fraction based on a published procedure[162], and the

melting temperatures were determined at the 0.5 folded fraction.

4.2.7. CD Titration

Aliquots of a stock solution containing compounds were gradually added into a solution

(200 μL) of **G1 (G2 or G3)** (5 μM) in lithium cacodylate (40 mM, pH 7.0) and 150 mM KCl at

25 °C. The final molar ratios of ligands to G-quadruplex were varied from 0 to 5.5. After each addition, the solution was gently mixed and incubated for 5 min for equilibrium before collecting the spectrum. The overall volume change during the titration was kept less than 1% of the initial volume. CD spectra were recorded as a function of wavelength (200-400 mm, 50 nm/min scan speed, 0.5 nm data pitch) using a 1 cm pathlength quartz cuvette. Each spectrum was averaged over three scans.

4.3.8. G-quadruplex Induction

The solution of **G1** (10 μM) in 50 mM Tris.HCl, pH 7.4 buffer was heated at 90 °C for 5 mins and immediately cooled to 0 °C using ice-water slush. The solution was kept in ice and taken out just before CD spectra measurement. Aliquots of stock solution of the ligands were added to vary the molar equivalents in solution from 0 to 10. After each addition, the solution was gently mixed and incubated for 5 min for equilibrium before collecting the spectrum. CD spectra were recorded as a function of wavelength (200-400 mm, 50 nm/min scan speed, 0.5 nm data pitch) using a 1 cm pathlength quartz cuvette. Each spectrum was averaged over three scans and blank subtracted.

4.2.9. Thermal Denaturation Analysis Monitored by UV

The mixture solutions (1 mL) of **D1** (2 μM) in lithium cacodylate buffer (40 mM, pH 7.1) and KCl (150 mM) in the absence or presence of a ligand (8 μM) were prepared. The UV melting curves were recorded at 260 nm as a function of temperature (20-85 °C, heating rate: 0.5 °C/min). Melting profiles were converted into a folded fraction, and the melting temperatures were determined at the 0.5 folded fraction.

4.2.10. Fluorescent Intercalator Displacement Assay

A reaction solution containing 0.25 μM DNA (**G1**, **G2**, **G3**, or **D1**) in lithium cacodylate (10 mM, pH 7.0) and KCl (50 mM) was incubated with 0.5 μM thiazole orange (TO) for 30 min in the dark. Aliquots of a ligand were added to this mixture solution and equilibrated for 5 min in the dark before measurement. The fluorescence spectra (Ex: 501 nm, slit width: 5.0 nm, scan speed 120 nm/min) were recorded from 510-650 nm after each addition. The percent fluorescence displacement (%F) was calculated at λ_{max} by using the formula % FI. displacement=100-[(FA/FA$_0$)×100], where FA$_0$ is the initial fluorescence intensity in the absence of a ligand, and FA is the fluorescence intensity upon addition of a ligand.

4.2.11. Docking Procedure

The geometry optimization of the representative ligand **11** was carried out at the hybrid density functional B3LYP level with 6-311G(d) basis set using Gaussian'09 program package. The Natural bond order (NBO) analysis was carried out to calculate the atomic charges on the optimized geometries. Docking simulations of **11** were carried out on the antiparallel, parallel, and hybrid G-quadruplexes. The structures of hybrid G-quadruplex, antiparallel G-quadruplex, and parallel G-quadruplex were modeled after the solution structures with the PDB id (the RCSB protein data bank) of 5Z80, 2MCC, and 2L7V, respectively. Each initial solution structure is ligand-bound, and the corresponding ligand was manually removed to obtain the G-quadruplex structure for docking. The flexible ligand docking experiments were carried out using the DOCK 6.9 program. Throughout the docking of all three ligands, the same receptor grid was used. Analyses of docked structures and docked image creations were carried out using the UCSF Chimera program.

4.2.12. ESI-MS Analysis of the Complex Formation Between G1 and 11/13

G1 (20 µM) and compound **11/13** (80 µM) were mixed and methanol was added to make

33% methanol-water final concentration. The mixed solutions were injected into a Thermo

Scientific LTQ XL Mass Spectrometer at a rate of 5 µL/min. The ion spray voltage was set to -

4500 V and ion transfer tube temperature to 300 °C. Nitrogen sheath and auxiliary gas flows

were 12 and 0 arbitrary, respectively. MS spectra were recorded in the m/z range from 1400 to

1800 with an average of 150 scans. Data were collected using Thermo Xcalibur 2.2 and

analyzed using the Thermo Xcalibur 4.0 software.

4.3. Results and Discussion

4.3.1 Synthesis of 2-Clip -phen Derivatives

The final compounds were synthesized in three steps, as shown in Schemes 1-3. 2-Clip -

phen was synthesized based on a published procedure[206]. We used 2-bromo-1,10-phenanthroline

instead of 2-chloro-1,10-phenanthroline as the starting material because it is readily available

and cost-effective. Reaction of 2-bromo-1,10-phenanthroline with serinol in the presence of

NaH in anhydrous DMF followed by recrystallization yielded pure 2-Clip-phen (Scheme 4.1). It

is noteworthy that 2-Clip-phen streaks in the silica gel column so that it cannot be easily purified

using column chromatography. In separated reactions, amines 5a-L, 5b-S and AN1-9 were

converted into the corresponding isothiocyanates by reacting with 1,1'-thiocarbonyldi-2(1H)-

pyridone (TCDP) in dry dichloromethane/dimethylformamide. In the final step, 2-Clip-phen

react with isothiocyanates to yield the target products via the thiourea linkage formation in good

yields. The target compounds had very similar mobility to 2-Clip-phen on the silica gel column;

therefore, purification with column chromatography was very time-consuming. Using a large

excess (4-5 molar equivalents) of isocyanates to consume most 2-Clip-phen in the coupling reaction greatly helped the purification step.
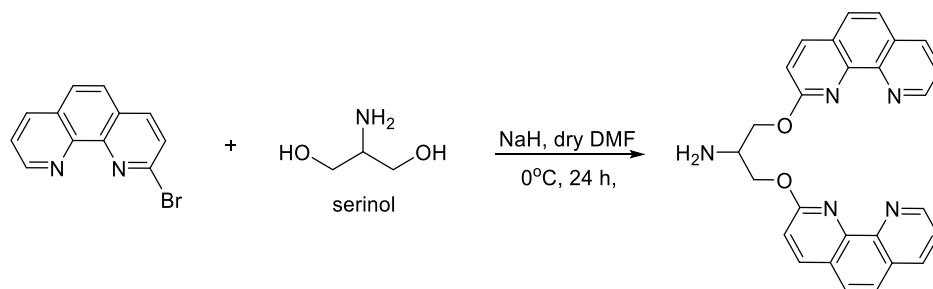
Scheme 4.1. Synthesis of 2-Clip-phen.
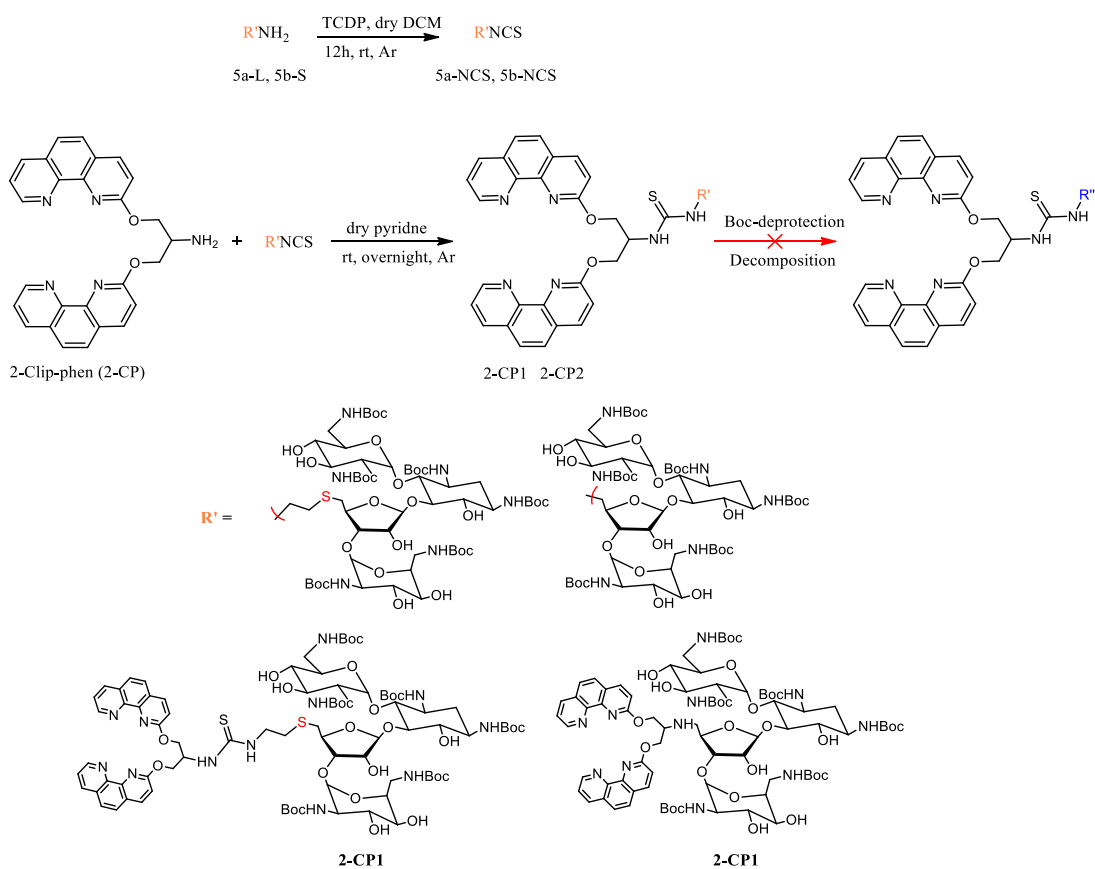
Scheme 4.2. Synthesis of 2-Clip-neomycin derivatives.

R-NH$_2$ + TCDP →[dry DCM, DMAP / rt, N$_2$, 12h] R'-NCS

R =

AN1 AN2 AN3 AN4 AN5 AN6 AN7 AN8 AN9

R' =

AM1 AM2 AM3 AM4 AM5 AM6 AM7 AM8 AM9
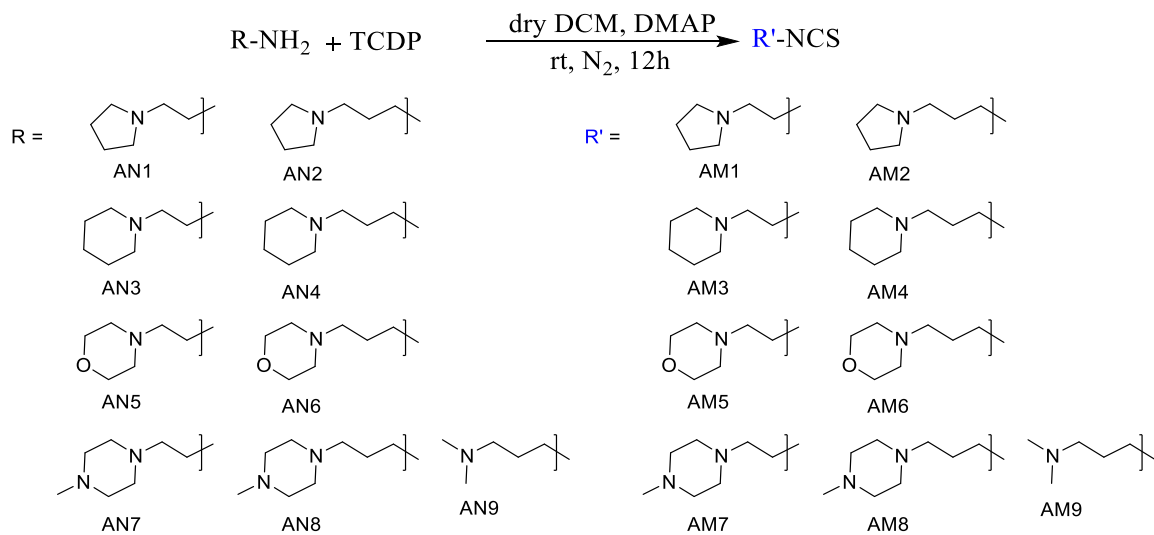
Scheme 4.3. Synthesis of isothiocyanate (**AM1-9**) from primary amines (**AN1-9**). (TCDP is 1-1'-thiocarbonyldi-2(1H)-pyridone).

R-NCS + (2-Clip-phen amine) →[dry Pyridine / rt, N$_2$, 12 h] (thiourea product)

R =

10 11 12 13 14 15 16 17 18
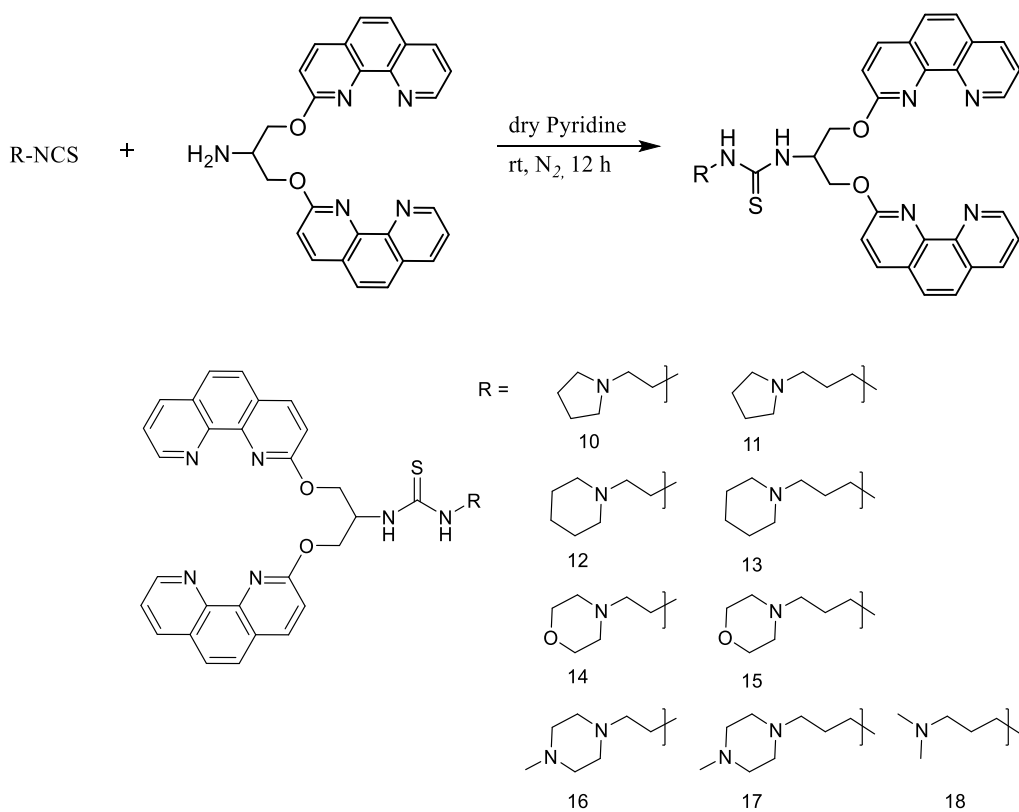
Scheme 4.4. Coupling of isothiocyanates and 2-Clip-phen moieties.

4.3.2. 2-Clip-phen Derivatives Stabilize the Human Telomeric Antiparallel G-quadruplexes Over Other G-quadruplexes

The stabilization of G-quadruplex DNA by 2-Clip-phen derivatives was determined by thermal denaturation studies monitored by circular dichroism (CD). CD spectroscopy is a common technique that is used to analyze the conformation of proteins and DNA. DNA G-quadruplexes can adopt three major conformations, characterized by unique signature peaks in circular dichroism. A typical parallel G-quadruplex topology exhibits a maximum peak (+) around 260 nm and a peak (-) with a maximum around 240 nm[216]. A hybrid G-quadruplex topology[188] typically shows a maximum (+) around 290 nm, a shoulder (+) centered around 260 nm, a weak band (+) with a maximum around 250 nm, and a band (-) with a maximum around 240 nm. An antiparallel G-quadruplex topology shows a maximum (+) around 290 nm and a band (-) with a maximum around 260 nm. CD signals of G-quadruplexes depend on the sugar-phosphate backbone strand orientation, loops composition, helical handedness, and nucleobase orientations[217,186]. For instance, the positive CD peak at 260 nm indicates the parallel strand orientation or anti-anti base steps, and the CD peak at 290 nm most likely shows the antiparallel strand orientation or syn-anti, anti-syn base steps. Hybrid G-quadruplex DNA can have more than one conformation. Its CD peaks at 290 nm and 260 nm come from a combination of nucleobase orientations. The relative intensity of the signals at 290 nm and 260 nm reflects the percentage of nucleobase orientations contributed to the conformation(s). For example, human telomeric DNA and BCL-2 proto-oncogene G-quadruplexes both adopt hybrid conformation(s) under physiological conditions, but their CD signals are slightly different. The CD spectrum of the BCL-2 G-quadruplex is contributed from anti-anti base steps, and that of the human telomeric G-quadruplex is contributed from syn-anti or anti-syn base steps[200].

In the present study, three G-quadruplexes were chosen for thermal denaturation monitored by CD. A hybrid human telomeric G-quadruplex (**G1**) was formed in the presence of 150 mM K$^+$. The same sequence was used to create an antiparallel G-quadruplex in the presence of 150 mM Na$^+$. To form a parallel G-quadruplex (**G3**) conformation, we used a 27mer c-MYC DNA sequence annealed in a mixture of 99 mM Li$^+$ with 1 mM K$^+$. The conformations of three G-quadruplexes were confirmed using circular dichroism **(Figure 4.4)**.
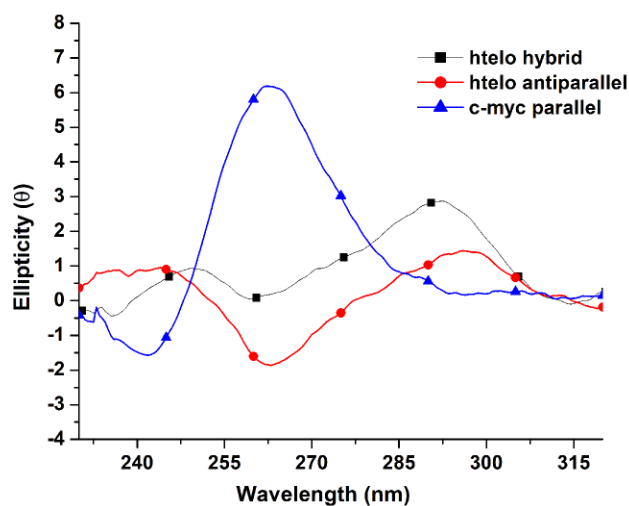
Figure 4.4. The CD spectra of htelo hybrid (**G1**), htelo antiparallel (**G2**) and c-MYC parallel (**G3**) G-quadruplex conformations.
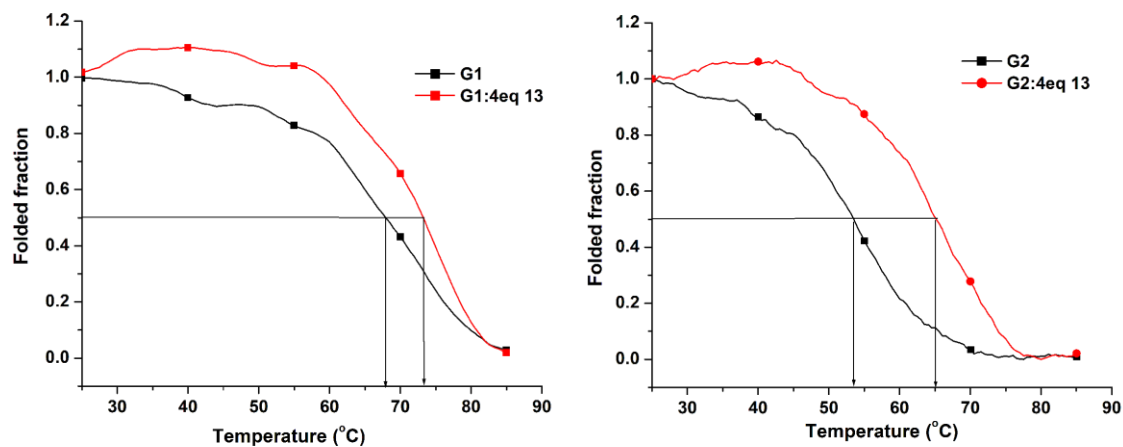
Figure 4.5. Effect of compound **13** on the T_m of htelo hybrid (**G1**) (left), and antiparallel G-quadruplex conformation (**G2**) (right).
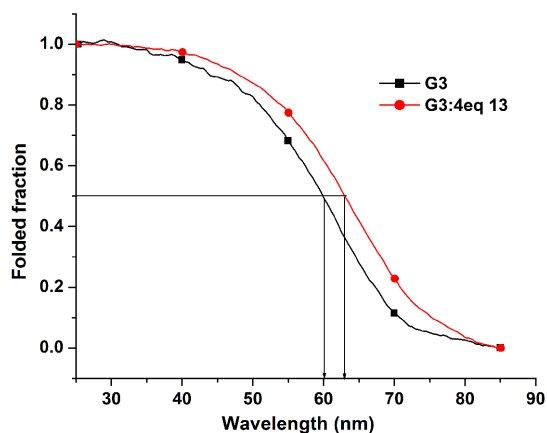
Figure 4.6. Effect of compound **13** on the T_m of c-MYC parallel (**G3**) G-quadruplex.

Thermal stability of G-quadruplex DNA can be probed by analyzing CD ellipticity at 290 nm or 265 nm as a function of increasing temperature. The intensity of the ellipticity decreases as a G-quadruplex dissociates into random coils at high temperatures. The temperature at which 50% of the G-quadruplex dissociates into random coils is defined as melting temperature (T_m).

A melting profile is normally converted into a folded fraction as a function of increasing temperature. The melting temperature can be determined at the folded fraction of 0.5.

The thermal denaturation experiments were conducted with a DNA: ligand ratio of 1:4. All the experiments were duplicated. The representative normalized melting profiles of compound **13** with three G-quadruplex conformations are shown in **figure 4.5** and **4.6**. The melting curves for all the compounds are presented in the appendix C.

All 2-Clip-phen derivatives showed a more significant stabilization effect (Table 4.1) on the antiparallel telomeric G-quadruplex (**G2**) than the other two G-quadruplexes (**G1** and **G3**). 2-Clip-phen had a minor effect on the stabilization of three G-quadruplexes, which increased the T_ms of **G1**, **G2**, and **G3** by 0.8, 2.3, and 2.9 °C, respectively. Amongst all the derivatives, compounds **11** and **13** had the best stabilization effect, which increased the T_m of **G2** by more than 10 °C. All the derivatives slightly increased the T_ms of **G1** and **G3** (1-3 °C in most cases), suggest that they preferentially bind to the antiparallel telomeric G-quadruplex. The effect of amino-containing side chains of the derivatives was determined by comparing the ΔT_ms of the derivatives with that of 2-Clip-phen for each G-quadruplex ($\Delta\Delta T_m$ in Table 4.1). Conjugation of amino-containing side chains to 2-Clip-phen could not enhance its binding to the parallel c-MYC G-quadruplex (**G3**) because the $\Delta\Delta T_m$ values were within the experimental errors. The amino-containing side chains drastically increased the binding of 2-Clip-phen to the antiparallel telomeric G-quadruplex (**G1**) and slightly increased its binding to the hybrid telomeric G-quadruplex (**G2**). The positively charged alicyclic bases or (aliphatic amine) make favorable interactions with the N-H of G-tetrads in the grooves or with negatively charged phosphate backbones. The alicyclic rings pyrrolidine and piperidine showed the most stabilizing effect followed by piperazine and morpholine. The morpholine ring contains an oxygen atom which

may cause electrostatic repulsive interactions with the phosphate backbones and thus decrease the ligand binding to G-quadruplexes[176]. The N-methyl group of the piperazine ring might be involved in some steric interactions within the groove, which renders the stacking interaction with the G-tetrads. The length of the linker seems to play an important role in thermal stabilization. The derivatives with a longer linker showed a slightly better stabilization effect when comparing the corresponding T_m values.

Table 4.1

G-quadruplex Stabilization Potential (ΔT_m) of the 2-Clip-phen (2-CP) Derivatives by CD Thermal Denaturation Study.

	Hybrid G-quadruplex (G1)			Antiparallel G-quadruplex (G2)			Parallel G-quadruplex (G3)		
	T_m	ΔT_m	$\Delta\Delta T_m$	T_m	ΔT_m	$\Delta\Delta T_m$	T_m	ΔT_m	$\Delta\Delta T_m$
C	67.4 ± 0.4			53.3 ± 0.3			60.3 ± 0.6		
2-CP	68.2 ± 1.2	0.8	-	55.6 ± .5	2.3	-	63.2 ± 0.3	2.9	-
10	71.6 ±0.4	4.2	3.4	63.3 ± 1	10	7.7	64.2 ± 0.7	3.9	1.0
11	72.7 ± 0.01	5.3	4.5	63.4 ± 1	10.1	7.8	62.3 ± 0.1	2.0	-0.9
12	71.2 ± 0.4	3.8	3.0	62.9 ± 0.8	9.6	7.3	65.2 ± 1.6	4.9	2.0
13	72.8 ± 0.4	5.4	4.6	64.7 ± 0.5	11.4	8.1	63.2 ± 0.6	2.9	0
14	68.6 ± 0.6	1.2	0.4	58.4 ± 1	5.1	2.8	62.1 ± 0.1	1.8	-0.9
15	70.3 ± 0.4	2.9	2.1	60.4 ± 0.8	7.1	4.8	62.5 ± 0.7	2.2	-0.7
16	70.7 ± 0.7	3.3	2.5	61.4 ± 0.01	8.1	5.8	62.9 ± 1.7	2.6	-0.3
17	70.5 ± 0.1	3.1	2.3	60.5 ± 1	7.2	4.9	62.1 ± 0.5	1.8	-1.1
18	71.1 ± 0.6	3.7	2.9	61.3 ± 1	8.0	5.7	65.5 ± 0.3	5.2	2.3

ΔT_m = melting temperature (T_m) of G-quadruplex with the 4 molar equivalents of compound – melting temperature of G-quadruplex without ligand (C). $\Delta\Delta T_m$ = melting temperature (T_m) of G-quadruplex with the 4 molar equivalents of compound – melting temperature of G-quadruplex 2-CP. C= control

All 2-Clip-phen derivatives showed the binding selectivity for G-quadruplexes over duplex DNA as none of them could stabilize an oligonucleotide duplex DNA (**D1**) (**Figure 4.7**).
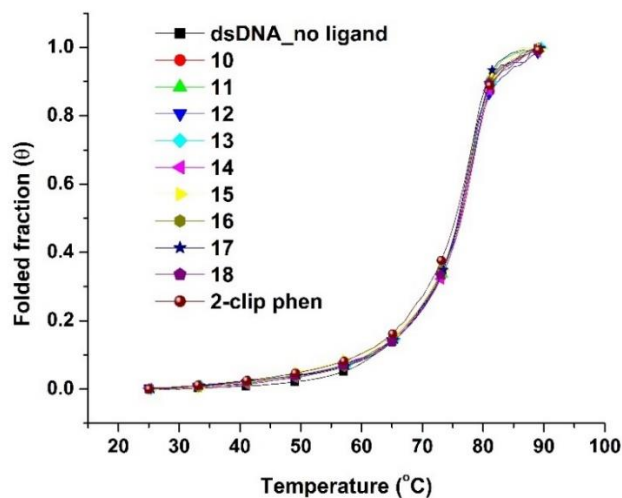
Figure 4.7. Melting of the dsDNA (26 mer) with 2-Clip-phen derivatives.

4.3.3. Binding of 2-Clip-phen Derivatives With G-quadruplex DNA Investigated Using the FID Assay

The binding to 2-Clip-phen derivatives was further investigated with the fluorescent intercalator displacement (FID) assay. The nature of the FID assay has been discussed in Chapter 3. FID assay is a well-established assay to complement the results of thermal denaturation studies[182]. In our study, thiazole orange (TO) was first incubated with G-quadruplex DNA for equilibrium binding. The pre-bound TO was gradually displaced from DNA by the addition of a ligand. The displacement of TO from the bound state leads to a dramatic loss of its quantum yield. The concentration of a ligand required to displace 50% of pre-bound TO from G-quadruplex DNA is defined as $^{G4}DC_{50}$ (**Figure 4.8**).

As shown in Table 4.2, the $^{G4}GC_{50}$ value of 2-Clip-phen is much higher than those of 2-Clip-phen derivatives for each G-quadruplex, suggesting that tethering an amino side chain to 2-Clip-phen can greatly enhance its binding to G-quadruplex DNA. All the 2-Clip -phen

derivatives showed stronger affinities to the antiparallel G-quadruplex (**G2**) than the hybrid G-quadruplex (**G1**). This result is in accordance with the thermal denaturation results mentioned above. The $^{G4}GC_{50}$ value of each 2-Clip-phen derivative for **G3** is much higher than those for **G1** and **G2**, indicating relatively weaker affinities of these ligands towards parallel c-MYC G-quadruplex.

For **G1**, the compounds **11**, **13**, **15**, **17** containing a 3-carbon linker between an alicyclic base and 2-Clip-phen showed lower $^{G4}DC_{50}$ as compared to their analogs **10**, **12**, **14**, **16** with a 2-carbon linker, respectively. The compounds **14** and **15** containing a 2-carbon and 3-carbon linker with the morpholine moiety showed a negligible difference in the $^{G4}DC_{50}$ for the **G1**. The 2-Clip-phen showed a $^{G4}DC_{50}$ value of 33.3 µM, which is around ~18 times weaker than the least potent derivative **12** for **G1**, suggesting the effect of side chains on increasing the binding affinity of 2-Clip-phen. The alicyclic base moiety may provide additional ionic interactions with G-quadruplexes in the grooves. The following trend was found in terms of the alicyclic base moiety in the FID assay for **G2**: **pyrrolidine > piperidine > morpholine > n-methyl piperazine** which is also the trend of the pKa of the alicyclic bases[218,219].
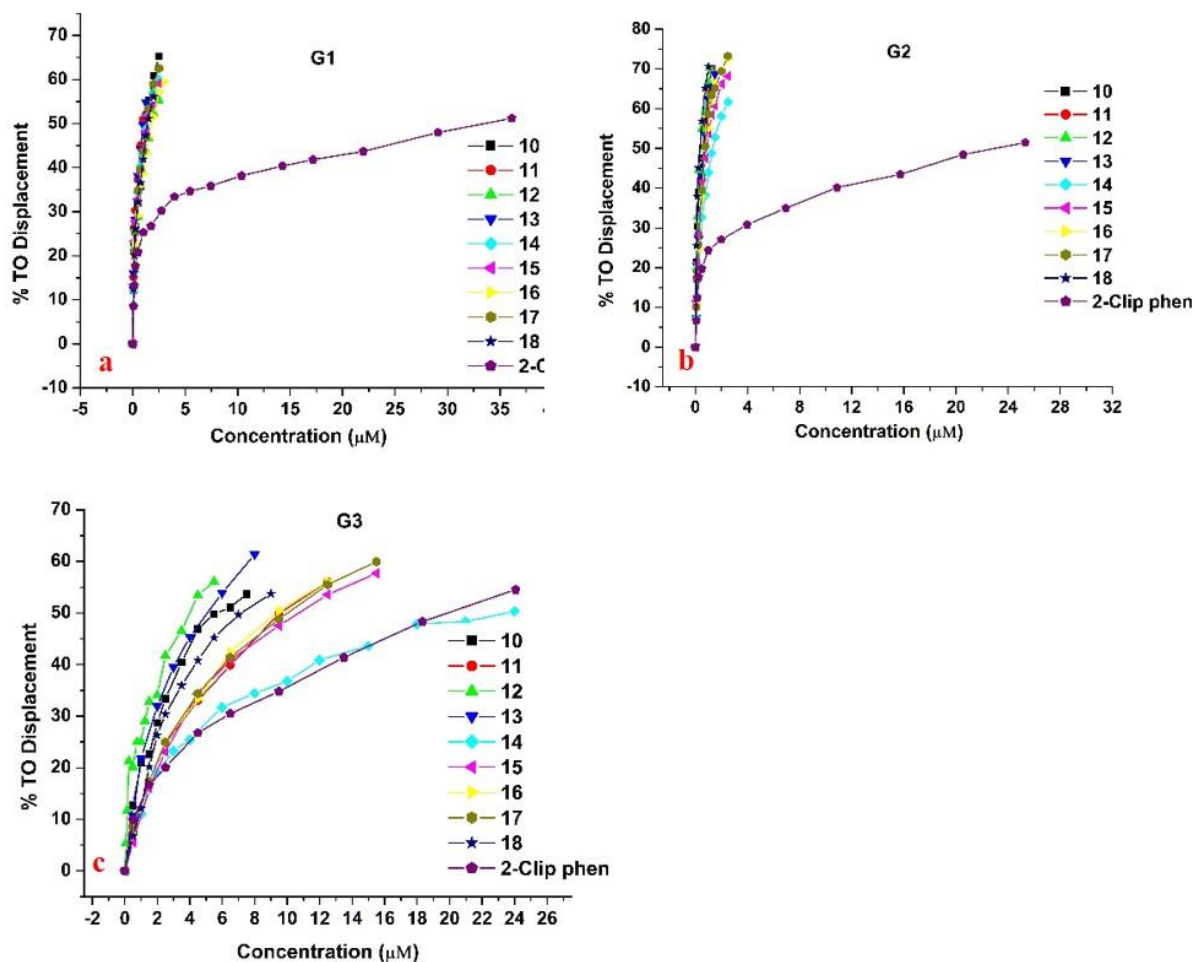
Figure 4.8. Displacement of thiazole orange (TO) by 2-Clip-phen derivatives from a) **G1**, b) **G2** c) **G3** in Fluorescent Intercalator displacement (FID) assay.

G-quadruplex DNA in Na^+ folds into the antiparallel conformation (**G2**). The compounds

11 and **13** with a pyrrolidine and piperidine alicyclic base showed no significant difference in

$^{G4}DC_{50}$ values for **G2**, which is in accordance with the thermal denaturation studies. The 2-

carbon linker containing derivatives of 2-Clip-phen overall showed better $^{G4}DC_{50}$ than the three

carbon linker analogues except for morpholine derivatives. One possible reason could be the

ligand induced G-quadruplex conformation has different groove features in K^+ and Na^+ salt

conditions. The exact reason for this reversal of trend is not known yet. 2-Clip-phen showed a

high $^{G4}DC_{50}$ value of 22.90 µM, which is in accordance with a small thermal stabilization observed.

Table 4.2
$^{G4}DC_{50}$ *Values of 2-Clip-phen Derivatives Against* **G1**, **G2** *and* **G3**.

$^{G4}DC_{50}$ of 2-Clip-phen derivatives against different G4 conformations in µM			
Compounds	**G1**	**G2**	**G3**
10	1.3 ± 0.1	0.6 ± 0.03	5.81 ± 0.78
11	01 ± 0.1	0.6 ± 0.02	9.67 ± 0.47
12	1.8 ± 0.3	0.4 ± 0.07	3.94 ± 0.16
13	1.0 ± 0.1	0.7 ± 0.02	5.13 ± 0.26
14	1.3 ± 0.1	1.3 ± 0.10	23.2 ± 2.5
15	1.4 ± 0.1	0.8 ± 0.08	10.76 ± 0.46
16	1.8 ± 0.2	0.7 ± 0.04	9.49 ± 0.73
17	1.4 ± 0.1	0.7 ± 0.02	10.23 ± 0.63
18	1.4 ± 0.1	0.4 ± 0.03	7.25 ± 1.07
2-CP	33.3 ± 2.9	22.9 ± 0.7	19.9 ± 2.2

2-CP = 2-Clip-phen

2-Clip-phen derivatives showed the least affinity towards parallel stranded c-MYC conformation. Out of all the derivatives, the piperidine-containing derivatives were the most potent, followed by pyrrolidine, N-methyl piperazine and morpholine which is in accordance with thermal denaturation studies. Derivatives with C-2 linker performed better than the ligands with C-3 linker except morpholine containing derivatives. The trend is again in line with the thermal denaturation studies.

The binding of the 2-Clip derivatives to dsDNA was also evaluated using the FID assay. As shown in Table 4.3, none of the derivatives showed the 50% displacement of the at 20 µM concentration. Both **11** and **13** showed the highest displacement of TO at 20 µM, which is 34.3 $\pm$ 2.4 % and 34.5 $\pm$ 3.0 %, respectively. Using a rough estimate, **11** and **13** are at least 28 times

more selective to **G2** than the duplex DNA. The corresponding selectivity to **G1** is at least 20 times over duplex DNA. Compounds showed better selectivity for **G2/G1** over duplex DNA than Phen-DC$_6$. The selectivity profile is equivalent or better than Phen-DC$_3$, one of the best G-quadruplex ligand[220].

Table 4.3
Percentage Displacement of the TO from the Duplex DNA at 20 µM Ligand Concentration.

Compound	TO displacement at 20 µM concentration
10	23.3 ± 3.0
11	34.3 ± 2.4
12	16.7 ± 1.2
13	34.5 ± 3.0
14	24.0 ± 2.1
15	34.8 ± 2.5
16	27.4 ± 4.8
17	28.3 ± 2.90
18	19.8 ± 1.7
2-CP	16.2 ± 1.1

4.3.4. 2-Clip-phen Derivatives Induced Conformational Change in a Hybrid Telomeric G-quadruplex

CD titrations were performed to study the conformation preference of the ligands to three G-quadruplex conformations mentioned above. As shown in **figure 4.4**, the CD spectrum of the htelo hybrid G-quadruplex (**G1**) showed a maximum at 293 nm, a shoulder peak at 270 nm, and a negative band at 235 nm, which signifies a hybrid conformation. The addition of ligand **11** and **13** to the DNA solution led to the increase in the intensity of the peak at 293 nm with a concomitant blue shift of the maximum (287.5 nm at 5.5 molar equivalents of ligand). There is a clear emergence of a new peak at 263 nm (negative band), which intensifies as more ligand stock

is added to the G-quadruplex solution. This observation suggested a ligand-induced transition from the hybrid conformation to the antiparallel conformation in solution. As a representative example, an isoelliptic point at 273 nm when titrating a ligand (**11** or **13**) to **G1** was observed, showing the clear transition to the antiparallel conformation[117]. In addition, titrating the hybrid conformation G-quadruplex (**G1**) with **11** or **13**, the intensity of the peak at ~293 nm increased as a function of increasing ligand concentrations in solution, signifying the conformation transition to the antiparallel conformation (**Figure 4.9a** and **4.9c**). It is known that a concomitant increase in the peak at 295 nm (+) with a blue shift at this wavelength and an increase in the peak at 265 (-) are a signature feature of the antiparallel conformation, resulting from the syn-anti and anti-syn base steps[186]. The change in conformation of hybrid htelo (**G1**) to antiparallel conformation (**G2**) has also been observed with other ligands such as 1,10-phenanthroline scaffold[214,221], natural product Telomestatin[222], metal porphyrins[117] and bis quinolinium derivative of pyridine[221].

Previous studies suggested that the **G1** sequence in K^+ can form an ensemble of different conformations in rapid equilibrium, with the hybrid conformation being the most stable[62,223]. Therefore, the CD spectrum of a hybrid G-quadruplex could be the weighted average spectrum of all conformers present in solution. The shoulder peak at 270 nm results from a minor 2-tetrad antiparallel G-quadruplex[136], in which guanines stack in exclusively alternating orientations and a triplet of guanines stacks on one terminal G-tetrad[136,223]. Two-tetrad G-quadruplex also called hybrid-3 is well characterized and is found to exist in equilibrium with hybrid conformations in the solution[223,224]. The presence of 5'-flanking sequences is required for its stabilization in solution[223]. The disappearance of the shoulder peak at 270 nm upon addition of **11** or **13** (**Figure 4.9b** and **4.9d**) indicates the interaction of the ligand with this 2-tetrad G-quadruplex. This

disappearance was observed at a low ligand concentration (0.5 equivalent of Compound **11**),

suggesting that the percentage of this 2-tetrad G-quadruplex in solution is small. A previous

study reported a less than 10% contribution[226] of the 2-tetrad G-quadruplex to the CD spectrum

of the hybrid G-quadruplex, which is consistent with our result. The disappearance of the peak

at 270 nm also suggests that **11** or **13** could stack on top of the end G-tetrad in the 2-tetrad G-

quadruplex and consequently disrupt the stacking of the guanine triplet at this location. The

ligand ability to induce the hybrid to antiparallel conformational change is weaker for 2-Clip-

phen (without any side chains) because higher ligand concentrations are needed. 2-Clip-phen

derivatives with an amino side chain can bind to antiparallel G-quadruplexes more tightly than 2-

Clip-phen, facilitates the transformation. At this point, it is inconclusive whether the hybrid G-

quadruplex transforms into the 3-tetrad or 2-tetrad antiparallel G-quadruplex or both in solution.

More characterization will be discussed in the later section when we discuss the results of ESI

mass spectrometry.
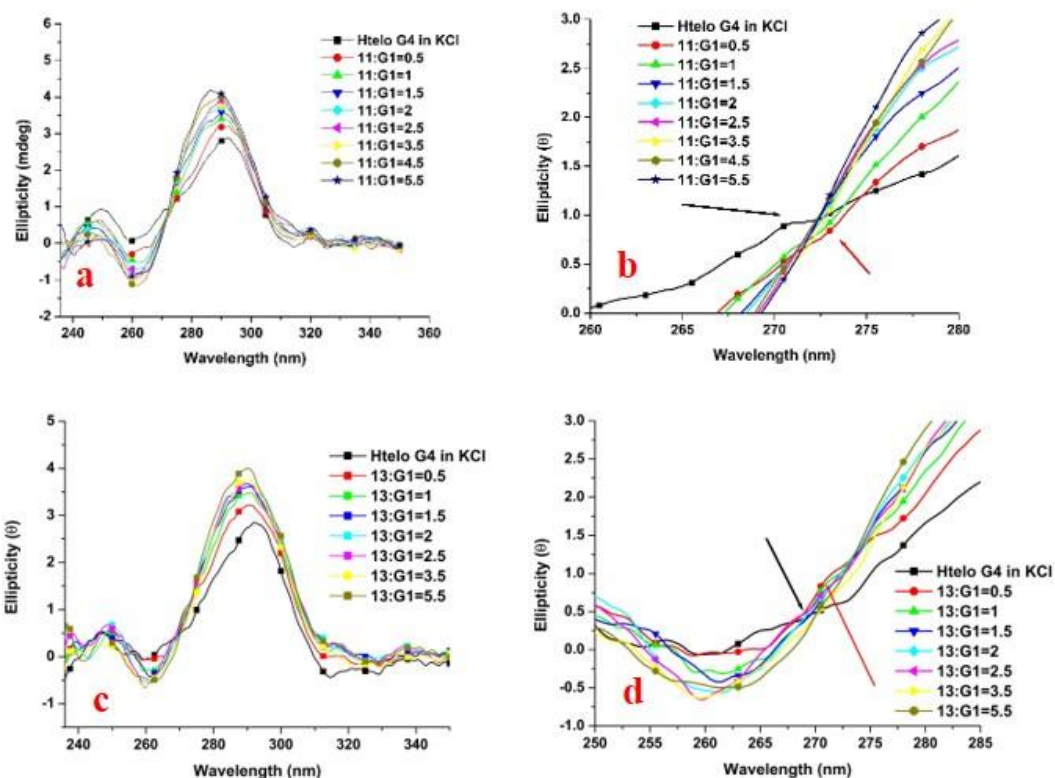
Figure 4.9. CD titrations of compound **11** with **13** to **G1** in 150 mM KCl. a) Ligand **11** to **G1**. b) Disappearance of shoulder peak at 270 nm with 0.5 equivalent of **11** in solution. c) Ligand **13** to **G1**. d) Disappearance of shoulder peak at 270 nm at 0.5 equivalent of **13** in solution.

By contrast, the addition of 2-Clip-phen derivatives to the c-MYC parallel G-quadruplex did not show any conformational changes even in the presence of more than 5.5 equivalents of ligands (**Figure 4.10**). This G-quadruplex is very stable, so the transformation from parallel to antiparallel could be thermodynamically unfavorable. The weak binding of 2-Clip-phen derivatives cannot promote the transformation as well.
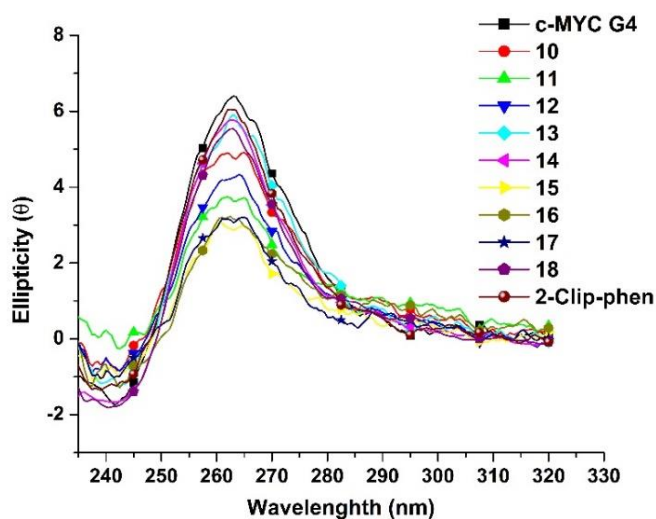
Figure 4.10. c-MYC G-quadruplex (black line) with 5.5 equivalents of 2-Clip-phen derivatives.

4.3.5. 2-Clip-phen Derivatives Can Even Induce Antiparallel G-quadruplex Formation in the Salt (K$^+$) Free Conditions

The ability of the ligands to induce the G-quadruplex formation from random coils in a salt-deficient condition was studied. Compound **11** and **13** were chosen as a representative example. Monovalent alkali metal ions like Na$^+$ and K$^+$ are necessary to stabilize G-quadruplex structures by sitting in/near the center cavity of G-tetrads and coordinating with O6 of guanines. The energy of this coordination almost entirely drives the folding of G-quadruplex with little or no contribution from other weak interactions[225]. The ligands were analyzed for their ability to induce the folding of G-quadruplex DNA from random coils in the absence of K$^+$/Na$^+$. The htelo DNA (10 µM) in 50mM Tris•HCl (K$^+$ free buffer) was heated to 90°C for 5 mins and then rapidly cooled in ice water to ensure the random coil of the DNA, characterized by the positive maximum at 255 nm and negative maximum ~240 nm (red line in **figure 4.11**). The addition of 0.5 equivalent of **11 and 13** followed by 5 min equilibrium time leads to induction of antiparallel

G-quadruplex characterized by the appearance of peak (positive) at 295 nm along with the peak (negative) at 265 nm. The addition of stoichiometric amounts of the ligand leads to further amplification of the peaks at 295 nm and 265 nm. The transition from random coils to folded G-quadruplex DNA is fast and takes only a few mins. The presence of the isoelliptic point at 272 nm suggests a clean transition to the antiparallel G-quadruplex upon the addition of **11.** A similar but much weaker effect was reported by Balasubramanian group[117] on a tetramethylpyridiniumporphyrazines class of G-quadruplex ligands,

To determine the effect of the side chain in **11** and **13**, the ability of the 2-Clip-phen for the induction of G-quadruplex from random coils was analyzed (**Figure 4.12**).
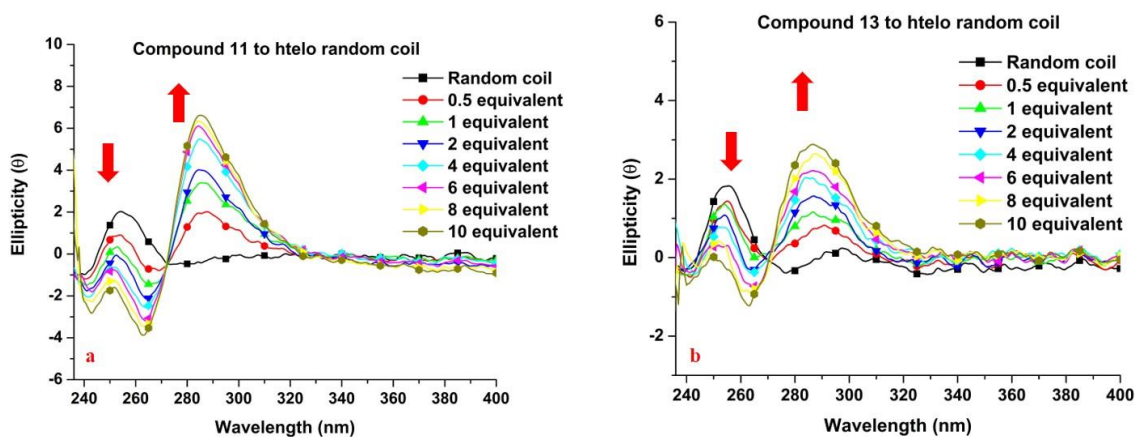
Figure 4.11. CD spectrum of non-annealed 10µM htelo DNA in a K$^+$/Na$^+$ free Tris.HCl buffer, pH 7.4 (black) and in the increasing stoichiometric amounts of ligands a) **11** and (b) **13**.

The comparable CD spectrum of 2-Clip-phen was weaker at all the concentrations as seen from the intensity of the peak at 295 nm (due to antiparallel G-quadruplex) and the peak at 255 nm (due to random coil), which is due to the antiparallel G-quadruplex and random coils of DNA, respectively. The peak at 255 nm at 10 equivalents of 2-Clip-phen has relatively the same

amplitude as peak at 295 nm, signifying much weaker influence of 2-Clip-phen scaffold alone on the induction of G-quadruplex from random coil.

Compound **11** and **13** were able to induce the formation of G-quadruplex from random coils even at a very low ligand to DNA molar ratio (0.5). There is not much change in intensity at 295 nm when th ligand to DNA molar ratio is over 6, as shown in **figure 4.9**. All the 2-Clip-phen derivatives were tested for their ability to induce G-quadruplex formation in K$^+$/Na$^+$ free buffer at both 0.5 and 6 molar equivalents for comparison (**Figure 4.13a** and **Figure 4.13b**). It is evident that **11** and **13** are the most potent ligands based on the intensity changes of peaks at 295 nm at both 0.5 and 6 molar equivalents of ligand concentrations. This data supports the thermal denaturation studies.
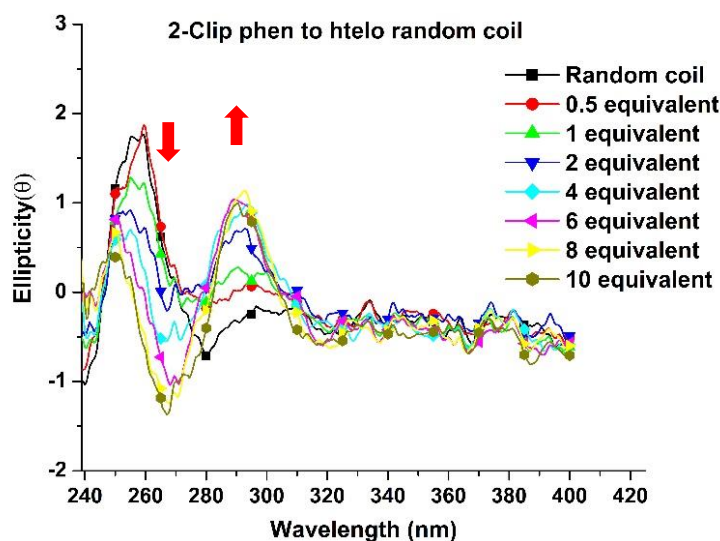
Figure 4.12. CD spectrum of random coil in a K$^+$ free buffer (black). The addition of aliquots of 2-Clip-phen to the solution lead to the appearance of the peaks at 295 nm (+ ve), 260 nm (- ve) with concomitant decrease in intensity at 255 nm. There is no significant increase in the amplitude of the peak at 295 nm after addition of 6 molar equivalents of 2-Clip-phen.
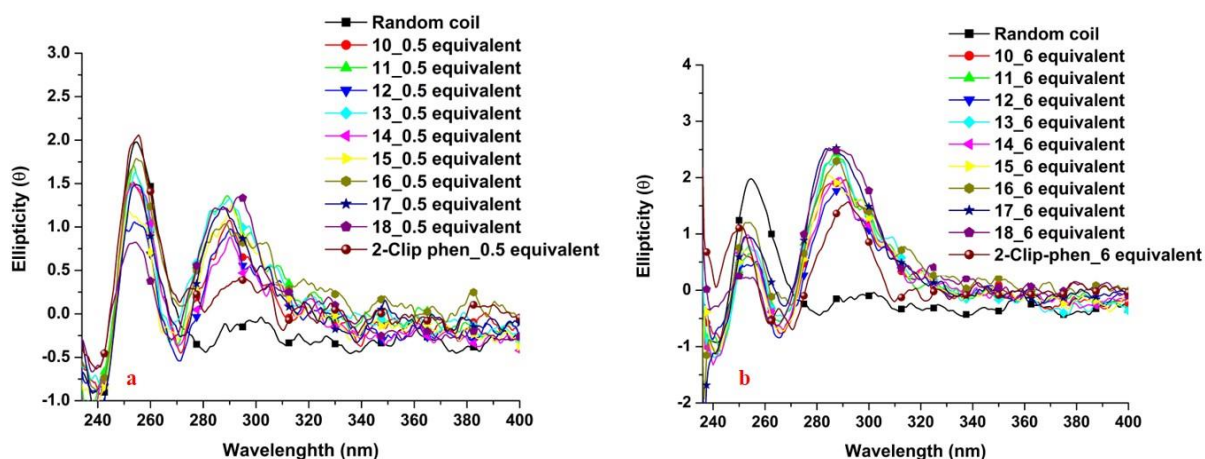
Figure 4.13. CD spectrum of the random coil in a K^+ free buffer (black). a) Addition of the 0.5 molar equivalents of 2-Clip-phen derivatives lead to appearance of peaks at 295 nm (+ve), 260 nm (-ve) with concomitant decrease in intensity at 255 nm. b) CD spectra of the mixture solutions with 6 molar equivalents of ligands.

The G-quadruplex (the **G1** sequence) induced by both **11** and **13** in a salt-free condition are stable, which can be readily observed by thermal denaturation. The melting temperature (T_m) of the induced G-quadruplex in the presence of both **11** and **13** (a 4:1 ligand to DNA ratio) is 60 °C (**Figure 4.14**). By contrast, without **11** or **13**, this sequence is random coils in solution. No melting transitions were observed.
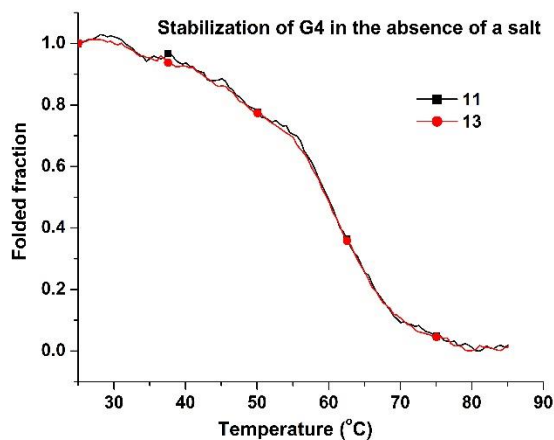
Figure 4.14. CD thermal denaturation of induced G-quadruplex.

4.3.6. Compound 11 Binds the Antiparallel Htelo G-quadruplex With 2:1 Binding Stoichiometry

The changes in ellipticity in CD spectrum were plotted against the ligand/DNA ratio[226]. As shown in the previous section, **11** can induce the formation of the antiparallel G-quadruplex in the absence of salts but has a little effect on the antiparallel G-quadruplex in 150 mM NaCl. Therefore, the CD spectra for the induction of G-quadruplex from random coils in a salt-free buffer and in 150 mM KCl were used to investigate the binding stoichiometry. The ellipticity at 292 nm was plotted against the ligand to DNA molar ratios. The graphs are shown **in figure 4.15**. The intersection point in the graphs corresponds to an approximately 2:1 ligand to DNA binding ratio, suggesting that two ligand molecules bind to one G-quadruplex molecule. UV/Vis titrations were not used for this study because the ligands have a UV λ_{max} at 270 nm, which interferes with the absorption of DNA in the same region.
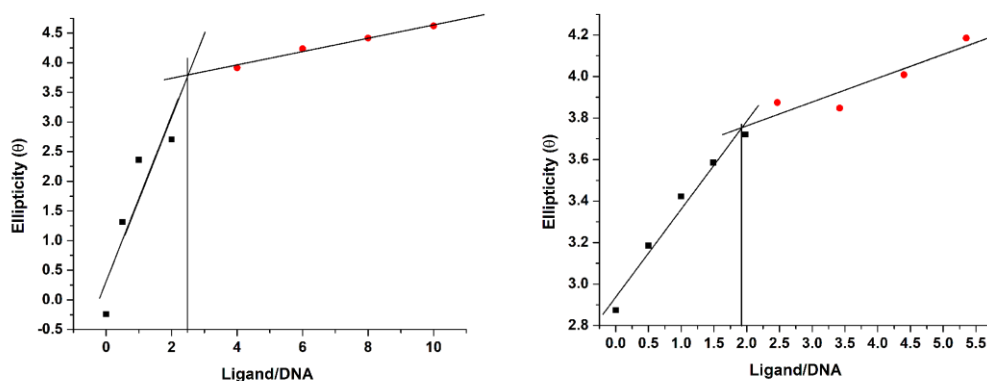
Figure 4.15. CD Job plot of **11** against **G1** random coil in 50mM Tris.HCl buffer, pH7.4 (left) and against folded **G1** in 150 mM KCl and 40 mM LiCaco buffer (right).

4.3.7. 2-Clip-phen Derivatives Induced a Two-tetrad G-quadruplex Conformation in Solution

The binding stoichiometry of the ligands to DNA and the nature of bound complexes were further investigated with electrospray ionization (ESI) mass spectrometry. As explained in Chapter 3, the number of ammonium ions complexed with a G-quadruplex DNA provides the nature of the G-quadruplex under study. If the G-quadruplex complex contains only one NH_4^+, it is a two-tetrad structure because two G-tetrads can only sandwich one NH_4^+ ion. If the G-quadruplex complex contains two NH_4^+ ions, it is composed of three G-tetrads. Previous CD studies suggested that 2-Clip-phen derivatives can induce the transformation of **G1** (hybrid) into antiparallel G-quadruplex(es) although the identity of the antiparallel G-quadruplex (2-tetrad or 3-tetrad) is uncertain.

In the study, compounds **11** and **13** were analyzed using an ESI-LTQ mass spectrometer for their binding stoichiometry and relative binding affinity for G-quadruplex DNA. A solution of G-quadruplex sequence (G_{seq}) annealed in 150 mM NH_4^+ was used as a control (**G1**) (**Figure 4.16a**). The ion at m/z 1460.36 corresponds to the **G1** $[G_{seq} + 2NH_4^+ - 7H^+]^{5-}$ anion.

The inclusion of two NH_4^+ ions suggests the formation of a 3-tetrad G-quadruplex (**G1**) in solution. Previous CD result shows this sequence folds into a hybrid G-quadruplex in NH_4^+, which is preserved in the gas phase[227]. The addition of 4 equivalents of ligand **11** to the G-quadruplex solution leads to the complexation of nearly all the G-quadruplex in solution. Only a small peak representing the free **G1** was observed (**Figure 4.16a** and **b**). Two new ion peaks at m/z 1432.64 and 1719.18 emerge, representing the $[G_{seq} + NH_4^+ + DMF + 2x(\textbf{11}))]^{6-}$ and $[G_{seq} + NH_4^+ + DMF + 2x(\textbf{11})]^{5-}$ ions, respectively, suggesting the formation of a 2:1 ligand to DNA complex with only one ammonium ion associated with the complex ion peaks. It shows that the G-quadruplex folds into a two-tetrad conformation (**Q1**) as only one NH_4^+ ion is associated with the ligand-DNA complex. We also observed the peaks at m/z 1435.91 and 1723.36, $[G_{seq} + 2NH_4^+ + DMF + 2x(\textbf{11}))]^{6-}$ and $[G_{seq} + 2NH_4^+ + DMF + 2x(\textbf{11}))]^{5-}$ respectively which showed the inclusion of two NH_4^+ ions in the complex ion peak along with two molecules of **11** in the complex ion peak (**Figure 4.16b**). **Q1** and **G1** were calculated according to the formula, **Q1** or **G1** = $[(G_{seq}) + (NH_4^+)_y - x(H^+)]^{-x+y}$, where G_{seq} is a unfolded DNA sequence. The ligands do not contribute to the overall charge of the complex ion peak in our case.
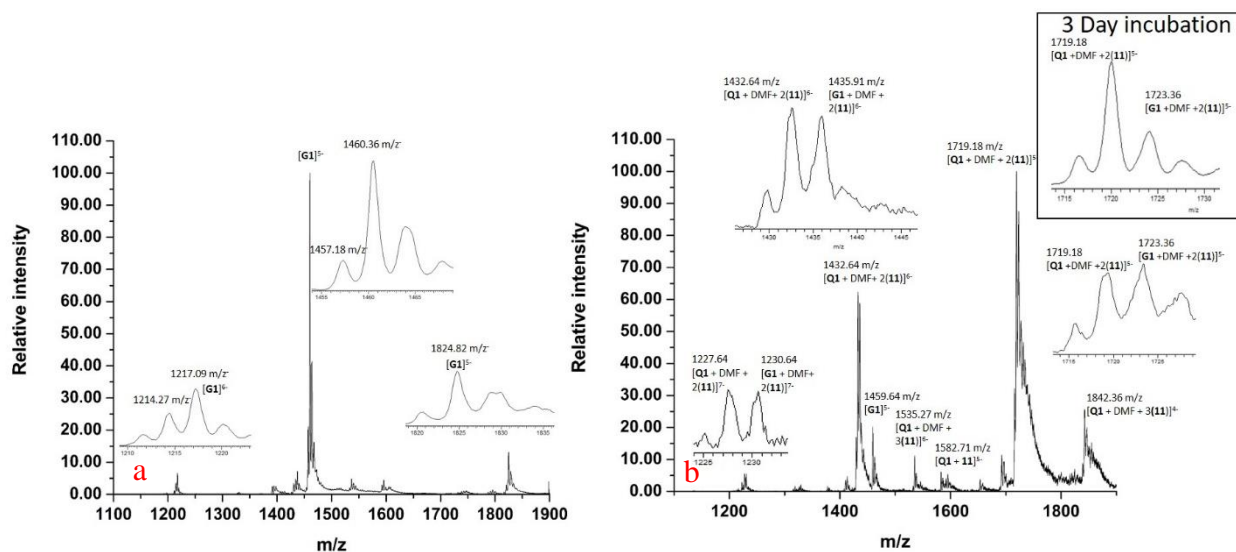
Figure 4.16. Mass spectrometry analysis of the G-quadruplex. a) **G1** without ligand. b) Mixture solution of **G1** and **11** ([**G1**]:[**11**] = 1:4). The peak with maximum intensity is set as the base peak (100%).

To confirm the result is not due to the experimental condition, an ESI experiment on the complexation of a phenanthroline-neomycin conjugate (Chapter 3, **7b**) with **G1** was carried out. As expected, this ligand forms a 1:1 ligand and **G1** complex with the inclusion of two NH_4^+ ions, indicating the presence of a 3-tetrad G-quadruplex (**Figure 4.17a**). It is noteworthy that our previous studies (Chapter 3) showed no change in conformation of the hybrid G-quadruplex occur on the addition of **7b** to **G1**[227].

The relative intensities of **Q1** and **G1** in the mass spectra changed when we prolonged the incubation time (from 30 minutes to 3 days) of the ligand (**11** or **13**) with **G1**. As shown in **figure 4.15b** and 4.16b, after a 3-day incubation time, the intensities of complex ions representing 2-tetrad G-quadruplex (**Q1**) increased, suggesting that 2-Clip-phen derivatives could bind more preferentially to the 2-tetrad antiparallel G-quadruplex over the 3-tetrad G-quadruplex (antiparallel or hybrid).

Overall, both **11** and **13** binds with the 2-tetrad and 3-tetrad G-quadruplexes in a 2:1 ligand to DNA stoichiometry. The mass spectral data is consistent with the results of CD studies.
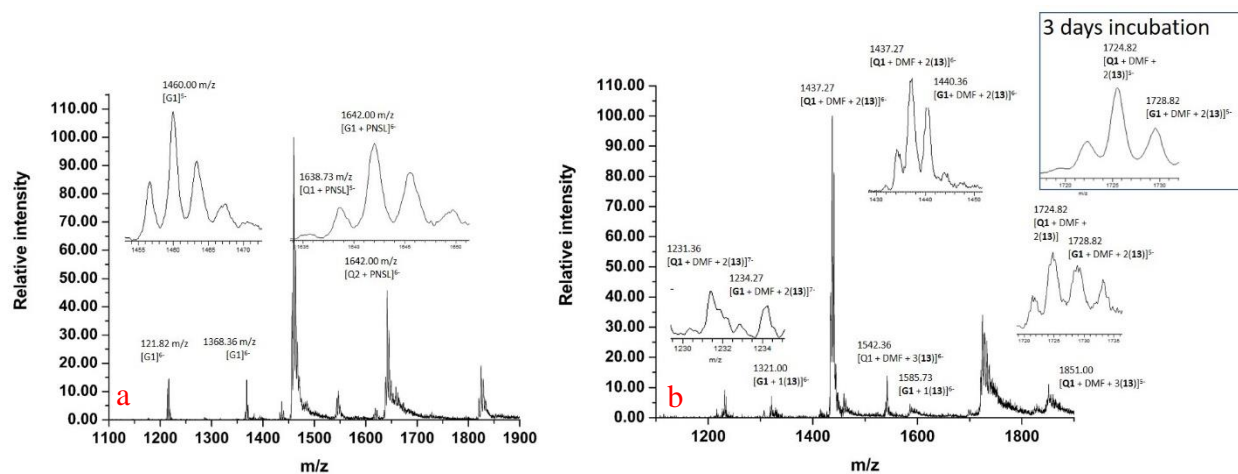
Figure 4.17. Mass spectrum of the mixture solution of a) **G1** and **7b** ([**G1**]:[**7b**] =1:2) and b) **G1** and **13** ([**G1**]:[**13**] =1:4) mixture solution on the right.

4.3.8. Docking Studies

The binding interactions of 2-Clip-phen derivatives with three G-quadruplexes were further studied with docking simulations. Compound **11** was chosen as a representative ligand. The Grid score and energy attribution of compound **11** with antiparallel, parallel, and hybrid G-quadruplexes are shown in Table 4.5. Relative binding interactions involved in stabilization of the ligand are primarily non-polar in nature due to the large contribution of the van der Waals interactions to the binding (Table 4.5).

Table 4.4

*DOCK 6 Grid Score and Energy Attribution of Ligand **11** With Three Different G-quadruplexes.*

	G-quadruplex conformation		
Receptor	**Antiparallel**	**Parallel**	**Hybrid**
van der Waals	-50.3	-58.0	-48.9
Electrostatic	-12.1	-10.0	-15.2

Ligand **11** binds with a relatively higher electrostatic contribution to hybrid G-quadruplex (**Figure 4.18a**). The two nitrogen atoms at the 1 and 10 positions of phenanthroline are probably interacting with the N-H of the guanines (G10, G18 in the top and G11, G17 in the middle tetrad) in the grooves. The amino-containing side chains point up, making no bonding contribution in this state. For the antiparallel G-quadruplex (**Figure 4.18b**), ligand **11** binds with the top G-tetrad through favorable π- π interactions, and One of the phenanthroline rings interacts with the lateral loop nucleobases with the ligand side chain pointing into the groove. We previously observed that 2-Clip-phen stabilizes **G2** poorly; therefore, the strong polar interactions in the groove between the basic nitrogen of the side chain are a plausible explanation for the effect of side chains on antiparallel G-quadruplex (**G2**) stabilization. The docking results indicate more contacting points of **11** with antiparallel G-quadruplex (**G2**), clearly showing that the binding of **11** to antiparallel G-quadruplex is better than that to hybrid G-quadruplex. This result is consistent with our CD and thermal denaturation studies. Parallel c-MYC G-quadruplex used in the study has propeller loops, which might be hampering the binding of the side chain in the groove. One of the phenanthroline rings and the side chain interact with flanking sequences (**Figure 4.18c**). The binding to the c-MYC G-quadruplex is governed mainly by the van der Waals interactions means that the side chain is not able to make favorable electrostatic

interactions. It is necessary to point out that the docking studies on the 2-tetrad G-quadruplex cannot be carried out at this moment due to the lack of reliable PDB structures.
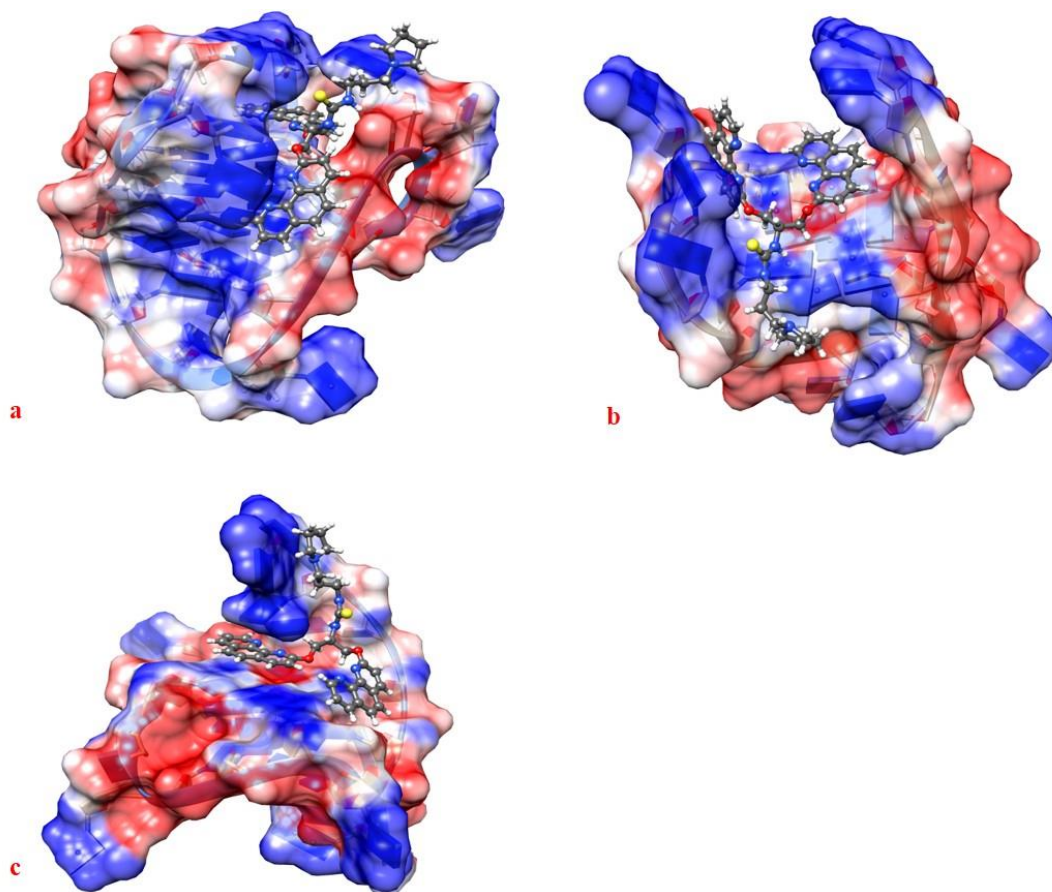
Figure 4.18. Docked structure of ligand **11** to different G-quadruplex conformations. The G-quadruplex is represented as surface colored by electrostatic potential and ligand **11** is represented as a ball and stick. a) Ligand interactions with the hybrid conformation, the two rings are involved in stacking and groove interactions. b) Ligand **11** interacting with antiparallel conformation and c) **11** is interacting with the parallel G-quadruplex preferably through π-π interactions.

4.4. Conclusion

The 2-Clip-phen derivatives show a remarkable preference to G-quadruplexes. They can induce antiparallel G-quadruplex from the random coils in a Na^+/K^+ free buffer and stabilize it to a great extent (Tm = 60 °C). In addition, the ligands can induce the formation of antiparallel conformation from the already folded hybrid conformation under the conditions suitable for the formation of hybrid conformation. A plausible mode of action is shown in **figure 4.19**. The 2-Clip-phen derivatives could bind to both 2-tetrad and 3-tetrad antiparallel G-quadruplexes. The binding of a 2-Clip-phen derivative to the 2-tetrad G-quadruplex disrupts its guanine triplet capping. When binding to the antiparallel G-quadruplex conformation, two phenanthroline moieties of 2-Clip-phen derivatives can interact individually with the loops nucleotides as observed in the docking study. The inability to be in a plane also greatly affects its ability to interact with c-MYC parallel G-quadruplex. The additional specificity could result from steric interactions between the side chain and edge reversal loop in the groove, which forces the side chain interacts with the flanking sequence rather than being in the groove (**Figure 4.18c**).

The discovery of 2-Clip-phen preferentially binding to antiparallel G-quadruplex DNA is encouraging. 2-Clip-phen derivatives are capable of coordinating with Cu^{2+}, which makes a fixed planar aromatic surface. The resulting metal complexes should have even better G-quadruplex binding affinities. In addition, The Cu^{2+}-2-Clip-phen complexes have the ability to cleave DNA; therefore, our 2-Clip-phen derivatives could be used as G-quadruplex specific cleaving agents. The evaluation will be reported in due course.
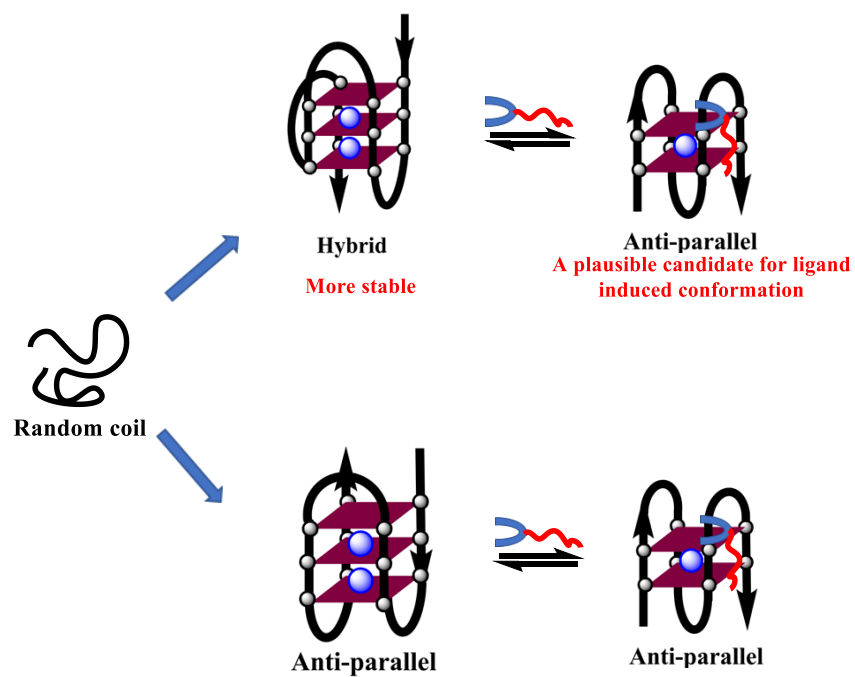
Figure 4.19. Proposed model to explain the different extent of thermal stabilization observed in NaCl and KCl salt solutions.

REFERENCES

1.   CRICK, F. Central Dogma of Molecular Biology. Nature **227**, 561–563 (1970).
2.   STEHELIN, D., VARMUS, H. E., BISHOP, J. M. & VOGT, P. K. DNA related to the transforming gene(s) of avian sarcoma viruses is present in normal avian DNA. Nature **260**, 170–173 (1976).
3.   Reddy, E. P., Reynolds, R. K., Santos, E. & Barbacid, M. A point mutation is responsible for the acquisition of transforming properties by the T24 human bladder carcinoma oncogene. Nature **300**, 149–152 (1982).
4.   Taparowsky, E., Shimizu, K., Goldfarb, M. & Wigler, M. and Activation of the Human N-MS Gene. Cell **34**, 581–586 (1983).
5.   GOODMAN, L. S. *et al*. NITROGEN MUSTARD THERAPY: Use of Methyl-Bis(Beta-Chloroethyl)amine Hydrochloride and Tris(Beta-Chloroethyl)amine Hydrochloride for Hodgkin's Disease, Lymphosarcoma, Leukemia and Certain Allied and Miscellaneous Disorders. J. Am. Med. Assoc. **132**, 126–132 (1946).
6.   Hubbard, R. D. & Fidanze, S. Alkylating and Platinum Antitumor Compounds. Compr. Med. Chem. II **7**, 129–148 (2006).
7.   Farber, S., Diamond, L. K., Mercer, R. D., Sylvester, R. F. & Wolff, J. A. Temporary Remissions in Acute Leukemia in Children Produced by Folic Acid Antagonist, 4-Aminopteroyl-Glutamic Acid (Aminopterin). N. Engl. J. Med. **238**, 787–793 (1948).
8.   Elion, G. B., Singer, S. & Hitchings, G. H. ANTAGONISTS OF NUCLEIC ACID DERIVATIVES: VIII. SYNERGISM IN COMBINATIONS OF BIOCHEMICALLY RELATED ANTIMETABOLITES. J. Biol. Chem. **208**, 477–488 (1954).
9.   HEIDELBERGER, C. *et al*. Fluorinated Pyrimidines, A New Class of Tumour-Inhibitory Compounds. Nature **179**, 663–666 (1957).
10.   DeVita, V. T. & Chu, E. A history of cancer chemotherapy. Cancer Res. **68**, 8643–8653 (2008).
11.   Johnson, I. S., Armstrong, J. G., Gorman, M. & Burnett, J. P. The Vinca Alkaloids: A New Class of Oncolytic Agents. Cancer Res. **23**, 1390 LP – 1427 (1963).
12.   Weaver, B. A. How Taxol/paclitaxel kills cancer cells. Mol. Biol. Cell **25**, 2677–2681 (2014).
13.   Lightfoot, H. L., Goldberg, F. W. & Sedelmeier, J. Evolution of Small Molecule Kinase Drugs. ACS Med. Chem. Lett. **10**, 153–160 (2019).
14.   Capdeville, R., Buchdunger, E., Zimmermann, J. & Matter, A. Glivec (ST1571, imatinib), a rationally developed, targeted anticancer drug. Nat. Rev. Drug Discov. **1**, 493–502 (2002).
15.   Slamon, D. J. *et al*. Human breast cancer: Correlation of relapse and survival with amplification of the HER-2/neu oncogene. Science (80-.). **235**, 182–191 (1987).
16.   Hudis, C. A. Trastuzumab — Mechanism of Action and Use in Clinical Practice. N. Engl. J. Med. **357**, 39–51 (2007).
17.   Sanderson, B. J. S. & Shield, A. J. Mutagenic damage to mammalian cells by therapeutic alkylating agents. Mutat. Res. - Fundam. Mol. Mech. Mutagen. **355**, 41–57 (1996).
18.   Zhou, J. & Giannakakou, P. Targeting microtubules for cancer chemotherapy. Curr. Med. Chem. - Anti-Cancer Agents **5**, 65–71 (2005).

19. Kaye, S. B. New antimetabolites in cancer chemotherapy and their clinical impact. Br. J. Cancer **78**, 1–7 (1998).
20. Ho, D., Yan, L., Iwatsubo, K., Vatner, D. E. & Vatner, S. F. Modulation of β-adrenergic receptor signaling in heart failure and longevity: Targeting adenylyl cyclase type 5. Heart Fail. Rev. **15**, 495–512 (2010).
21. Gorre, M. E. *et al*. Clinical Resistance to STI-571 Cancer Therapy Caused by BCR-ABL Gene Mutation or Amplification. Science (80-.). **293**, 876 (2001).
22. Hochhaus, A. *et al*. Molecular and chromosomal mechanisms of resistance to imatinib (STI571) therapy. Leukemia **16**, 2190–2196 (2002).
23. Torka, P., Barth, M., Ferdman, R. & Hernandez-Ilizaliturri, F. J. Mechanisms of Resistance to Monoclonal Antibodies (mAbs) in Lymphoid Malignancies. Curr. Hematol. Malig. Rep. **14**, 426–438 (2019).
24. Pohlmann, P. R., Mayer, I. A. & Mernaugh, R. Resistance to trastuzumab in breast cancer. Clin. Cancer Res. **15**, 7479–7491 (2009).
25. McLuckie, K. I. E. *et al*. G-Quadruplex DNA as a Molecular Target for Induced Synthetic Lethality in Cancer Cells. J. Am. Chem. Soc. **135**, 9640–9643 (2013).
26. Rodriguez, R. *et al*. A novel small molecule that alters shelterin integrity and triggers a DNA-damage response at telomeres. J. Am. Chem. Soc. **130**, 15758–15759 (2008).
27. Piazza, A. *et al*. Genetic instability triggered by G-quadruplex interacting Phen-DC compounds in Saccharomyces cerevisiae. Nucleic Acids Res. **38**, 4337–4348 (2010).
28. Ohnmacht, S. A. *et al*. A G-quadruplex-binding compound showing anti-tumour activity in an in vivo model for pancreatic cancer. Sci. Rep. **5**, 1–11 (2015).
29. Balasubramanian, S., Hurley, L. H. & Neidle, S. Targeting G-quadruplexes in gene promoters: A novel anticancer strategy? Nat. Rev. Drug Discov. **10**, 261–275 (2011).
30. Berman, E. *et al*. Resistance to imatinib in patients with chronic myelogenous leukemia and the splice variant BCR-ABL1 35INS. Leuk. Res. **49**, 108–112 (2016).
31. Hoogsteen, K. The crystal and molecular structure of a hydrogen-bonded complex between 1-methylthymine and 9-methyladenine. Acta Crystallogr. **16**, 907–916 (1963).
32. GELLERT, M., LIPSETT, M. N. & DAVIES, D. R. Helix formation by guanylic acid. Proc. Natl. Acad. Sci. U. S. A. **48**, 2013–2018 (1962).
33. Bhattacharyya, D., Arachchilage, G. M. & Basu, S. Metal cations in G-quadruplex folding and stability. Front. Chem. **4**, 1–14 (2016).
34. Phan, A. T., Kuryavyi, V., Luu, K. N. & Patel, D. J. Structure of two intramolecular G-quadruplexes formed by natural human telomere sequences in K+ solution. Nucleic Acids Res. **35**, 6517–6525 (2007).
35. Wang, Y. & Patel, D. J. Solution structure of the human telomeric repeat d[AG3(T2AG3)3] G-tetraplex. Structure **1**, 263–282 (1993).
36. Seenisamy, J. *et al*. The Dynamic Character of the G-Quadruplex Element in the c-MYC Promoter and Modification by TMPyP4. J. Am. Chem. Soc. **126**, 8702–8709 (2004).
37. Ducani, C., Bernardinelli, G., Högberg, B., Keppler, B. K. & Terenzi, A. Interplay of three G-quadruplex units in the KIT promoter. J. Am. Chem. Soc. **141**, 10205–10213 (2019).
38. Harrell Jr, W. A., Neidle, S. & Balasubramanian, S. Quadruplex Nucleic Acids. Quadruplex Nucleic Acids (The Royal Society of Chemistry, 2006). doi:10.1039/9781847555298.
39. Johnson, F. B. Fundamentals of G-quadruplex biology. Annual Reports in Medicinal Chemistry vol. 54 (Elsevier Inc., 2020).

40. Webba Da Silva, M. Geometric formalism for DNA quadruplex folding. Chem. - A Eur. J. **13**, 9738–9745 (2007).
41. Winnerdy, F. R. & Phan, A. T. Quadruplex structure and diversity. in Annual Reports in Medicinal Chemistry vol. 54 45–73 (2020).
42. Guédin, A., Gros, J., Alberti, P. & Mergny, J. L. How long is too long? Effects of loop size on G-quadruplex stability. Nucleic Acids Res. **38**, 7858–7868 (2010).
43. Lech, C. J., Heddi, B. & Phan, A. T. Guanine base stacking in G-quadruplex nucleic acids. Nucleic Acids Res. **41**, 2034–2046 (2013).
44. Hänsel-Hertsch, R., Di Antonio, M. & Balasubramanian, S. DNA G-quadruplexes in the human genome: Detection, functions and therapeutic potential. Nat. Rev. Mol. Cell Biol. **18**, 279–284 (2017).
45. Agrawal, P., Lin, C., Mathad, R. I., Carver, M. & Yang, D. The major G-quadruplex formed in the human BCL-2 proximal promoter adopts a parallel structure with a 13-nt loop in k+ solution. J. Am. Chem. Soc. **136**, 1750–1753 (2014).
46. Eddy, J. & Maizels, N. Gene function correlates with potential for G4 DNA formation in the human genome. Nucleic Acids Res. **34**, 3887–3896 (2006).
47. Hänsel-Hertsch, R. *et al*. G-quadruplex structures mark human regulatory chromatin. Nat. Genet. **48**, 1267–1272 (2016).
48. Biffi, G., Tannahill, D., McCafferty, J. & Balasubramanian, S. Quantitative visualization of DNA G-quadruplex structures in human cells. Nat. Chem. **5**, 182–186 (2013).
49. Rhodes, D. & Lipps, H. J. Survey and summary G-quadruplexes and their regulatory roles in biology. Nucleic Acids Res. **43**, 8627–8637 (2015).
50. Mendoza, O., Bourdoncle, A., Boulé, J. B., Brosh, R. M. & Mergny, J. L. G-quadruplexes and helicases. Nucleic Acids Research vol. 44 1989–2006 (2016).
51. de Renty, C. & Ellis, N. A. Bloom's syndrome: Why not premature aging?: A comparison of the BLM and WRN helicases. Ageing Res. Rev. **33**, 36–51 (2017).
52. Gray, L. T., Vallur, A. C., Eddy, J. & Maizels, N. G quadruplexes are genomewide targets of transcriptional helicases XPB and XPD. Nat. Chem. Biol. **10**, 313–318 (2014).
53. Henderson, E. R. & Blackburn, E. H. An overhanging 3' terminus is a conserved feature of telomeres. Mol. Cell. Biol. **9**, 345–348 (1989).
54. Griffith, J. D. *et al*. Mammalian telomeres end in a large duplex loop. Cell **97**, 503–514 (1999).
55. Maestroni, L., Matmati, S. & Coulon, S. Solving the telomere replication problem. Genes (Basel). **8**, 1–16 (2017).
56. Palm, W. & De Lange, T. How shelterin protects mammalian telomeres. Annu. Rev. Genet. **42**, 301–334 (2008).
57. Loayza, D. & De Lange, T. POT1 as a terminal transducer of TRF1 telomere length control. Nature **423**, 1013–1018 (2003).
58. Denchi, E. L. & De Lange, T. Protection of telomeres through independent control of ATM and ATR by TRF2 and POT1. Nature **448**, 1068–1071 (2007).
59. Zahler, A. M., Williamson, J. R., Cech, T. R. & Prescott, D. M. Inhibition of telomerase by G-quartet DNA structures. Nature **350**, 718–720 (1991).
60. Yang, D. & Okamoto, K. Structural insights into G-quadruplexes: Towards new anticancer drugs. Future Med. Chem. **2**, 619–646 (2010).

61. Siddiqui-Jain, A., Grand, C. L., Bearss, D. J. & Hurley, L. H. Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription. Proc. Natl. Acad. Sci. U. S. A. **99**, 11593–11598 (2002).

62. Ambrus, A., Chen, D., Dai, J., Jones, R. A. & Yang, D. Solution structure of the biologically relevant G-quadruplex element in the human c-MYC promoter. Implications for G-quadruplex stabilization. Biochemistry **44**, 2048–2058 (2005).

63. Phan, A. T., Kuryavyi, V., Burge, S., Neidle, S. & Patel, D. J. Structure of an unprecedented G-quadruplex scaffold in the human c-kit promoter. J. Am. Chem. Soc. **129**, 4386–4392 (2007).

64. Kuryavyi, V., Phan, A. T. & Patel, D. J. Solution structures of all parallel-stranded monomeric and dimeric G-quadruplex scaffolds of the human c-kit2 promoter. Nucleic Acids Res. **38**, 6757–6773 (2010).

65. Dai, J., Chen, D., Jones, R. A., Hurley, L. H. & Yang, D. NMR solution structure of the major G-quadruplex structure formed in the human BCL2 promoter region. Nucleic Acids Res. **34**, 5133–5144 (2006).

66. Sun, H. *et al.* A newly identified G-quadruplex as a potential target regulating Bcl-2 expression. Biochim. Biophys. Acta - Gen. Subj. **1840**, 3052–3057 (2014).

67. Membrino, A., Cogoi, S., Pedersen, E. B. & Xodo, L. E. G4-DNA formation in the HRAS promoter and rational design of decoy oligonucleotides for cancer therapy. PLoS One **6**, (2011).

68. Cogoi, S., Shchekotikhin, A. E. & Xodo, L. E. HRAS is silenced by two neighboring G-quadruplexes and activated by MAZ, a zinc-finger transcription factor with DNA unfolding property. Nucleic Acids Res. **42**, 8379–8388 (2014).

69. Luu, K. N., Phan, A. T., Kuryavyi, V., Lacroix, L. & Patel, D. J. Structure of the human telomere in K+ solution: An intramolecular (3 + 1) G-quadruplex scaffold. J. Am. Chem. Soc. **128**, 9963–9970 (2006).

70. Eilers, M. & Eisenman, R. N. Myc's broad reach. Genes and Development vol. 22 2755–2766 (2008).

71. Salghetti, S. E., Young Kim, S. & Tansey, W. P. Destruction of Myc by ubiquitin-mediated proteolysis: cancer-associated and transforming mutations stabilize Myc. EMBO J. **18**, 717–726 (1999).

72. Menssen, A. & Hermeking, H. Characterization of the c-MYC-regulated transcriptome by SAGE: Identification and analysis of c-MYC target genes. Proc. Natl. Acad. Sci. U. S. A. **99**, 6274–6279 (2002).

73. Miller, D. M., Thomas, S. D., Islam, A., Muench, D. & Sedoris, K. c-Myc and cancer metabolism. Clin. Cancer Res. **18**, 5546–5553 (2012).

74. Greenberg, R. A. *et al.* Telomerase reverse transcriptase gene is a direct target of c-Myc but is not functionally equivalent in cellular transformation. Oncogene **18**, 1219–1226 (1999).

75. Simonsson, T., Pecinka, P. & Kubista, M. DNA tetraplex formation in the control region of c-myc. Nucleic Acids Res. **26**, 1167–1172 (1998).

76. Grand, C. L. *et al.* The cationic porphyrin TMPyP4 down-regulates c-MYC and human telomerase reverse transcriptase expression and inhibits tumor growth in vivo. Mol. Cancer Ther. **1**, 565–573 (2002).

77. Kuryavyi, V., Gaw, H. Y. & Patel, D. J. Small-molecule interaction with a five-guanine-tract G-quadruplex structure from the human MYC promoter. Nat. Chem. Biol. **1**, 167–174 (2005).

78. Hurley, L. H., Von Hoff, D. D., Siddiqui-Jain, A. & Yang, D. Drug Targeting of the c-MYC Promoter to Repress Gene Expression via a G-Quadruplex Silencer Element. Semin. Oncol. **33**, 498–512 (2006).

79. Han, F. X., Wheelhouse, R. T. & Hurley, L. H. Interactions of TMPyP4 and TMPyP2 with quadruplex DNA. Structural basis for the differential effects on telomerase inhibition. J. Am. Chem. Soc. **121**, 3561–3570 (1999).

80. Kim, M. Y., Gleason-Guzman, M., Izbicka, E., Nishioka, D. & Hurley, L. H. The different biological effects of Telomestatin and TMPyP4 can be attributed to their selectivity for interaction with intramolecular or intermolecular G-quadruplex structures. Cancer Res. **63**, 3247–3256 (2003).

81. Ren, J. & Chaires, J. B. Sequence and structural selectivity of nucleic acid binding ligands. Biochemistry **38**, 16067–16075 (1999).

82. Marchand, A., Rosu, F., Zenobi, R. & Gabelica, V. Thermal Denaturation of DNA G-Quadruplexes and Their Complexes with Ligands: Thermodynamic Analysis of the Multiple States Revealed by Mass Spectrometry. J. Am. Chem. Soc. **140**, 12553–12565 (2018).

83. Le, D. D., Di Antonio, M., Chan, L. K. M. & Balasubramanian, S. G-quadruplex ligands exhibit differential G-tetrad selectivity. Chem. Commun. **51**, 8048–8050 (2015).

84. Ruan, T. L. *et al*. Lowering the overall charge on TMPyP4 improves its selectivity for G-quadruplex DNA. Biochimie **132**, 121–130 (2017).

85. Oyekan, A. O. Role of the Endothelium and Cyclic GMP in Renal Vasodilator Responses to Cryptolepine in Rats. J. Cardiovasc. Pharmacol. **23**, (1994).

86. Bierer, D. E. Ethnobotanical-directed discovery of the antihyperglycemic properties of cryptolepine: Its isolation from Cryptolepis sanguinolenta, synthesis, and in vitro and in vivo activities. J. Med. Chem. **41**, 894–901 (1998).

87. Ou, T. M. *et al*. Stabilization of G-quadruplex DNA and down-regulation of oncogene c-myc by quindoline derivatives. J. Med. Chem. **50**, 1465–1474 (2007).

88. Zhou, J. L. *et al*. Synthesis and evaluation of quindoline derivatives as G-quadruplex inducing and stabilizing ligands and potential inhibitors of telomerase. J. Med. Chem. **48**, 7315–7321 (2005).

89. Dai, J., Carver, M., Hurley, L. H. & Yang, D. Solution structure of a 2:1 quindoline-c-MYC G-quadruplex: Insights into G-quadruplex-interactive small molecule drug design. J. Am. Chem. Soc. **133**, 17673–17680 (2011).

90. Funke, A., Dickerhoff, J. & Weisz, K. Towards the Development of Structure-Selective G-Quadruplex-Binding Indolo[3,2-b]quinolines. Chem. - A Eur. J. **22**, 3170–3181 (2016).

91. Boddupally, P. V. L. *et al*. Anticancer activity and cellular repression of c-MYC by the G-quadruplex-stabilizing 11-piperazinylquindoline is not dependent on direct targeting of the G-quadruplex in the c-MYC promoter. J. Med. Chem. **55**, 6076–6086 (2012).

92. Hu, M. H. *et al*. Discovery of a New Four-Leaf Clover-Like Ligand as a Potent c-MYC Transcription Inhibitor Specifically Targeting the Promoter G-Quadruplex. J. Med. Chem. **61**, 2447–2459 (2018).

93. Zeng, D. Y. *et al*. Discovery of Novel 11-Triazole Substituted Benzofuro[3,2-b]quinolone Derivatives as c-myc G-Quadruplex Specific Stabilizers via Click Chemistry. J. Med. Chem. **60**, 5407–5423 (2017).

94. Panda, D. *et al*. A Nucleus-Imaging Probe That Selectively Stabilizes a Minor Conformation of c-MYC G-quadruplex and Down-regulates c-MYC Transcription in Human Cancer Cells. Sci. Rep. **5**, 1–16 (2015).

95. Lee, H. M. *et al*. Identification of natural product Fonsecin B as a stabilizing ligand of c-myc G-quadruplex DNA by high-throughput virtual screening. Chem. Commun. **46**, 4680–4682 (2010).

96. Ma, Y. *et al*. 9-N-Substituted berberine derivatives: Stabilization of G-quadruplex DNA and down-regulation of oncogene c-myc. Bioorganic Med. Chem. **16**, 7582–7591 (2008).

97. Yang, H. *et al*. A Cyclometallated Iridium(III) Complex As a c-myc G-Quadruplex Stabilizer and Down-Regulator of c-myc Oncogene Expression. Curr. Med. Chem. **20**, 576–582 (2013).

98. Wu, P. *et al*. Stabilization of G-quadruplex DNA with platinum(II) Schiff base complexes: Luminescent probe and down-regulation of c-myc oncogene expression. Chem. - A Eur. J. **15**, 13008–13021 (2009).

99. Lennartsson, J. & Rönnstrand, L. Stem cell factor receptor/c-Kit: From basic Science to clinical implications. Physiol. Rev. **92**, 1619–1649 (2012).

100. Park, G. H., Plummer, H. K. & Krystal, G. W. Selective Sp1 binding is critical for maximal activity of the human c- kit promoter. Blood **92**, 4138–4149 (1998).

101. Rankin, S. *et al*. Putative DNA quadruplex formation within the human c-kit oncogene. J. Am. Chem. Soc. **127**, 10584–10589 (2005).

102. Fernando, H. *et al*. A conserved quadruplex motif located in a transcription activation site of the human c-kit oncogene. Biochemistry **45**, 7854–7860 (2006).

103. Wei, D., Husby, J. & Neidle, S. Flexibility and structural conservation in a c-KIT G-quadruplex. Nucleic Acids Res. **43**, 629–644 (2015).

104. Rocca, R. *et al*. In Silico Identification of Piperidinyl-amine Derivatives as Novel Dual Binders of Oncogene c-myc/c-Kit G-quadruplexes. ACS Med. Chem. Lett. **9**, 848–853 (2018).

105. Zorzan, E. *et al*. Screening of candidate G-quadruplex ligands for the human c-KIT promotorial region and their effects in multiple in-vitro models. Oncotarget **7**, 21658–21675 (2016).

106. Bejugam, M. *et al*. Trisubstituted isoalloxazines as a new class of G-quadruplex binding ligands: Small molecule regulation of c-kit oncogene expression. J. Am. Chem. Soc. **129**, 12926–12927 (2007).

107. Zuffo, M. *et al*. More is not always better: Finding the right trade-off between affinity and selectivity of a G-quadruplex ligand. Nucleic Acids Res. **46**, (2018).

108. McLuckie, K. I. E. *et al*. G-Quadruplex-Binding benzo[a]phenoxazines down-regulate c-KIT expression in human gastric carcinoma cells. J. Am. Chem. Soc. **133**, 2658–2663 (2011).

109. Waller, Z. A. E., Shirude, P. S., Rodriguez, R. & Balasubramanian, S. Triarylpyridines: A versatile small molecule scaffold for G-quadruplex recognition. Chem. Commun. 1467–1469 (2008) doi:10.1039/b718854d.

110. Diveshkumar, K. V. *et al*. Specific Stabilization of c-MYC and c-KIT G-Quadruplex DNA Structures by Indolylmethyleneindanone Scaffolds. Biochemistry **55**, 3571–3585 (2016).
111. Jantos, K., Rodriguez, R., Ladame, S., Shirude, P. S. & Balasubramanian, S. Oxazole-based peptide macrocycles: A new class of G-quadruplex binding ligands. J. Am. Chem. Soc. **128**, 13662–13663 (2006).
112. Fernández-Medarde, A. & Santos, E. Ras in cancer and developmental diseases. Genes and Cancer **2**, 344–358 (2011).
113. Miller, M. S. & Miller, L. D. RAS mutations and oncogenesis: Not all RAS mutations are created equally. Front. Genet. **2**, 1–9 (2012).
114. Forbes, S. A. *et al*. COSMIC: Somatic cancer genetics at high-resolution. Nucleic Acids Res. **45**, D777–D783 (2017).
115. Saha, A., Bombard, S., Granzhan, A. & Teulade-Fichou, M. P. Probing of G-Quadruplex Structures via Ligand-Sensitized Photochemical Reactions in BrU-Substituted DNA. Sci. Rep. **8**, 1–14 (2018).
116. Membrino, A. *et al*. Cellular uptake and binding of guanidine-modified phthalocyanines to KRAS/HRAS G-quadruplexes. Chem. Commun. **46**, 625–627 (2010).
117. Gonçalves, D. P. N., Rodriguez, R., Balasubramanian, S. & Sanders, J. K. M. Tetramethylpyridiniumporphyrazines—a new class of G-quadruplex inducing and stabilising ligands. Chem. Commun. 4685–4687 (2006) doi:10.1039/b611731g.
118. Liu, Z. Q. *et al*. Facile syntheses of disubstituted bis(vinylquinolinium)benzene derivatives as G-quadruplex DNA binders. Tetrahedron **69**, 4922–4932 (2013).
119. Wang, M. Q. *et al*. Development of a light-up fluorescent probe for HRAS G-quadruplex DNA. Tetrahedron Lett. **58**, 3296–3300 (2017).
120. Lu, Y. J. *et al*. Benzothiazole-substituted benzofuroquinolinium dyes as new fluorescent probes for G-quadruplex DNA. Dye. Pigment. **122**, 94–102 (2015).
121. Li, Q. *et al*. Stabilizing parallel G-quadruplex DNA by a new class of ligands: Two non-planar alkaloids through interaction in lateral grooves. Biochimie **91**, 811–819 (2009).
122. Wang, M. Q. *et al*. Development of a carbazole-based fluorescence probe for G-quadruplex DNA: The importance of side-group effect on binding specificity. Spectrochim. Acta - Part A Mol. Biomol. Spectrosc. **199**, 441–447 (2018).
123. Onel, B. *et al*. A New G-Quadruplex with Hairpin Loop Immediately Upstream of the Human BCL2 P1 Promoter Modulates Transcription. J. Am. Chem. Soc. **138**, 2563–2570 (2016).
124. Ou, T. M. *et al*. Inhibition of cell proliferation by quindoline derivative (SYUIQ-05) through its preferential interaction with c-myc promoter G-quadruplex. J. Med. Chem. **54**, 5671–5679 (2011).
125. Lu, Y. J. *et al*. 5-N-methylated quindoline derivatives as telomeric G-quadruplex stabilizing ligands: Effects of 5-N positive charge on quadruplex binding affinity and cell proliferation. J. Med. Chem. **51**, 6381–6392 (2008).
126. Wang, X. D. *et al*. Turning off transcription of the bcl-2 gene by stabilizing the bcl-2 promoter quadruplex with quindoline derivatives. J. Med. Chem. **53**, 4390–4398 (2010).
127. Pelliccia, S. *et al*. Bio-inspired dual-selective BCL-2/c-MYC G-quadruplex binders: Design, synthesis, and anticancer activity of drug-like imidazo[2,1-i]purine derivatives. J. Med. Chem. **63**, 2035–2050 (2020).

128. Han, H., Langley, D. R., Rangan, A. & Hurley, L. H. Selective interactions of cationic porphyrins with G-quadruplex structures. J. Am. Chem. Soc. **123**, 8902–8913 (2001).

129. Fedoroff, O. Y. *et al*. NMR-based model of a telomerase-inhibiting compound bound to G- quadruplex DNA. Biochemistry **37**, 12367–12374 (1998).

130. Read, M. *et al*. Structure-based design of selective and potent G quadruplex-mediated telomerase inhibitors. Proc. Natl. Acad. Sci. U. S. A. **98**, 4844–4849 (2001).

131. Burger, A. M. *et al*. The G-quadruplex-interactive molecule BRACO-19 inhibits tumor growth, consistent with telomere targeting and interference with telomerase function. Cancer Res. **65**, 1489–1496 (2005).

132. Incles, C. M. *et al*. A G-quadruplex telomere targeting agent produces p16-associated senescence and chromosomal fusions in human prostate cancer cells. Mol. Cancer Ther. **3**, 1201–1206 (2004).

133. Müller, S. *et al*. Pyridostatin analogues promote telomere dysfunction and long-term growth inhibition in human cancer cells. Org. Biomol. Chem. **10**, 6537–6546 (2012).

134. De Cian, A. *et al*. Reevaluation of telomerase inhibition by quadruplex ligands and their mechanisms of action. Proc. Natl. Acad. Sci. U. S. A. **104**, 17347–17352 (2007).

135. De Cian, A., DeLemos, E., Mergny, J. L., Teulade-Fichou, M. P. & Monchaud, D. Highly efficient G-quadruplex recognition by bisquinolinium compounds. J. Am. Chem. Soc. **129**, 1856–1857 (2007).

136. Marchand, A. *et al*. Ligand-induced conformational changes with cation ejection upon binding to human telomeric DNA G-quadruplexes. J. Am. Chem. Soc. **137**, 750–756 (2015).

137. Bončina, M. *et al*. Thermodynamic fingerprints of ligand binding to human telomeric G-quadruplexes. Nucleic Acids Res. **43**, 10376–10386 (2015).

138. Chung, W. J., Heddi, B., Hamon, F., Teulade-Fichou, M.-P. & Phan, A. T. Solution Structure of a G-quadruplex Bound to the Bisquinolinium Compound Phen-DC 3. Angew. Chemie **126**, 1017–1020 (2014).

139. Hampel, S. M., Sidibe, A., Gunaratnam, M., Riou, J. F. & Neidle, S. Tetrasubstituted naphthalene diimide ligands with selectivity for telomeric G-quadruplexes and cancer cells. Bioorganic Med. Chem. Lett. **20**, 6459–6463 (2010).

140. Vo, T. *et al*. Substituted Naphthalenediimide Compounds Bind Selectively to Two Human Quadruplex Structures with Parallel Topology. ACS Med. Chem. Lett. (2020) doi:10.1021/acsmedchemlett.0c00041.

141. Gunaratnam, M. *et al*. Targeting pancreatic cancer with a G-quadruplex ligand. Bioorganic Med. Chem. **19**, 7151–7157 (2011).

142. Micco, M. *et al*. Structure-based design and evaluation of naphthalene diimide G-quadruplex ligands as telomere targeting agents in pancreatic cancer cells. J. Med. Chem. **56**, 2959–2974 (2013).

143. Cuenca, F. *et al*. Tri- and tetra-substituted naphthalene diimides as potent G-quadruplex ligands. Bioorganic Med. Chem. Lett. **18**, 1668–1673 (2008).

144. Shin-ya, K. *et al*. Telomestatin, a novel telomerase inhibitor from Streptomyces anulatus. J. Am. Chem. Soc. **123**, 1262–1263 (2001).

145. Chung, W. J. *et al*. Solution structure of an intramolecular (3 + 1) human telomeric G-quadruplex bound to a telomestatin derivative. J. Am. Chem. Soc. **135**, 13495–13501 (2013).

146. Moazed, D., and H. F. N. Interaction of antibiotics with functional sites in 16S ribosomal RNA. Nature **327**, 389–394 (1987).
147. Fourmy, D., Recht, M. I., Blanchard, S. C. & Puglisi, J. D. Structure of the A site of Escherichia coli 16S ribosomal RNA complexed with an aminoglycoside antibiotic. Science (80-.). **274**, 1367–1371 (1996).
148. Oliveira, J. F. P. *et al*. Prevalence and risk factors for aminoglycoside nephrotoxicity in intensive care units. Antimicrob. Agents Chemother. **53**, 2887–2891 (2009).
149. Grau-Campistany, A. *et al*. Conjugation of a Ru(II) arene complex to neomycin or to guanidinoneomycin leads to compounds with differential cytotoxicities and accumulation between cancer and normal cells. Mol. Pharm. **10**, 1964–1976 (2013).
150. Arya, D. P., Xue, L. & Willis, B. Aminoglycoside (Neomycin) Preference Is for A-Form Nucleic Acids, Not Just RNA: Results from a Competition Dialysis Study. J. Am. Chem. Soc. **125**, 10148–10149 (2003).
151. Arya, D. P. *et al*. Neomycin Binding to Watson−Hoogsteen (W−H) DNA Triplex Groove: A Model. J. Am. Chem. Soc. **125**, 3733–3744 (2003).
152. Kirk, S. R., Luedtke, N. W. & Tor, Y. Neomycin-acridine conjugate: A potent inhibitor of Rev-RRE binding [16]. J. Am. Chem. Soc. **122**, 980–981 (2000).
153. Arya, D. P., Xue, L. & Tennant, P. Combining the best in triplex recognition: Synthesis and nucleic acid binding of a BQQ-neomycin conjugate. J. Am. Chem. Soc. **125**, 8070–8071 (2003).
154. Xue, L., Ranjan, N. & Arya, D. P. Synthesis and spectroscopic studies of the aminoglycoside (Neomycin)-perylene conjugate binding to human telomeric DNA. Biochemistry **50**, 2838–2849 (2011).
155. Ramirez, M. S. & Tolmasky, M. E. Aminoglycoside modifying enzymes. Drug Resist. Updat. **13**, 151–171 (2010).
156. Zhang, J. *et al*. Synthesis and combinational antibacterial study of 5″-modified neomycin. J. Antibiot. (Tokyo). **62**, 539–544 (2009).
157. Liu, M. *et al*. Tethered bisubstrate derivatives as probes for mechanism and as inhibitors of aminoglycoside 3'-phosphotransferases. J. Org. Chem. **65**, 7422–7431 (2000).
158. Luedtke, N. W., Liu, Q. & Tor, Y. RNA - Ligand interactions: Affinity and specificity of aminoglycoside dimers and acridine conjugates to the HIV-1 Rev response element. Biochemistry **42**, 11391–11403 (2003).
159. Lu, W., Sengupta, S., Petersen, J. L., Akhmedov, N. G. & Shi, X. Mitsunobu coupling of nucleobases and alcohols: An efficient, practical synthesis for novel nonsugar carbon nucleosides. J. Org. Chem. **72**, 5012–5015 (2007).
160. Edward Lindsell, W., Murray, C., Preston, P. N. & Woodman, T. A. J. Synthesis of 1,3-diynes in the purine, pyrimidine, 1,3,5-triazine and acridine series. Tetrahedron **56**, 1233–1245 (2000).
161. Liu, W., Minier, M. A., Franz, A. H., Curtis, M. & Xue, L. Synthesis of nucleobase-calix[4]arenes via click chemistry and evaluation of their complexation with alkali metal ions and molecular assembly. Supramol. Chem. **23**, 806–818 (2011).
162. Mergny, J. L. & Lacroix, L. Analysis of Thermal Melting Curves. Oligonucleotides **13**, 515–537 (2003).
163. Degtyareva, N. N. *et al*. Antimicrobial Activity, AME Resistance, and A-Site Binding Studies of Anthraquinone-Neomycin Conjugates. ACS Infect. Dis. **3**, 206–215 (2017).

164. Xue, L., Charles, I. & Arya, D. P. Pyrene-neomycin conjugate: dual recognition of a DNA triple helix. Chem. Commun. 70–71 (2002) doi:10.1039/B108171C.

165. Ambrus, A. *et al*. Human telomeric sequence forms a hybrid-type intramolecular G-quadruplex structure with mixed parallel/antiparallel strands in potassium solution. Nucleic Acids Res. **34**, 2723–2735 (2006).

166. Ranjan, N., Andreasen, K. F., Kumar, S., Hyde-Volpe, D. & Arya, D. P. Aminoglycoside Binding to Oxytricha nova Telomeric DNA. Biochemistry **49**, 9891–9903 (2010).

167. Balouiri, M., Sadiki, M. & Ibnsouda, S. K. Methods for in vitro evaluating antimicrobial activity: A review. J. Pharm. Anal. **6**, 71–79 (2016).

168. Jubeh, B., Breijyeh, Z. & Karaman, R. Resistance of gram-positive bacteria to current antibacterial agents and overcoming approaches. Molecules **25**, (2020).

169. Miller, S. I. Antibiotic resistance and regulation of the Gram-negative bacterial outer membrane barrier by host innate immune molecules. MBio **7**, 5–7 (2016).

170. Zhao, F. *et al*. Molecular recognition of RNA by neomycin and a restricted neomycin derivative. Angew. Chemie - Int. Ed. **44**, 5329–5334 (2005).

171. Becker, B. & Cooper, M. A. Aminoglycoside antibiotics in the 21st century. ACS Chem. Biol. **8**, 105–115 (2013).

172. Chittapragada, M., Roberts, S. & Ham, Y. W. Aminoglycosides: Molecular insights on the recognition of RNA and aminoglycoside mimics. Perspect. Medicin. Chem. **2009**, 21–37 (2009).

173. Lightfoot, H. L., Hagen, T., Cléry, A., Allain, F. H. T. & Hall, J. Control of the polyamine biosynthesis pathway by G2-quadruplexes. Elife **7**, 1–22 (2018).

174. Rosu, F., Gabelica, V., Shin-Ya, K. & De Pauw, E. Telomestatin-induced stabilization of the human telomeric DNA quadruplex monitored by electrospray mass spectrometry. Chem. Commun. **2**, 2702–2703 (2003).

175. Martino, L., Pagano, B., Fotticchia, I., Neidle, S. & Giancola, C. Shedding light on the interaction between TMPyP4 and human telomeric quadruplexes. J. Phys. Chem. B **113**, 14779–14786 (2009).

176. Wang, S. *et al*. Thiazole orange – Spermine conjugate: A potent human telomerase inhibitor comparable to BRACO-19. Eur. J. Med. Chem. **175**, 20–33 (2019).

177. Kaul, M. & Pilch, D. S. Thermodynamics of Aminoglycoside−rRNA Recognition: The Binding of Neomycin-Class Aminoglycosides to the A Site of 16S rRNA. Biochemistry **41**, 7695–7706 (2002).

178. Arya, D. P. & Coffee Jr, R. L. DNA Triple Helix Stabilization by Aminoglycoside Antibiotics. Bioorg. Med. Chem. Lett. **10**, 1897–1899 (2000).

179. Liu, W., Wang, S., Dotsenko, I. A., Samoshin, V. V & Xue, L. Arylsulfanyl groups - Suitable side chains for 5-substituted 1,10-phenanthroline and nickel complexes as G4 ligands and telomerase inhibitors. J. Inorg. Biochem. **173**, 12–20 (2017).

180. Reed, J. E., Neidle, S. & Vilar, R. Stabilisation of human telomeric quadruplex DNA and inhibition of telomerase by a platinum-phenanthroline complex. Chem. Commun. 4366–4368 (2007) doi:10.1039/b709898g.

181. Dotsenko, I. A., Curtis, M., Samoshina, N. M. & Samoshin, V. V. Convenient synthesis of 5-aryl(alkyl)sulfanyl-1,10-phenanthrolines from 5,6-epoxy-5,6-dihydro-1,10-phenanthroline, and their activity towards fungal β-d-glycosidases. Tetrahedron **67**, 7470–7478 (2011).

182. Monchaud, D., Allain, C. & Teulade-Fichou, M. P. Development of a fluorescent

intercalator displacement assay (G4-FID) for establishing quadruplex-DNA affinity and selectivity of putative ligands. Bioorganic Med. Chem. Lett. **16**, 4842–4845 (2006).

183. Boger, D. L., Fink, B. E., Brunette, S. R., Tse, W. C. & Hedrick, M. P. A simple, high-resolution method for establishing DNA binding affinity and sequence selectivity. J. Am. Chem. Soc. **123**, 5878–5891 (2001).

184. Lubitz, I., Zikich, D. & Kotlyar, A. Specific high-affinity binding of thiazole orange to triplex and g-quadruplex DNA. Biochemistry **49**, 3567–3574 (2010).

185. Jezowska-Bojczuk, M. *et al*. Identification of copper(II) binding sites in the aminoglycosidic antibiotic neomycin B. Eur. J. Inorg. Chem. 3063–3071 (2005) doi:10.1002/ejic.200500102.

186. del Villar-Guerra, R., Trent, J. O. & Chaires, J. B. G-Quadruplex Secondary Structure Obtained from Circular Dichroism Spectroscopy. Angew. Chemie - Int. Ed. **57**, 7171–7175 (2018).

187. Chung, W. J. *et al*. Structure of a left-handed DNA G-quadruplex. Proc. Natl. Acad. Sci. U. S. A. **112**, 2729–2733 (2015).

188. Dai, J. *et al*. Structure of the intramolecular human telomeric G-quadruplex in potassium solution: A novel adenine triple formation. Nucleic Acids Res. **35**, 2440–2450 (2007).

189. Carvalho, J., Queiroz, J. A. & Cruz, C. Circular dichroism of G-Quadruplex: A laboratory experiment for the study of topology and ligand binding. J. Chem. Educ. **94**, 1547–1551 (2017).

190. Gray, R. D., Trent, J. O. & Chaires, J. B. Folding and unfolding pathways of the human telomeric G-quadruplex. J. Mol. Biol. **426**, 1629–1650 (2014).

191. Rueda, M., Luque, F. J. & Orozco, M. G-quadruplexes can maintain their structure in the gas phase. J. Am. Chem. Soc. **128**, 3608–3619 (2006).

192. Guo, X., Liu, S. & Yu, Z. Bimolecular Quadruplexes and Their Transitions to Higher-Order Molecular Structures Detected by ESI-FTICR-MS. J. Am. Soc. Mass Spectrom. **18**, 1467–1476 (2007).

193. Rosu, F., Gabelica, V., Houssier, C., Colson, P. & De Pauw, E. Triplex and quadruplex DNA structures studied by electrospray mass spectrometry. Rapid Commun. Mass Spectrom. **16**, 1729–1736 (2002).

194. Su, H. & Xu, Y. Application of ITC-based characterization of thermodynamic and kinetic association of ligands with proteins in drug design. Front. Pharmacol. **9**, 1–7 (2018).

195. White, E. W. *et al*. Structure-specific recognition of quadruplex DNA by organic cations: Influence of shape, substituents and charge. Biophys. Chem. **126**, 140–153 (2007).

196. Seenisamy, J. *et al*. Design and Synthesis of an Expanded Porphyrin That Has Selectivity for the c-MYC G-Quadruplex Structure. J. Am. Chem. Soc. **127**, 2944–2959 (2005).

197. Wei, D., Parkinson, G. N., Reszka, A. P. & Neidle, S. Crystal structure of a c-kit promoter quadruplex reveals the structural role of metal ions and water molecules in maintaining loop conformation. Nucleic Acids Res. **40**, 4691–4700 (2012).

198. Kim, N. W. *et al*. Specific Association of Human Telomerase Activity with Immortal Cells and Cancer. Science (80-.). **266**, 2011–2015 (1994).

199. Pascolo, E. *et al*. Mechanism of human telomerase inhibition by BIBR1532, a synthetic, non-nucleosidic drug candidate. J. Biol. Chem. **277**, 15566–15572 (2002).

200. Dai, J. *et al*. An intramolecular G-quadruplex structure with mixed parallel/antiparallel G-strands formed in the human BCL-2 promoter region in solution. J. Am. Chem. Soc. **128**, 1096–1098 (2006).

201. Dexheimer, T. S., Sun, D. & Hurley, L. H. Deconvoluting the structural and drug-recognition complexity of the G-quadruplex-forming region upstream of the bcl-2 P1 promoter. J. Am. Chem. Soc. **128**, 5404–5415 (2006).

202. Mukundan, V. T. & Phan, A. T. Bulges in G-quadruplexes: Broadening the definition of G-quadruplex-forming sequences. J. Am. Chem. Soc. **135**, 5017–5028 (2013).

203. Xu, H. *et al*. CX-5461 is a DNA G-quadruplex stabilizer with selective lethality in BRCA1/2 deficient tumours. Nat. Commun. **8**, (2017).

204. Drygin, D. *et al*. Anticancer activity of CX-3543: A direct inhibitor of rRNA biogenesis. Cancer Res. **69**, 7653–7661 (2009).

205. Rodriguez, R. *et al*. Small-molecule-induced DNA damage identifies alternative DNA structures in human genes. Nat. Chem. Biol. **8**, 301–310 (2012).

206. Pitié, M., Donnadieu, B. & Meunier, B. Preparation of the New Bis(phenanthroline) Ligand 'Clip-Phen' and Evaluation of the Nuclease Activity of the Corresponding Copper Complex. Inorg. Chem. **37**, 3486–3489 (1998).

207. Pitié, M. *et al*. DNA cleavage by copper complexes of 2- and 3-Clip-Phen derivatives. Eur. J. Inorg. Chem. 528–540 (2003) doi:10.1002/ejic.200390075.

208. Pitié, M. & Meunier, B. Preparation of a spermine conjugate of the bis-phenanthroline ligand Clip-Phen and evaluation of the corresponding copper complex. Bioconjug. Chem. **9**, 604–611 (1998).

209. Pitie, M., Van Horn, J. D., Brion, D., Burrows, C. J. & Meunier, B. Targeting the DNA cleavage activity of copper phenanthroline and Clip-Phen to A·T tracts via linkage to a poly-N-methylpyrrole. Bioconjug. Chem. **11**, 892–900 (2000).

210. Boldron, C., Ross, S. A., Pitié, M. & Meunier, B. Acridine conjugates of 3-Clip-Phen: Influence of the linker on the synthesis and the DNA cleavage activity of their copper complexes. Bioconjug. Chem. **13**, 1013–1020 (2002).

211. De Hoog, P. *et al*. New approach for the preparation of efficient DNA cleaving agents: Ditopic copper-platinum complexes based on 3-clip-phen and cisplatin. J. Med. Chem. **50**, 3148–3152 (2007).

212. Özalp-Yaman, Ş. *et al*. Platinated copper(3-clip-phen) complexes as effective DNA-cleaving and cytotoxic agents. Chem. - A Eur. J. **14**, 3418–3426 (2008).

213. Panattoni, A., El-Sagheer, A. H., Brown, T., Kellett, A. & Hocek, M. Oxidative DNA Cleavage with Clip-Phenanthroline Triplex-Forming Oligonucleotide Hybrids. ChemBioChem **21**, 991–1000 (2020).

214. Bianco, S. *et al*. Bis-phenanthroline derivatives as suitable scaffolds for effective G-quadruplex recognition. Dalt. Trans. **39**, 5833–5841 (2010).

215. Wang, J., Liang, Y. L. & Qu, J. Boiling water-catalyzed neutral and selective N-Boc deprotection. Chem. Commun. 5144–5146 (2009) doi:10.1039/b910239f.

216. Seenisamy, J. *et al*. The dynamic character of the G-quadruplex element in the c-MYC promoter and modification by TMPyP4. J. Am. Chem. Soc. **126**, 8702–8709 (2004).

217. Chung, W. J. *et al*. Structure of a left-handed DNA G-quadruplex. Proc. Natl. Acad. Sci. **112**, 2729–2733 (2015).

218. Khalili, F., Henni, A. & East, A. L. L. PKa values of some Piperazines at (298, 303, 313, and 323) K. J. Chem. Eng. Data **54**, 2914–2917 (2009).

219. Perrin, D. D. Dissociation constants of organic bases in aqueous solution. (Butterworths, 1972).

220. Monchaud, D. *et al*. Ligands playing musical chairs with G-quadruplex DNA: A rapid and

simple displacement assay for identifying selective G-quadruplex binders. Biochimie **90**, 1207–1223 (2008).

221. Musetti, C. *et al*. Metal ion-mediated assembly of effective phenanthroline-based G-quadruplex ligands. Dalt. Trans. 3657–3660 (2009) doi:10.1039/b904630p.

222. Rezler, E. M. *et al*. Telomestatin and diseleno sapphyrin bind selectively to two different forms of the human telomeric G-quadruplex structure. J. Am. Chem. Soc. **127**, 9439–9447 (2005).

223. Zhang, Z., Dai, J., Veliath, E., Jones, R. A. & Yang, D. Structure of a two-G-tetrad intramolecular G-quadruplex formed by variant human telomeric sequence in K+ solution: Insights into the interconversion of human telomeric G-quadruplex structures. Nucleic Acids Res. **38**, 1009–1021 (2009).

224. Lim, K. W. *et al*. Structure of the human telomere in K + solution: A stable basket-type G-quadruplex with only two G-tetrad layers. J. Am. Chem. Soc. **131**, 4301–4309 (2009).

225. Gray, R. D. & Chaires, J. B. Linkage of cation binding and folding in human telomeric quadruplex DNA. Biophys. Chem. **159**, 205–209 (2011).

226. Głuszyńska, A., Juskowiak, B. & Rubiś, B. Binding study of the fluorescent carbazole derivative with human telomeric G-quadruplexes. Molecules **23**, 1–22 (2018).

227. Singh, M. *et al*. Use of neomycin as a structured amino-containing side chain motif for phenanthroline-based G-quadruplex ligands and telomerase inhibitors. Chem. Biol. Drug Des. **96**, 1292–1304 (2020).

APPENDIX A: ^{1}H AND ^{13}C NMR SPECTRA OF SELECTED COMPOUNDS
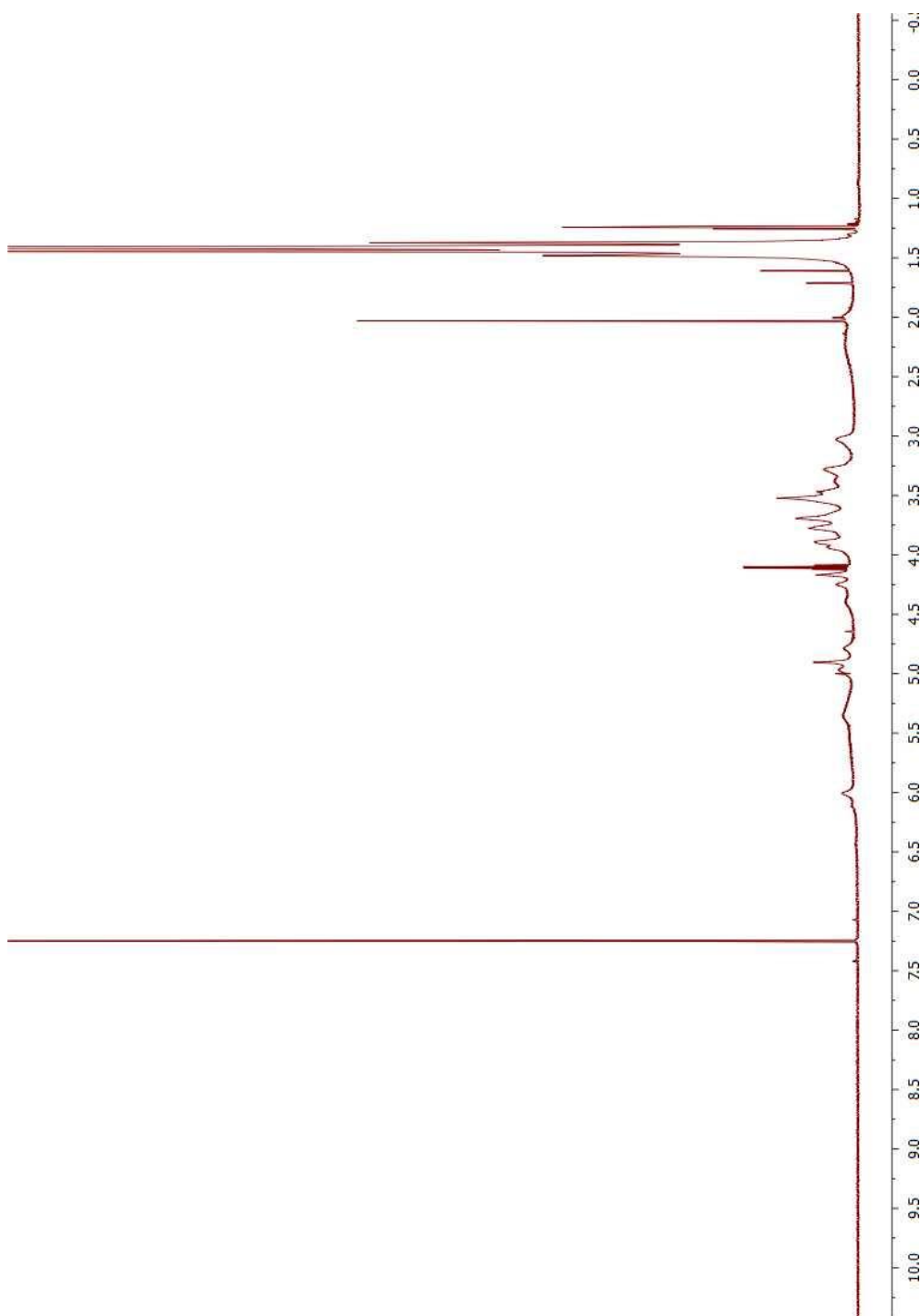
^{1}H NMR of compound **2**
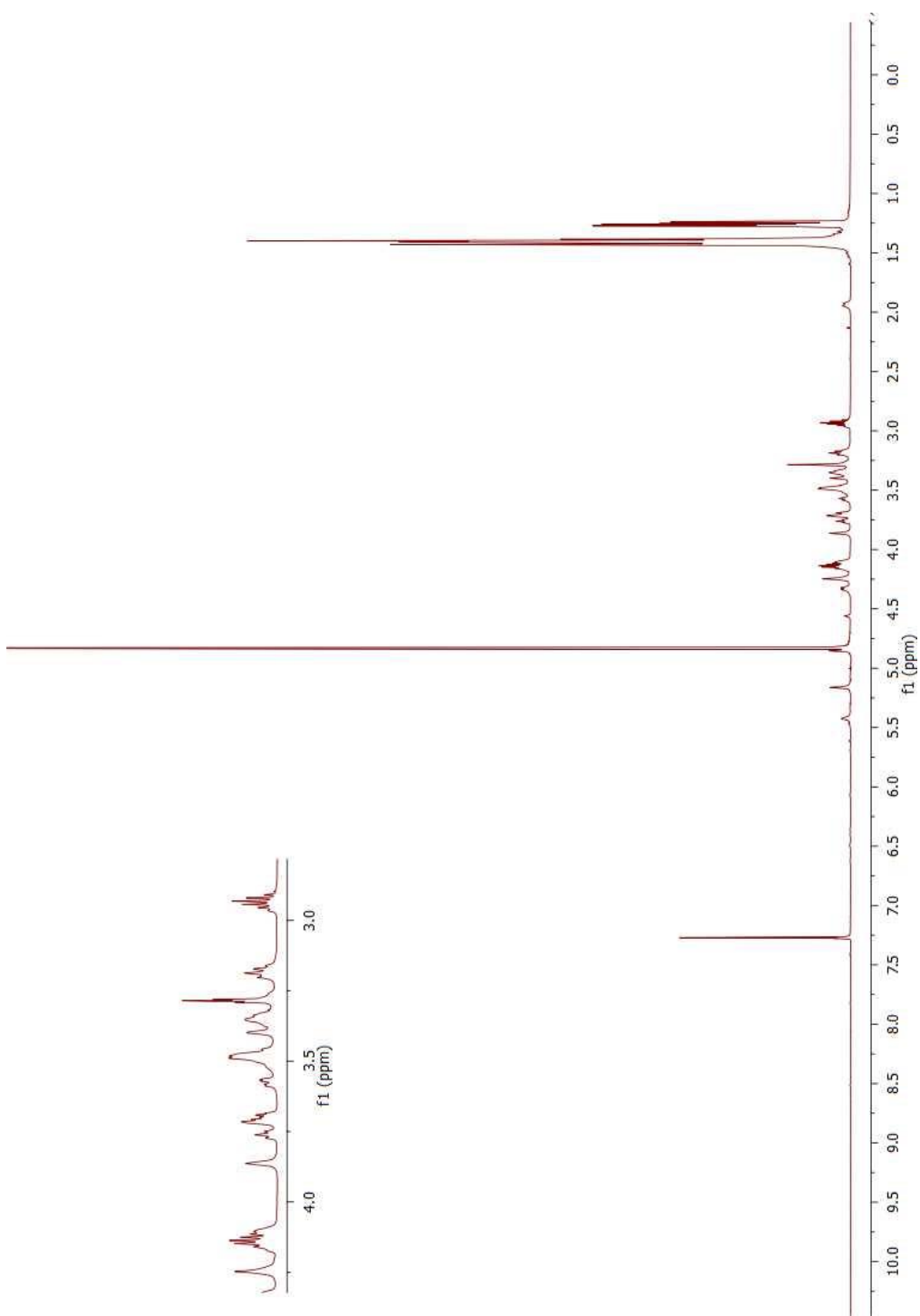
^{1}H NMR of compound **3**
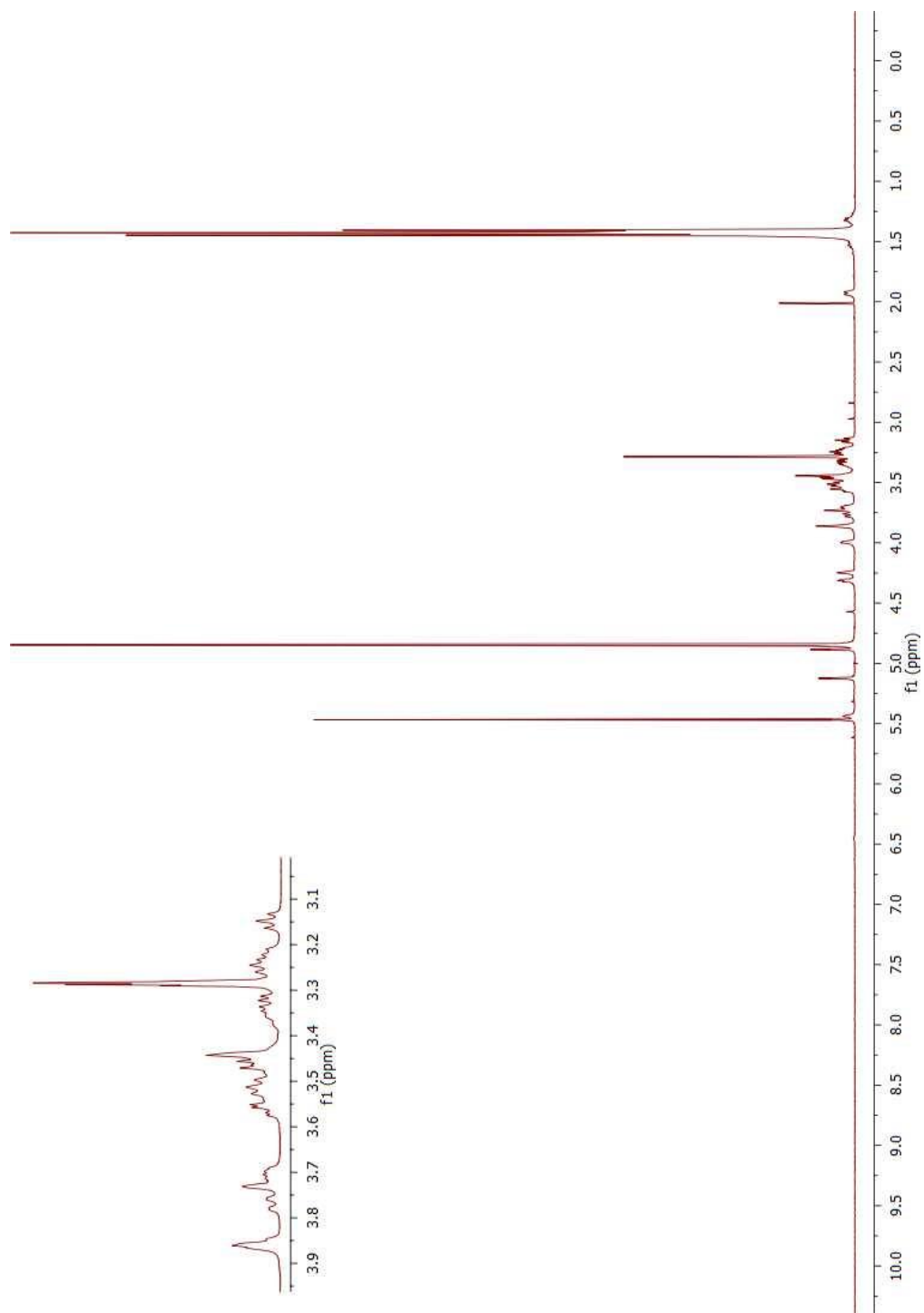
^{1}H NMR of compound **4**
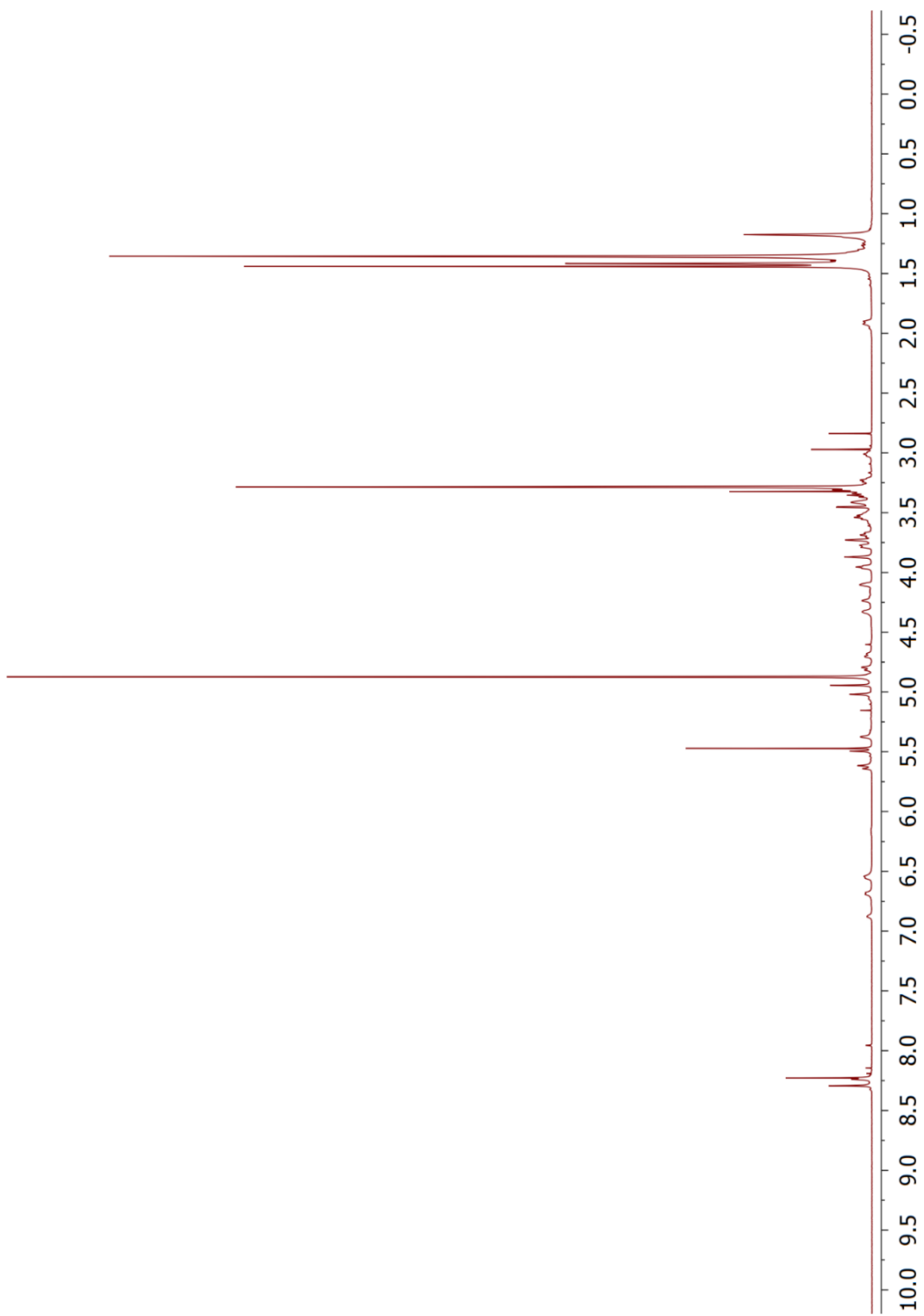
¹H NMR of compound **5a**
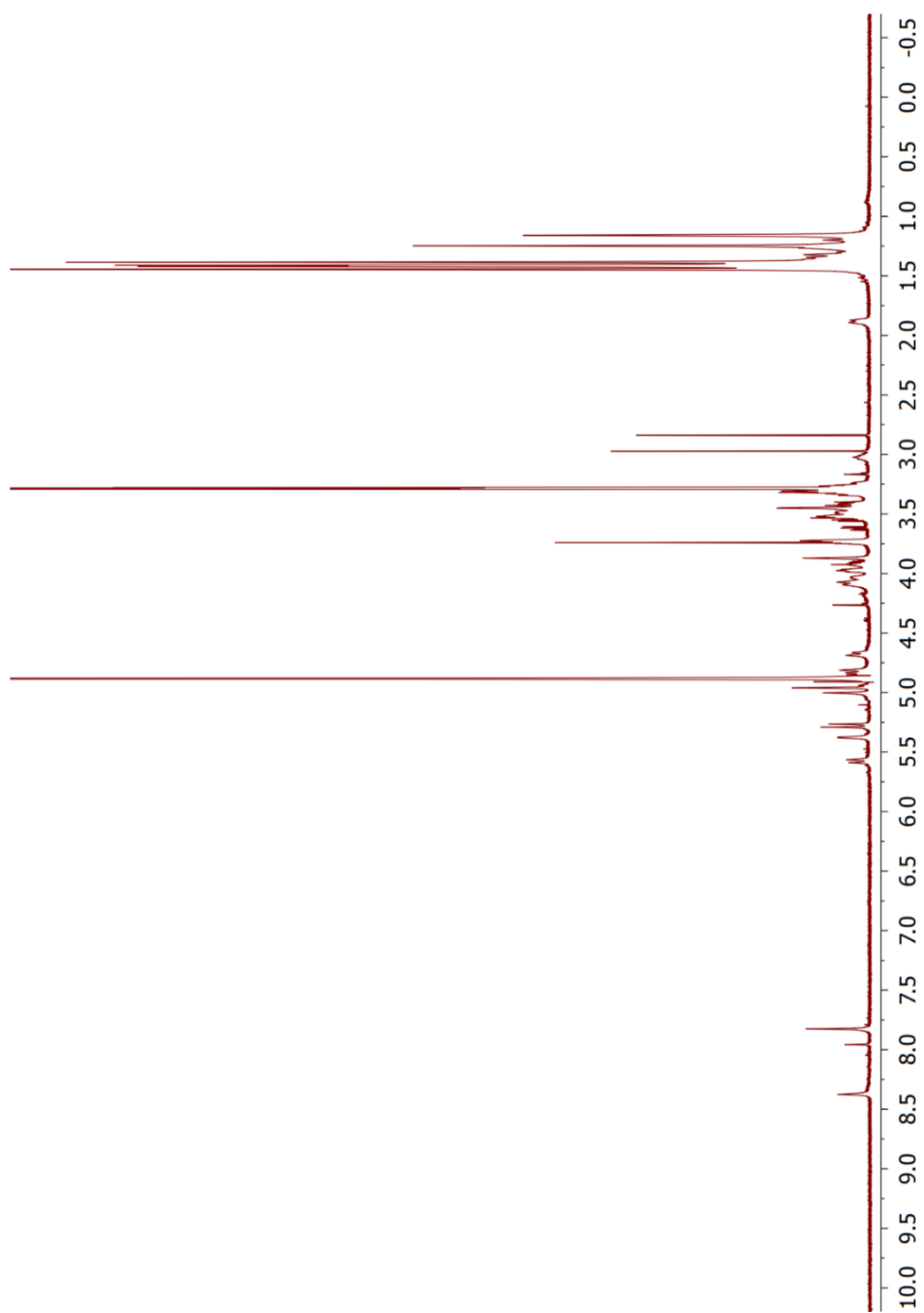
^{1}H NMR of compound **5b**
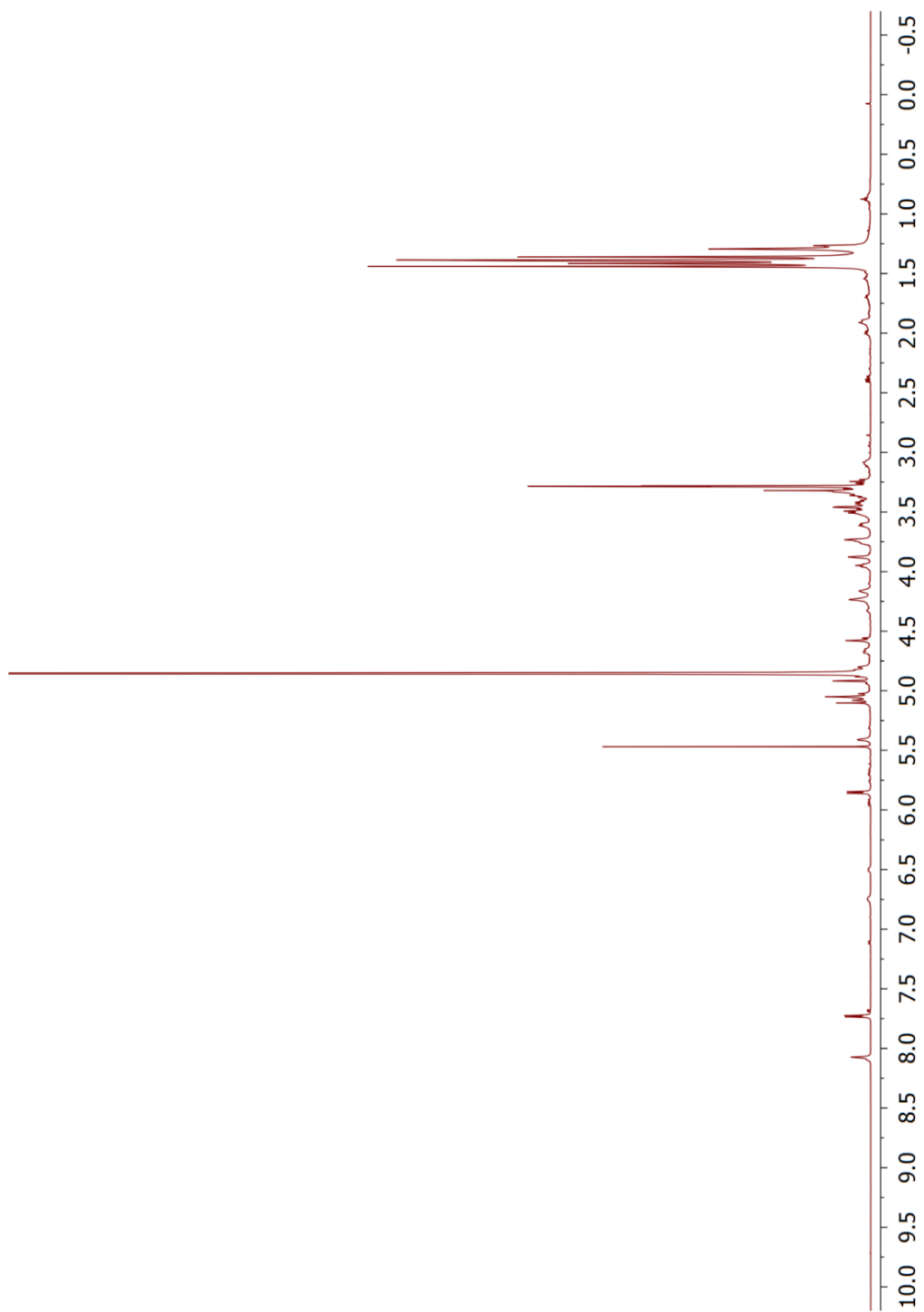
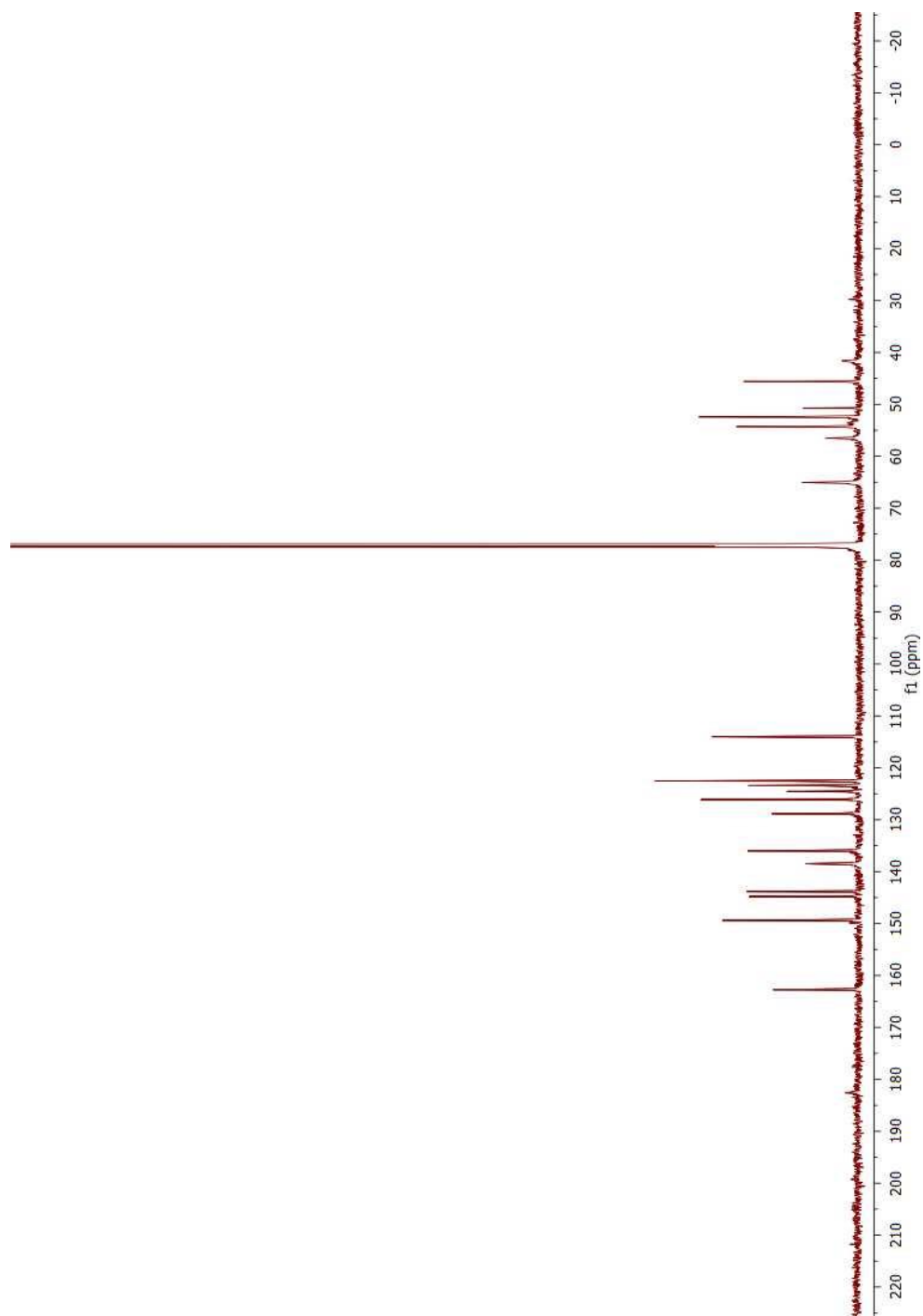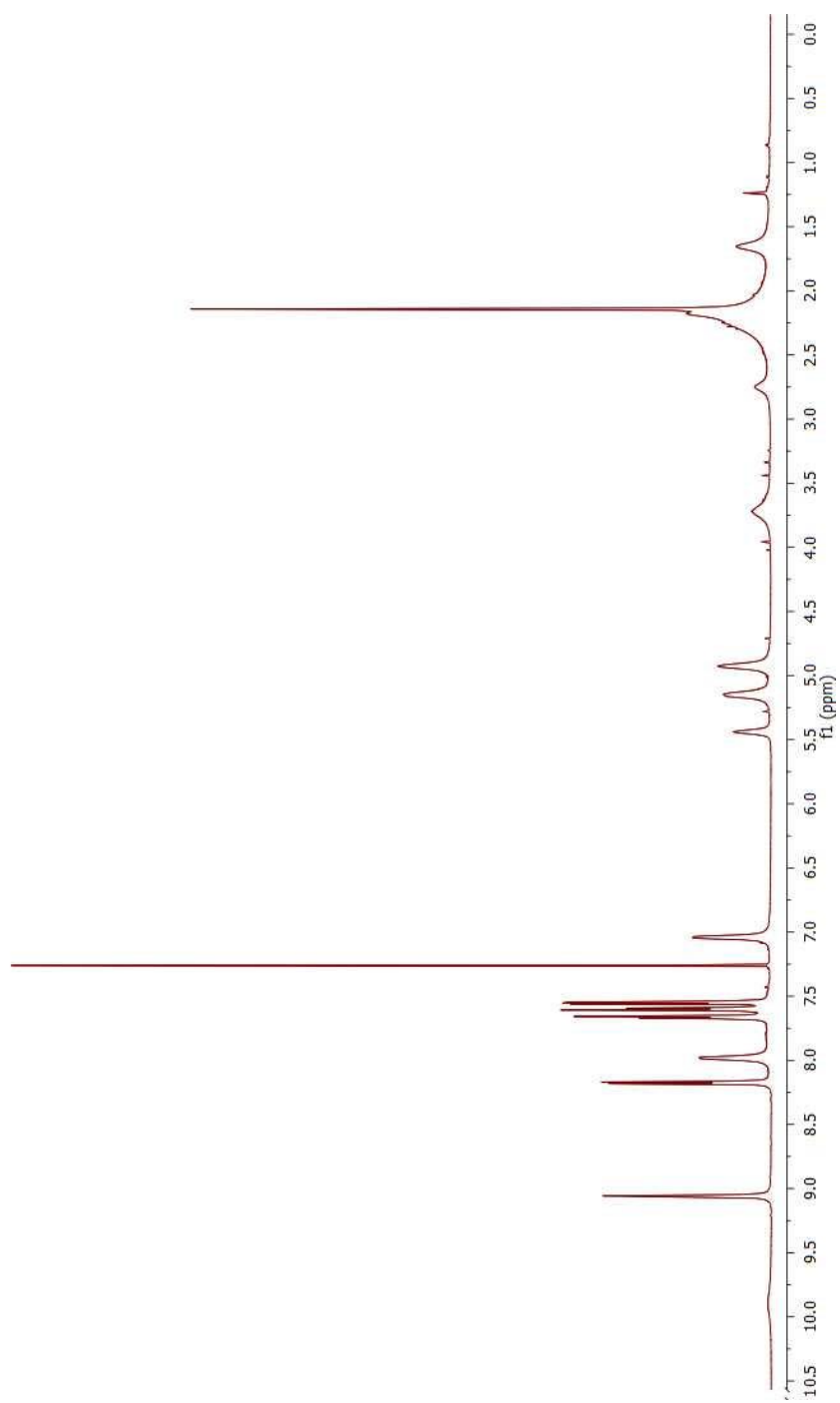^{1}H NMR of compound **5c**
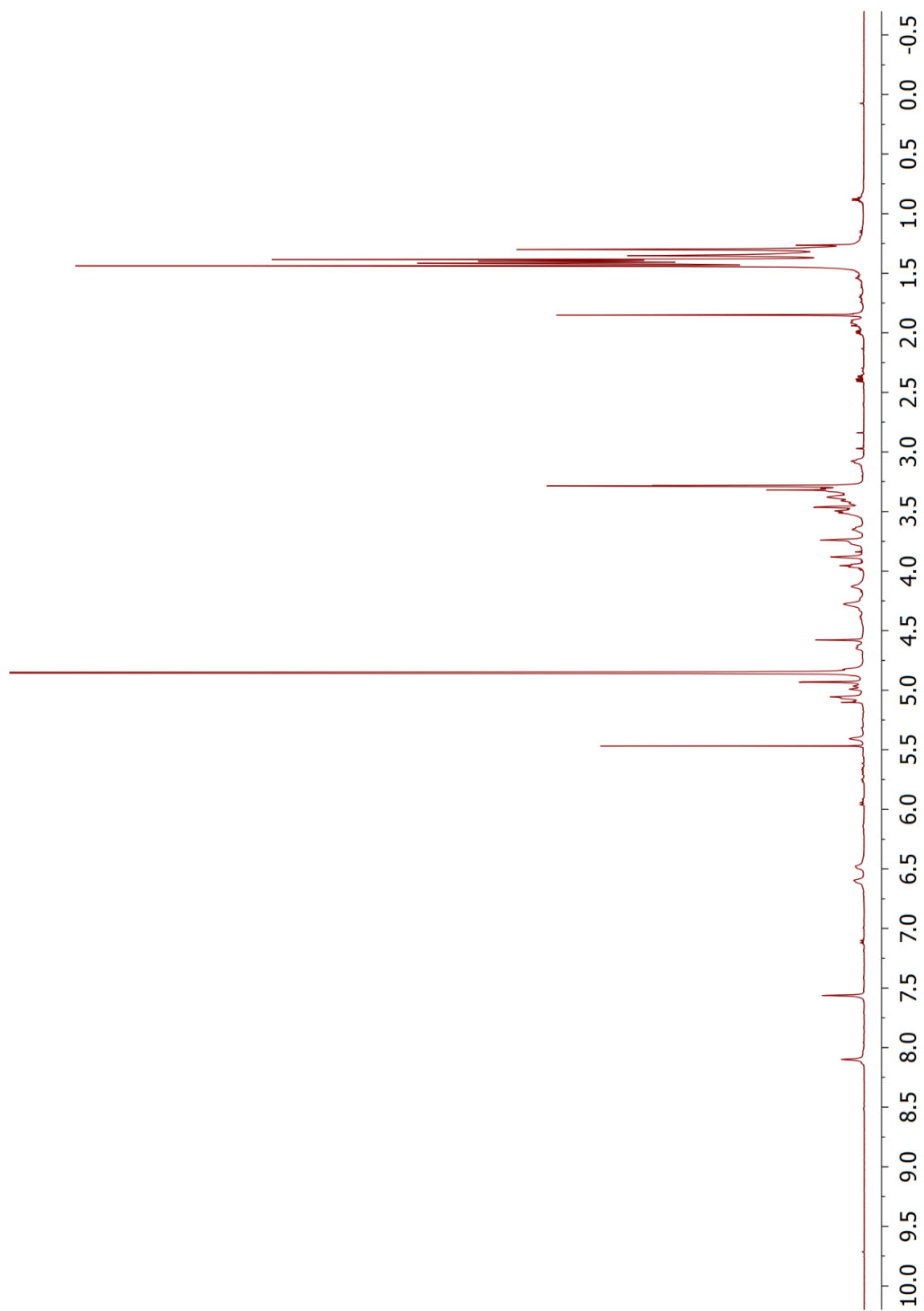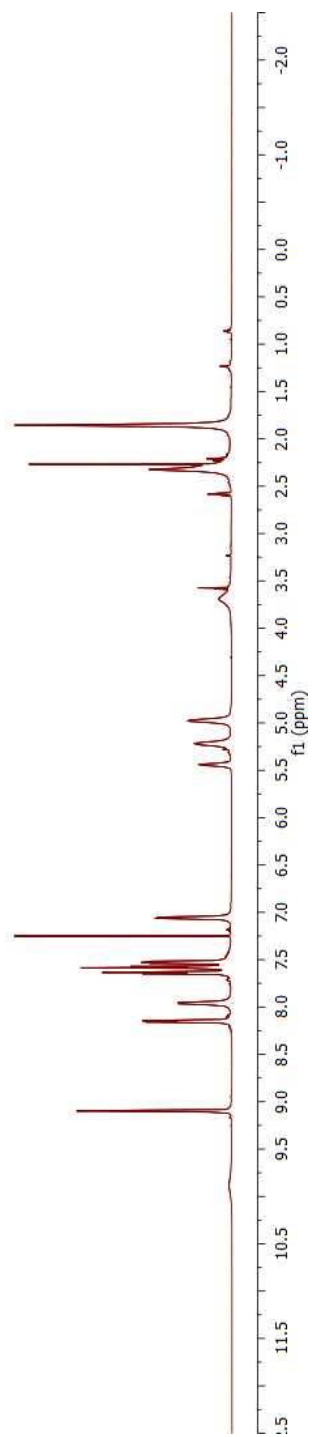
^{1}H NMR of compound **5d**
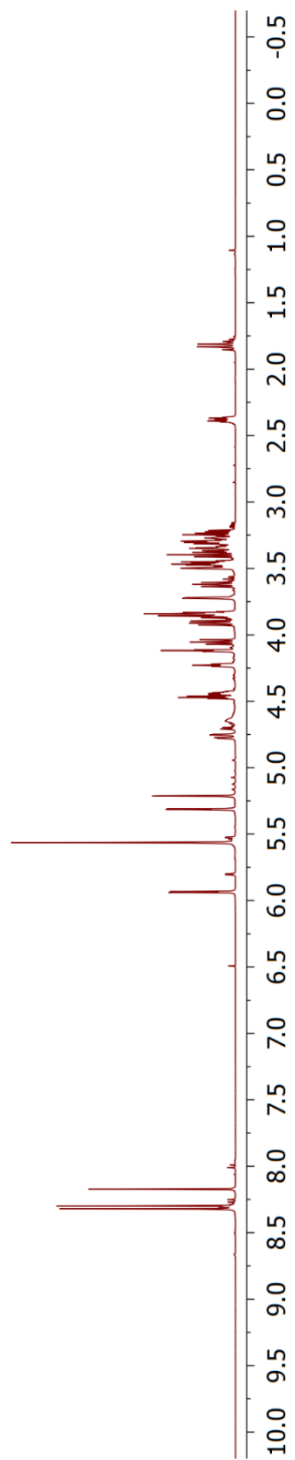
^{1}H NMR of compound **6a**
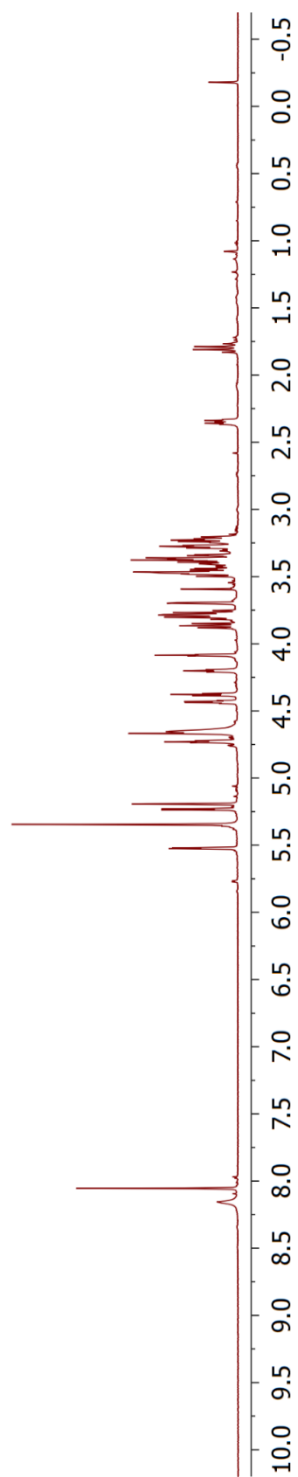
^{1}H NMR of compound **6b**
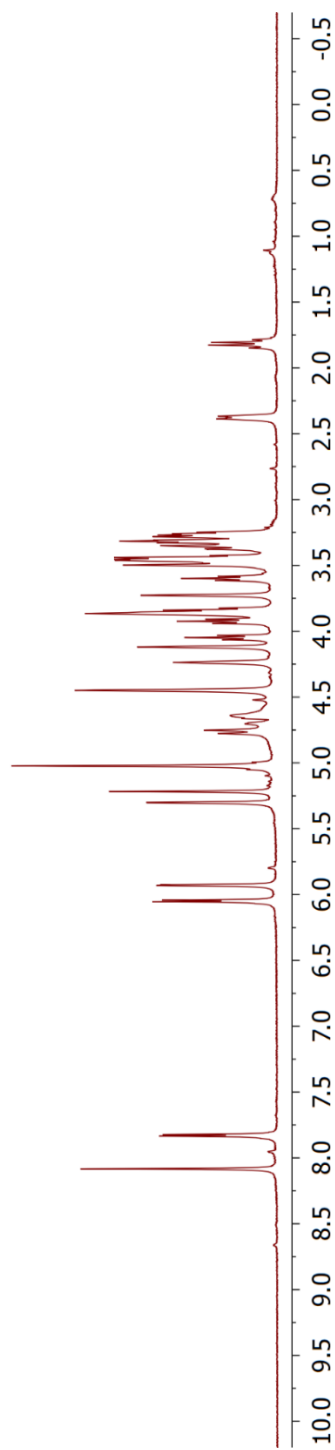
^{1}H NMR of compound **6c**
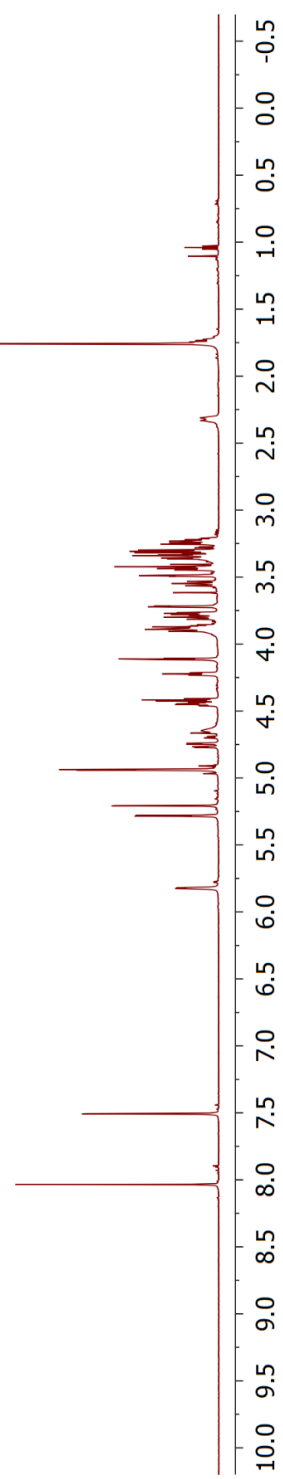
^{1}H NMR of compound **6d**
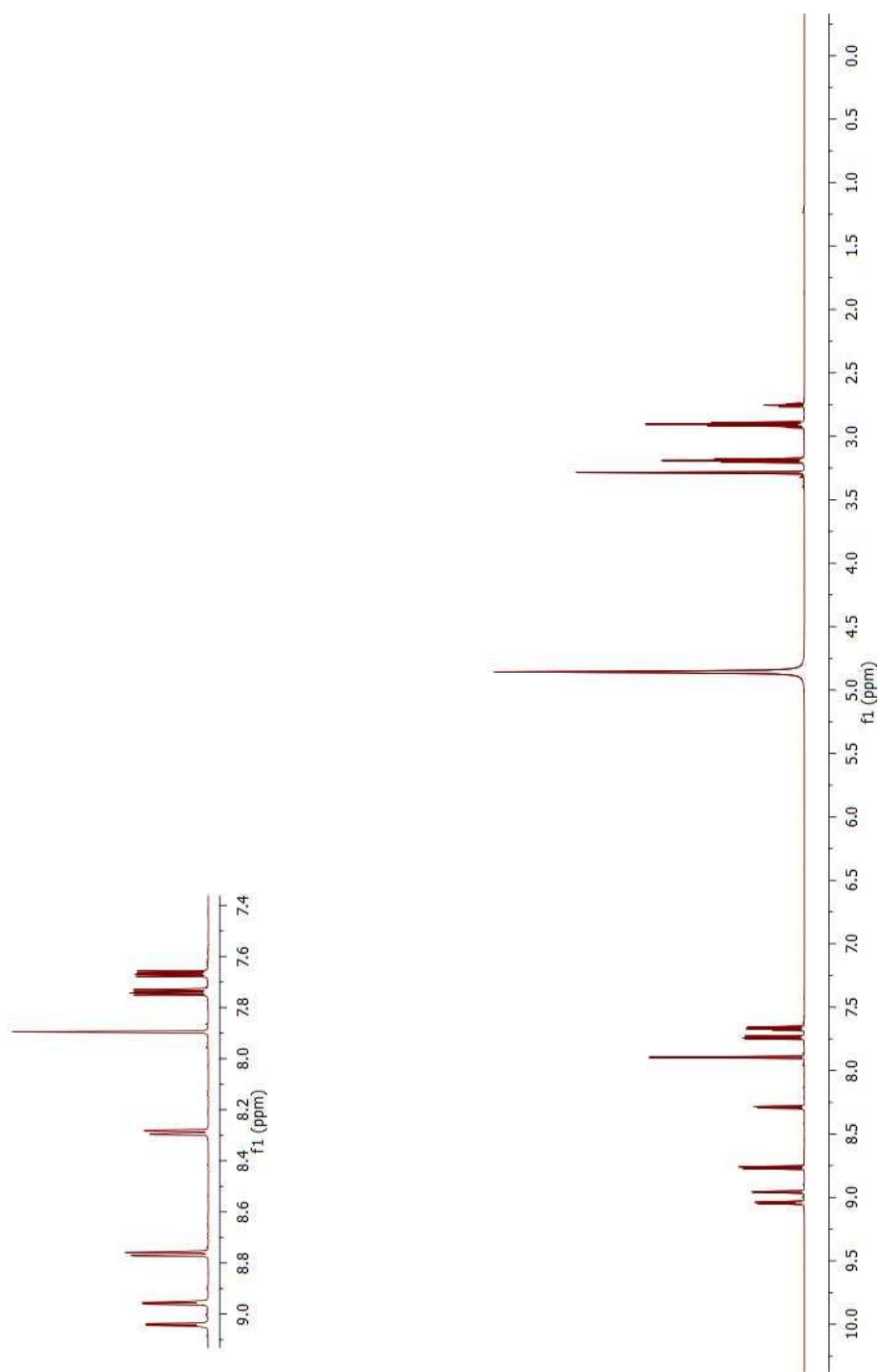
^{1}H NMR of compound **9**
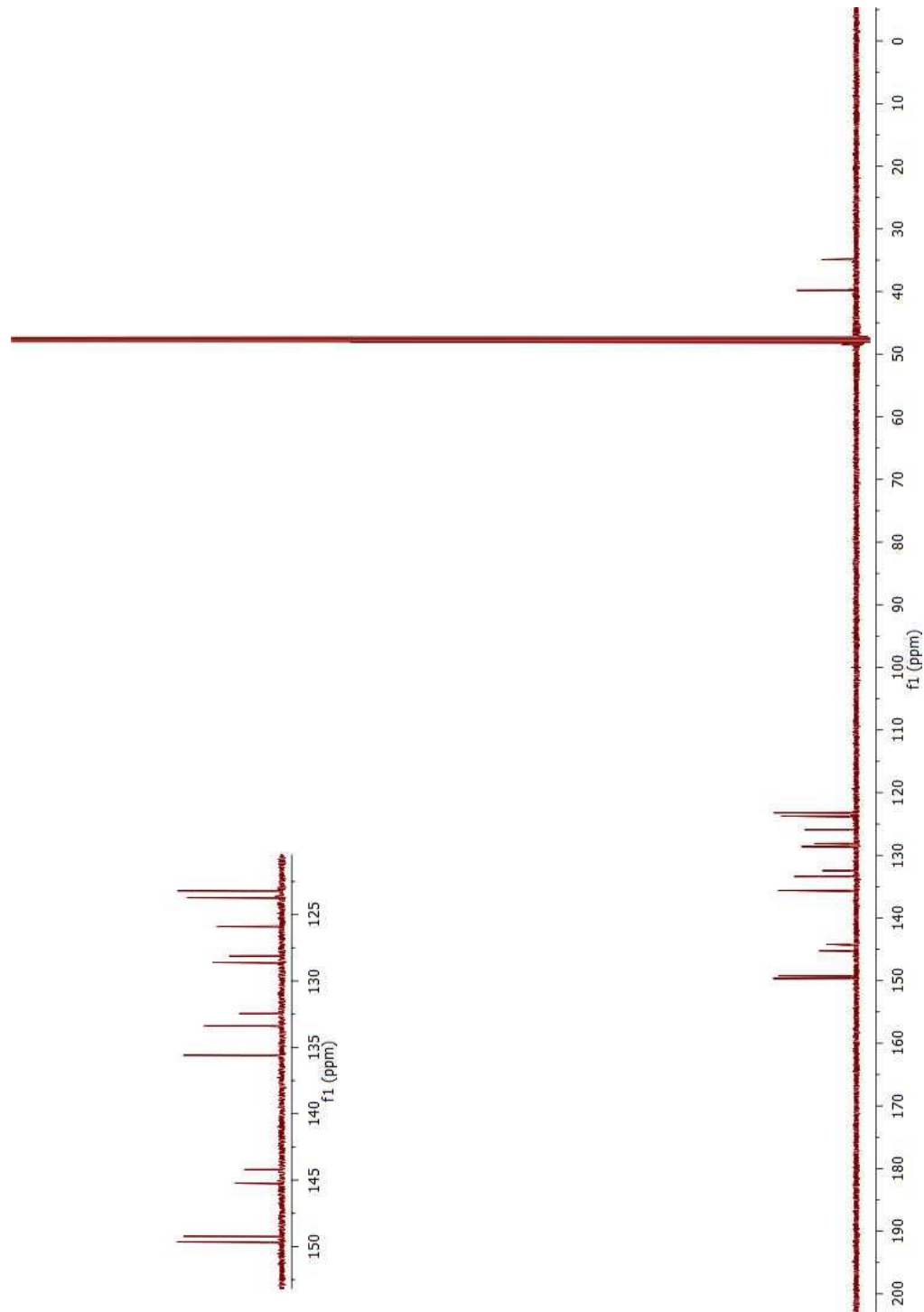
^{13}C NMR of compound **9**
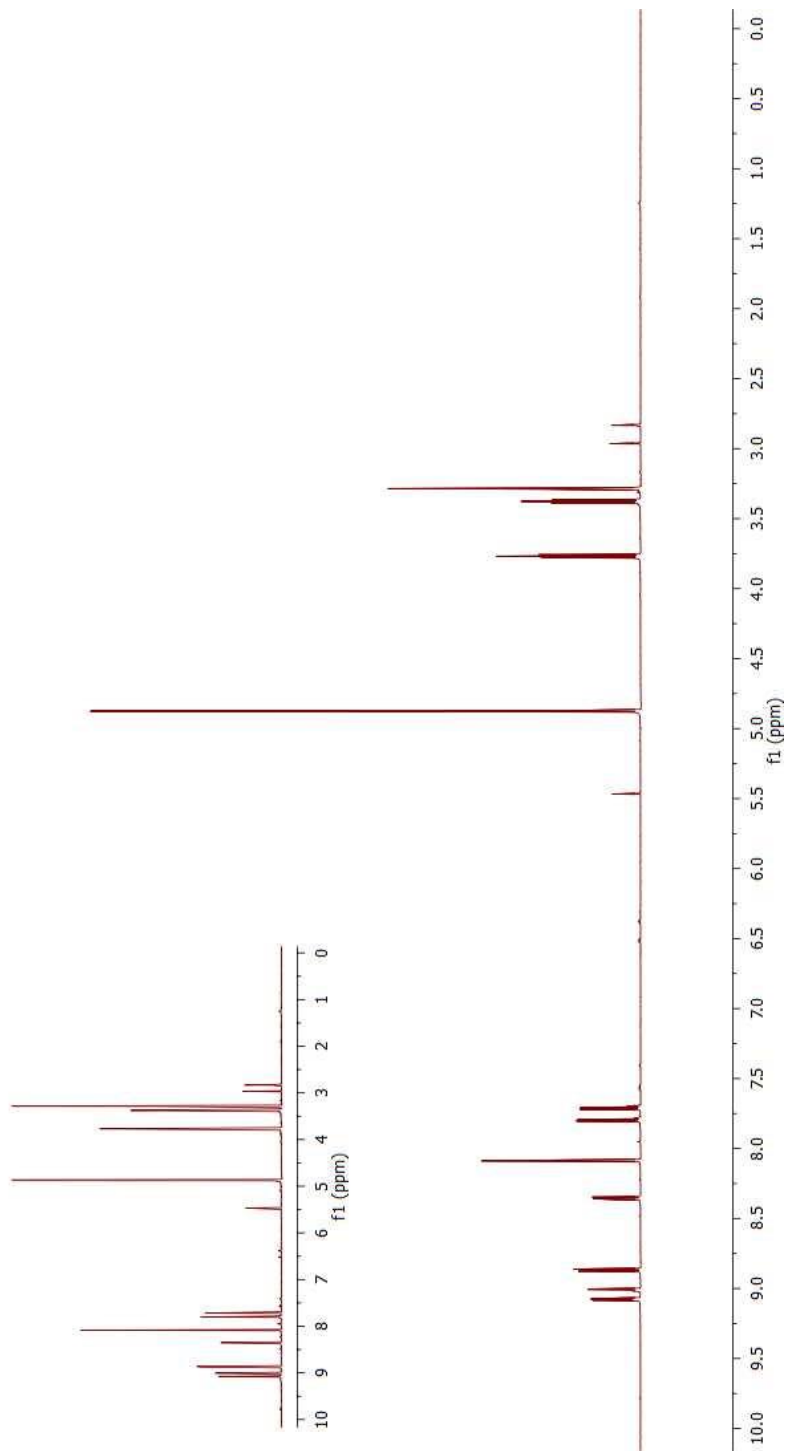
^{1}H NMR of compound **9NCS**
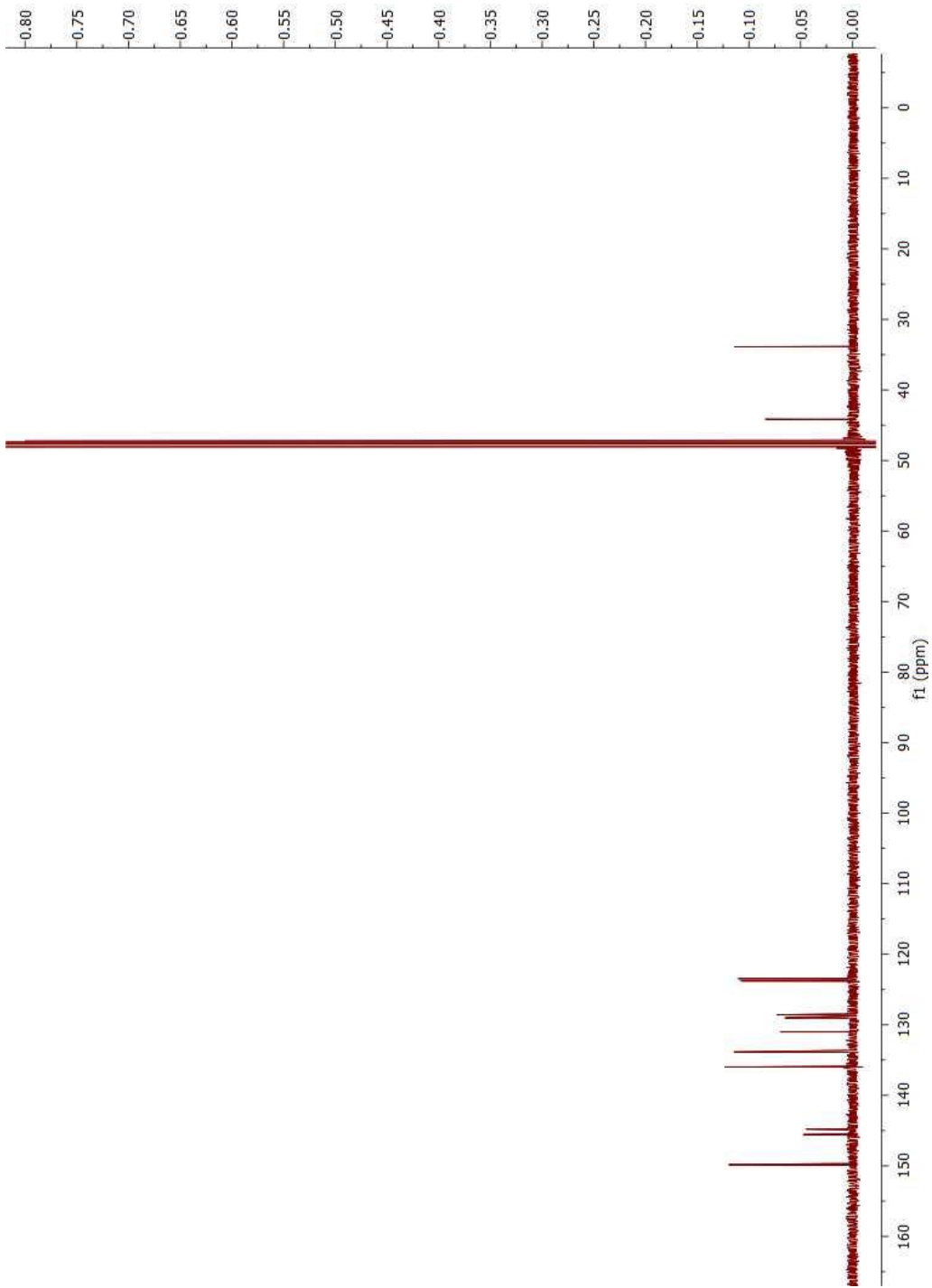
^{13}C NMR of compound **9NCS**
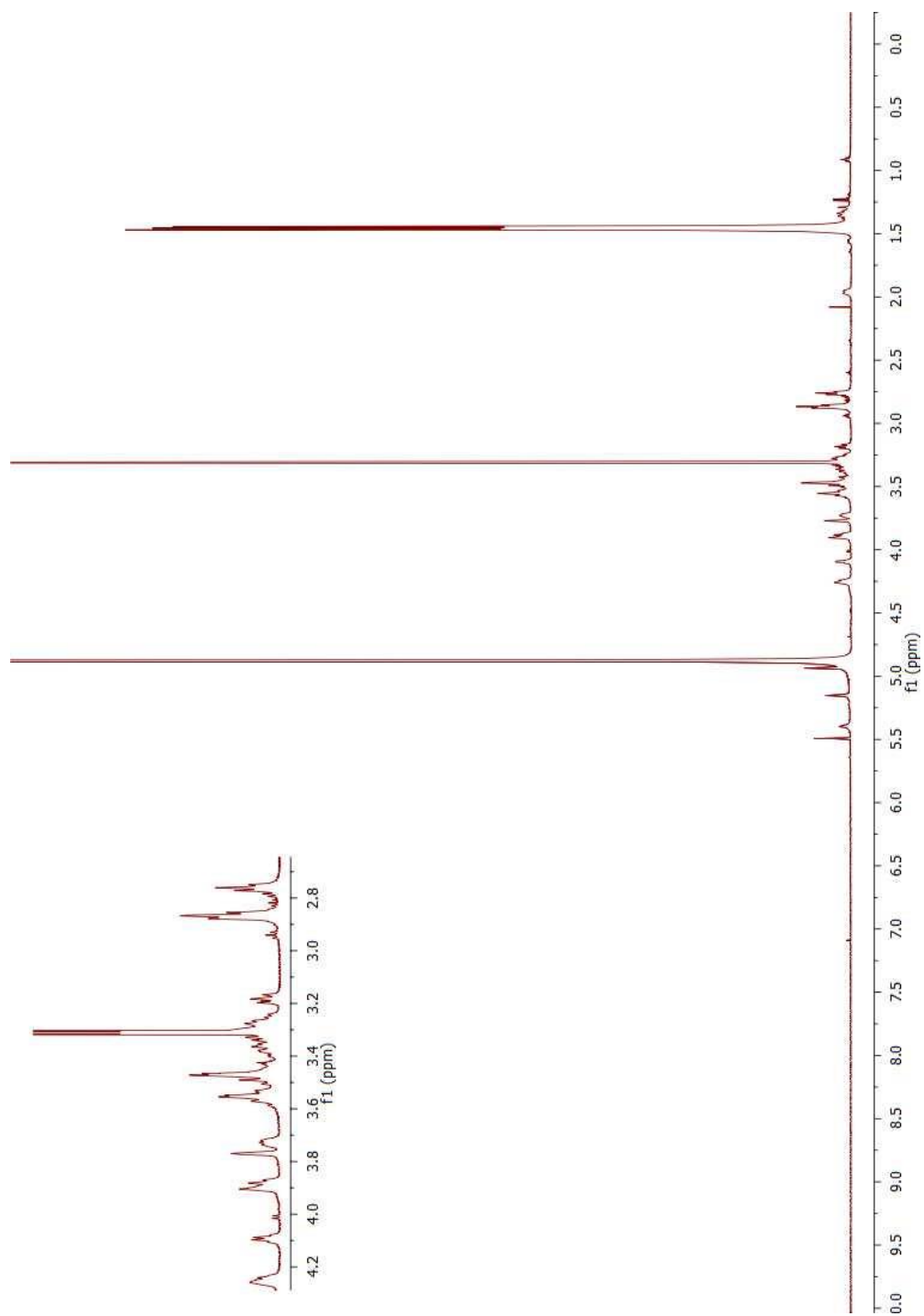
^{1}H NMR of compound **5a-L**
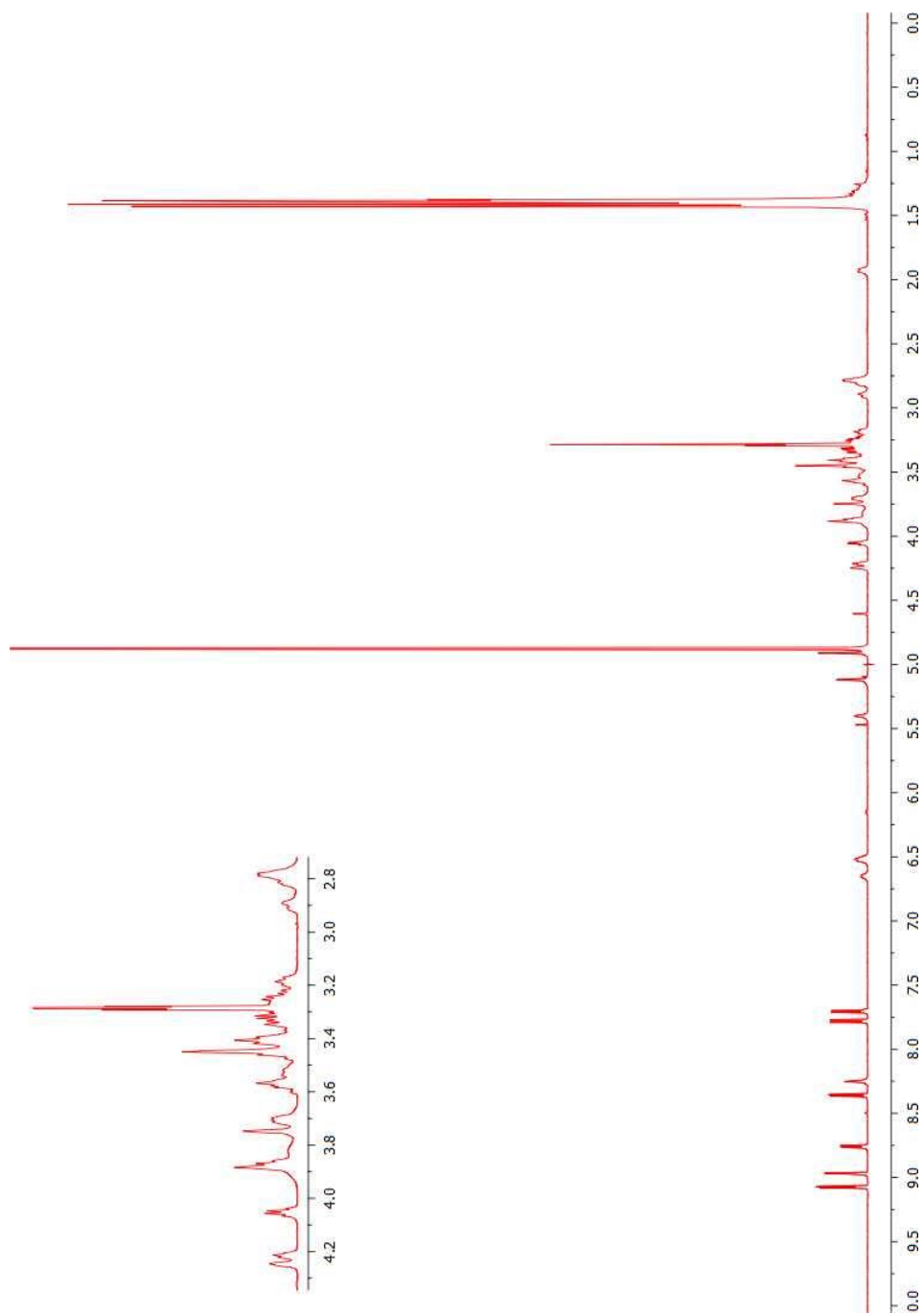
^{1}H NMR of compound **6a-L**
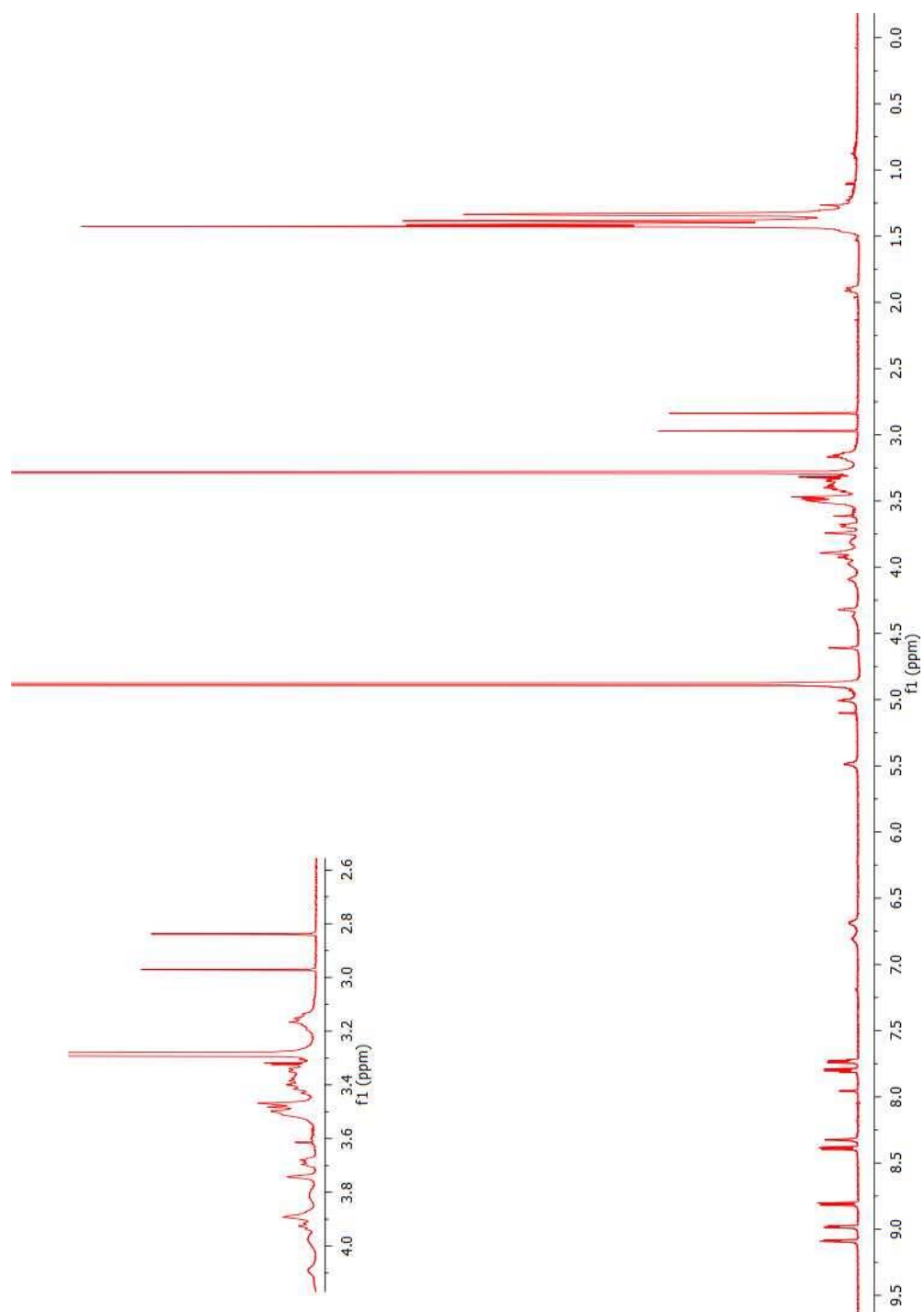
^{1}H NMR of compound **6b-S**
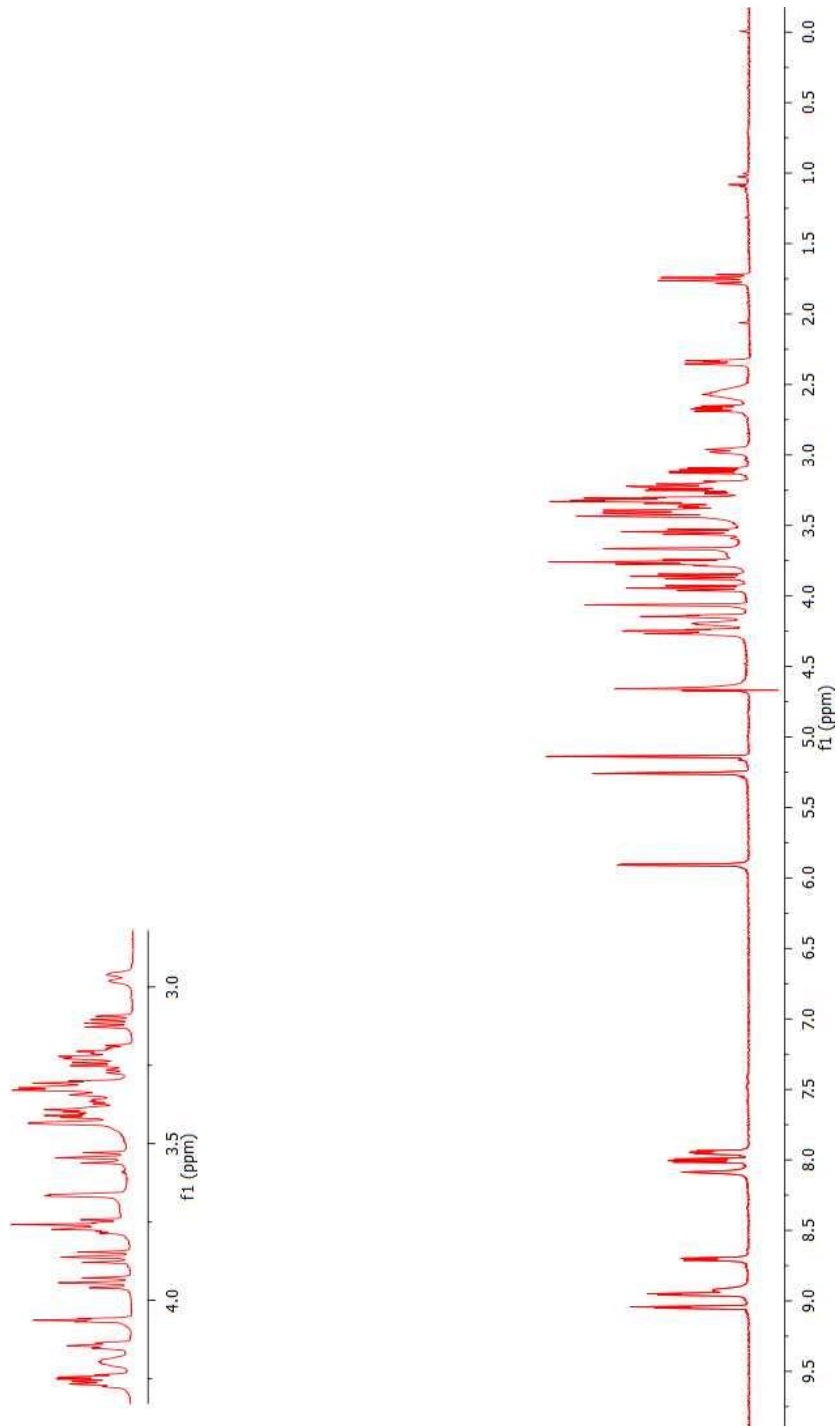
^{1}H NMR of compound **7a**
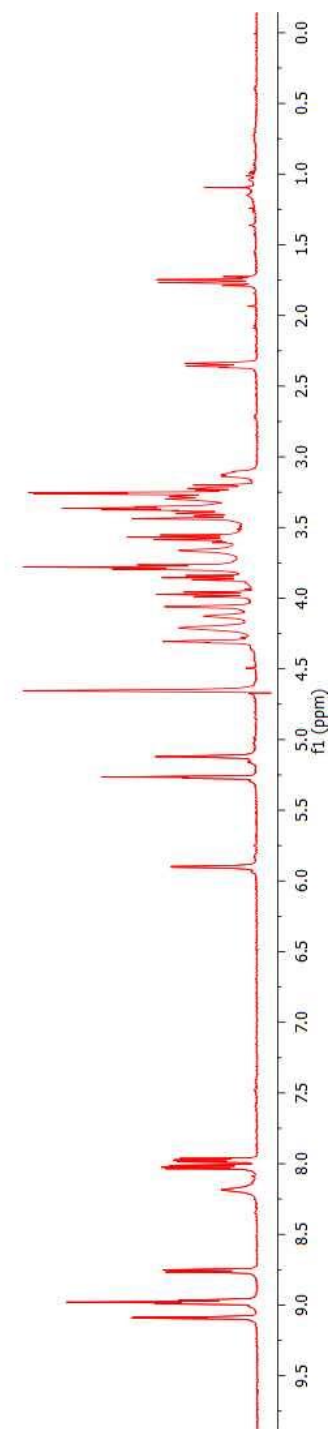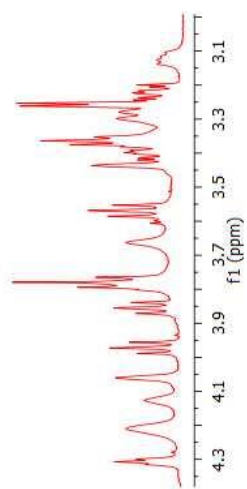
^{1}H NMR of compound **7b**
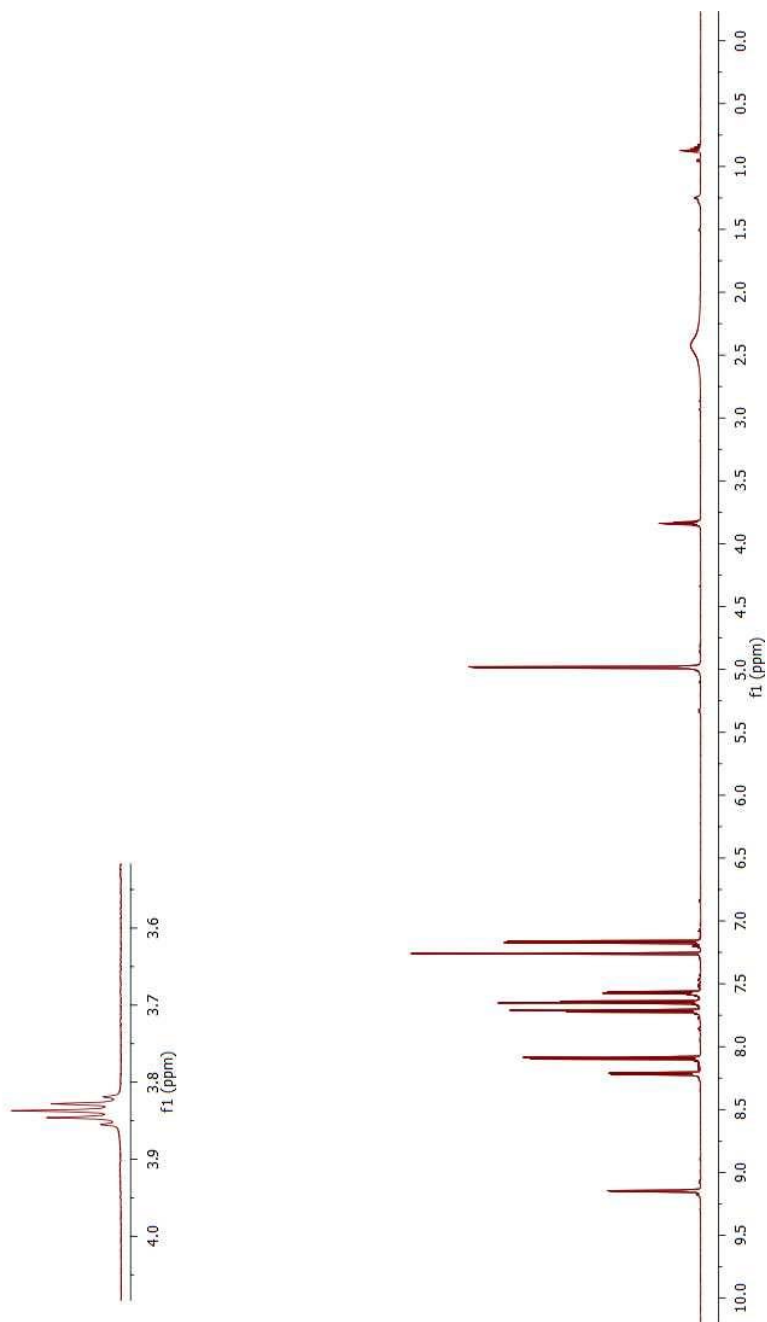
^{1}H NMR of compound **2-Clip-phen**
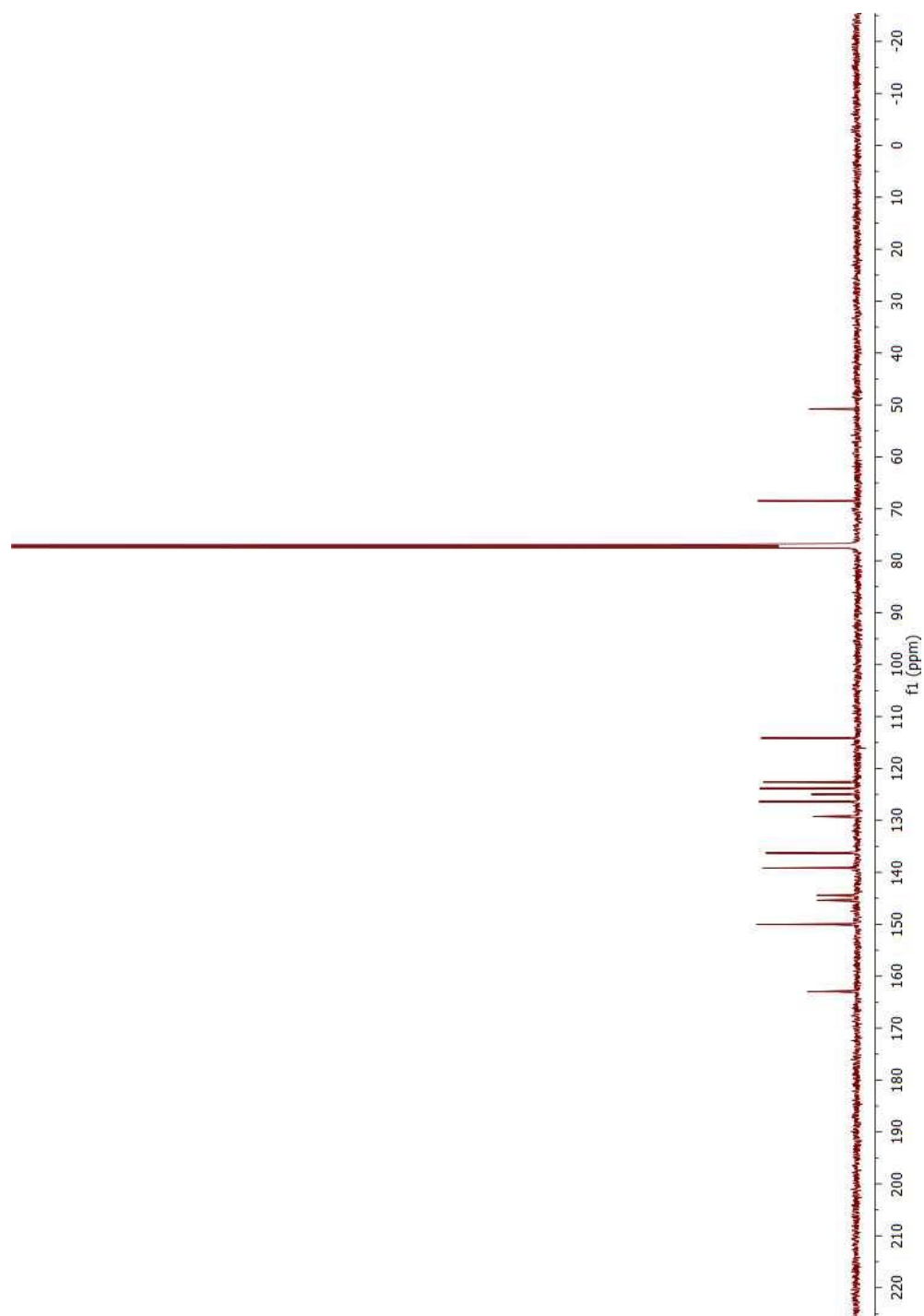
^{1}H NMR of compound **2-Clip-phen**
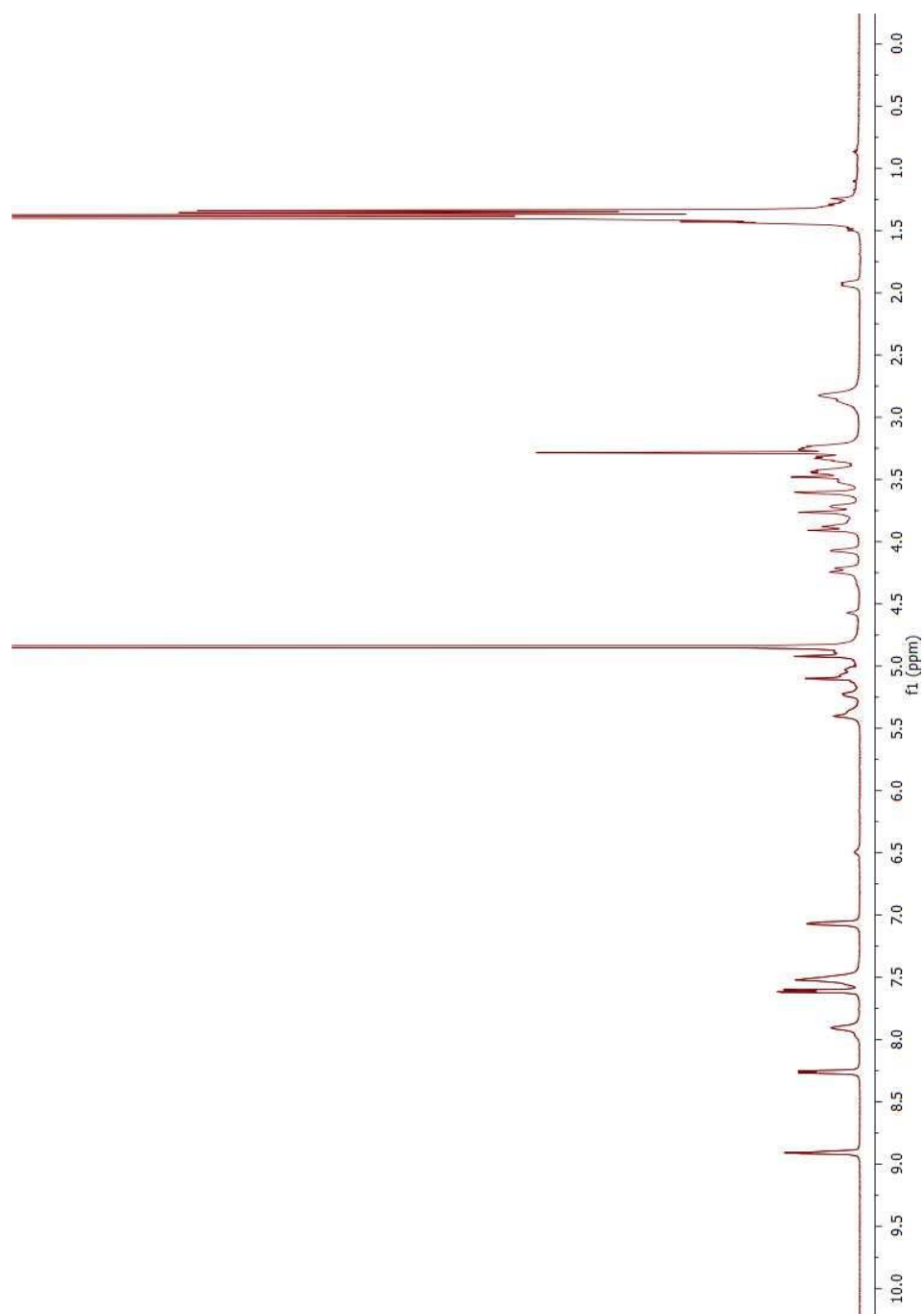
¹H NMR of compound **2-CP1**
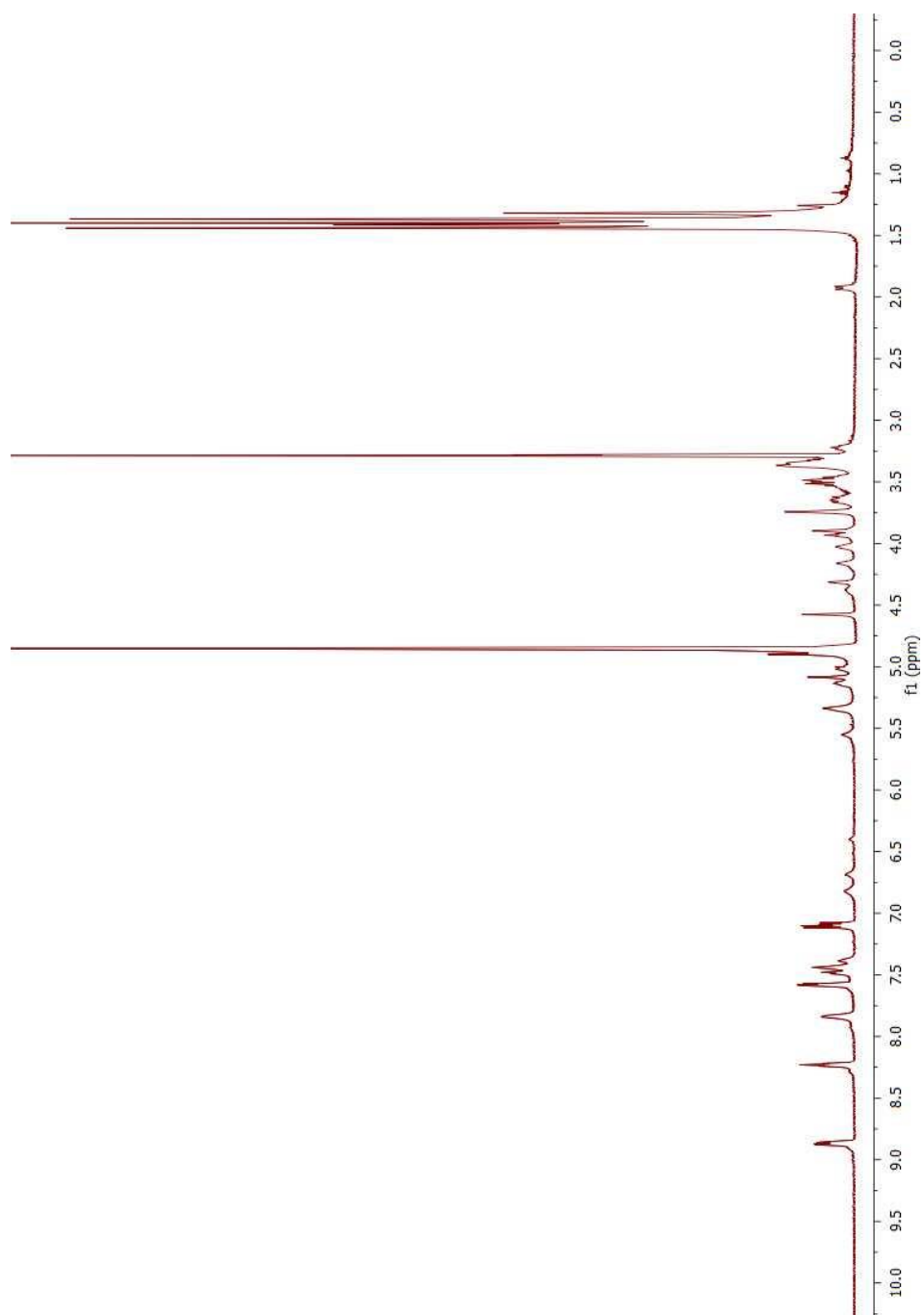
^{1}H NMR of compound **2-CP2**
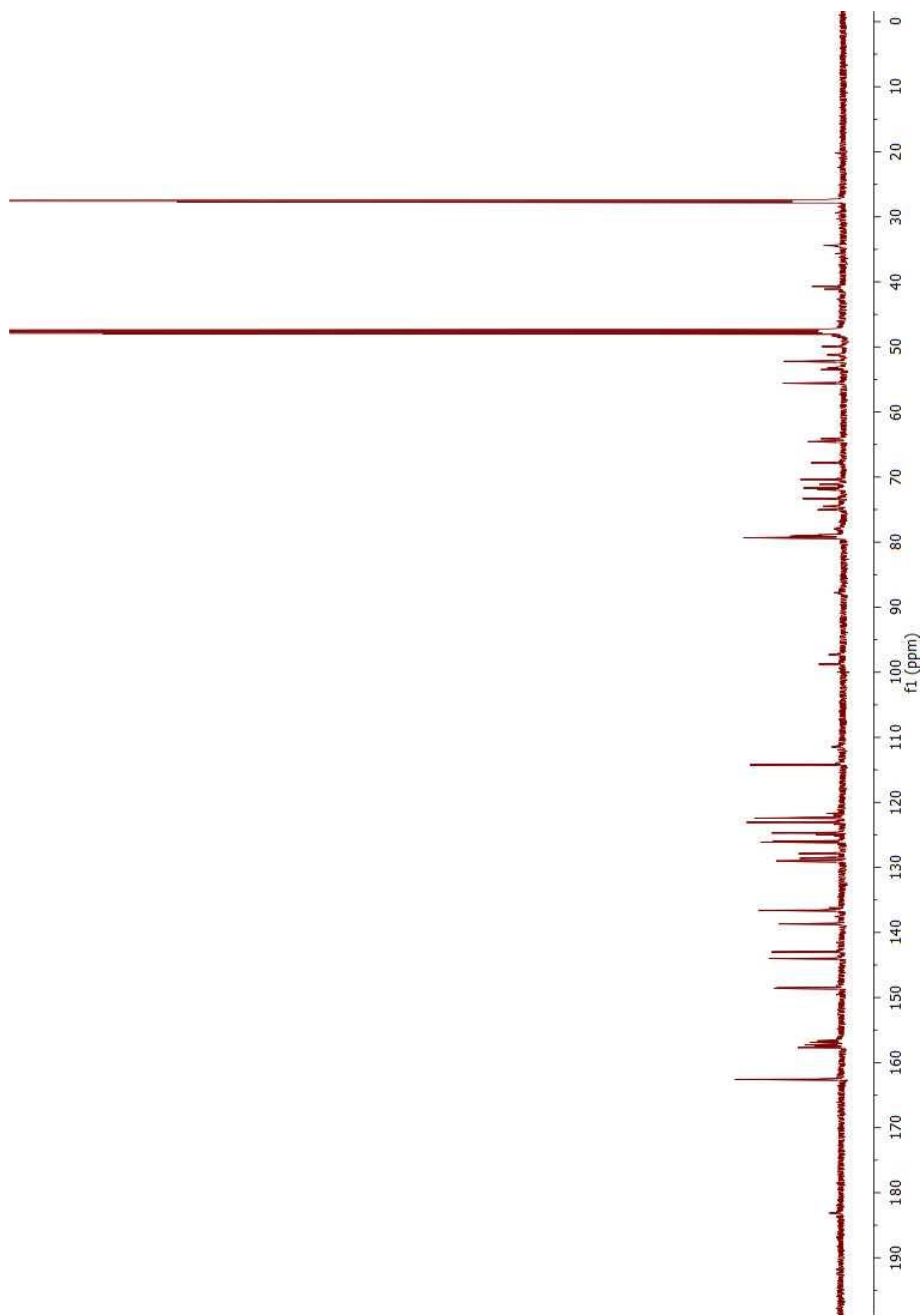
¹³C NMR of compound **2CP2**
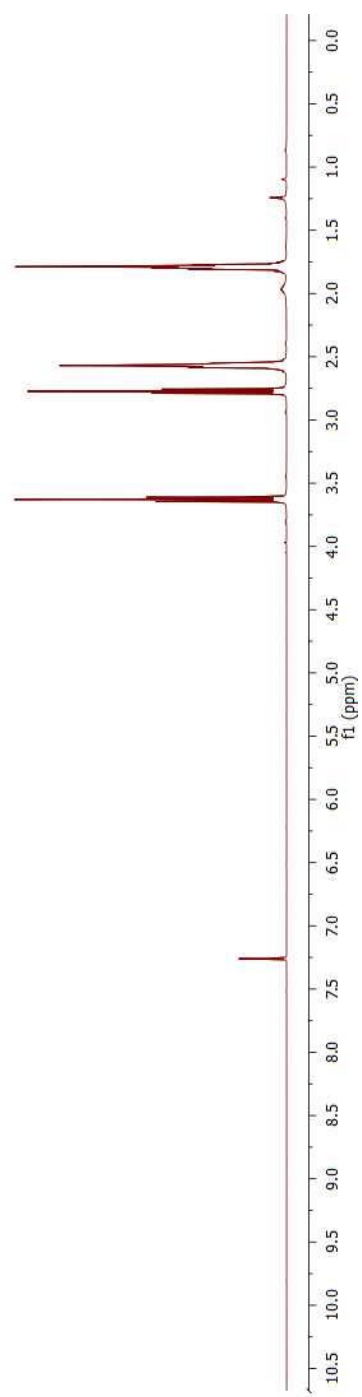
^{1}H NMR of compound **AM1**
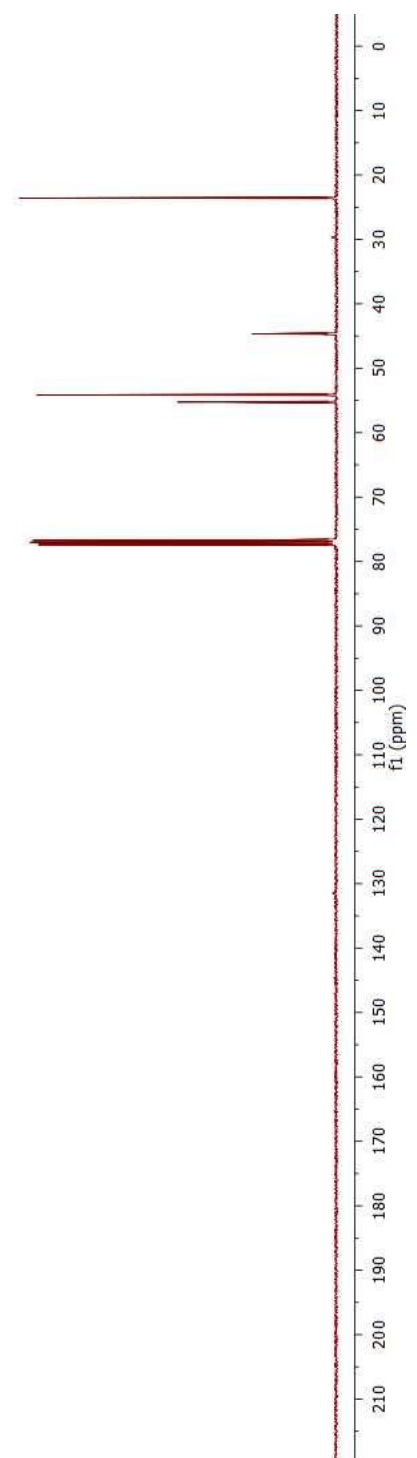
^{13}C NMR of compound **AM1**,
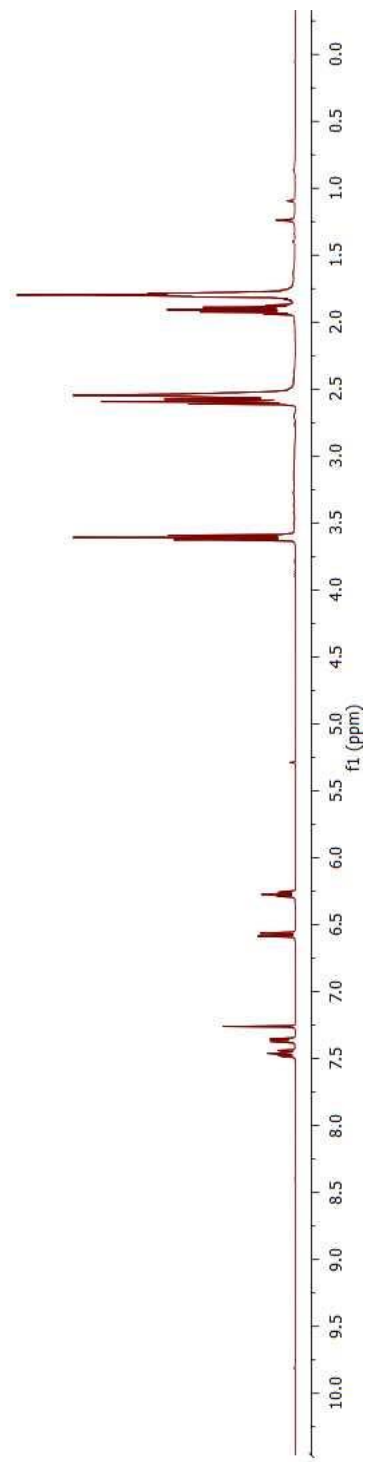
^{1}H NMR of compound **AM2**
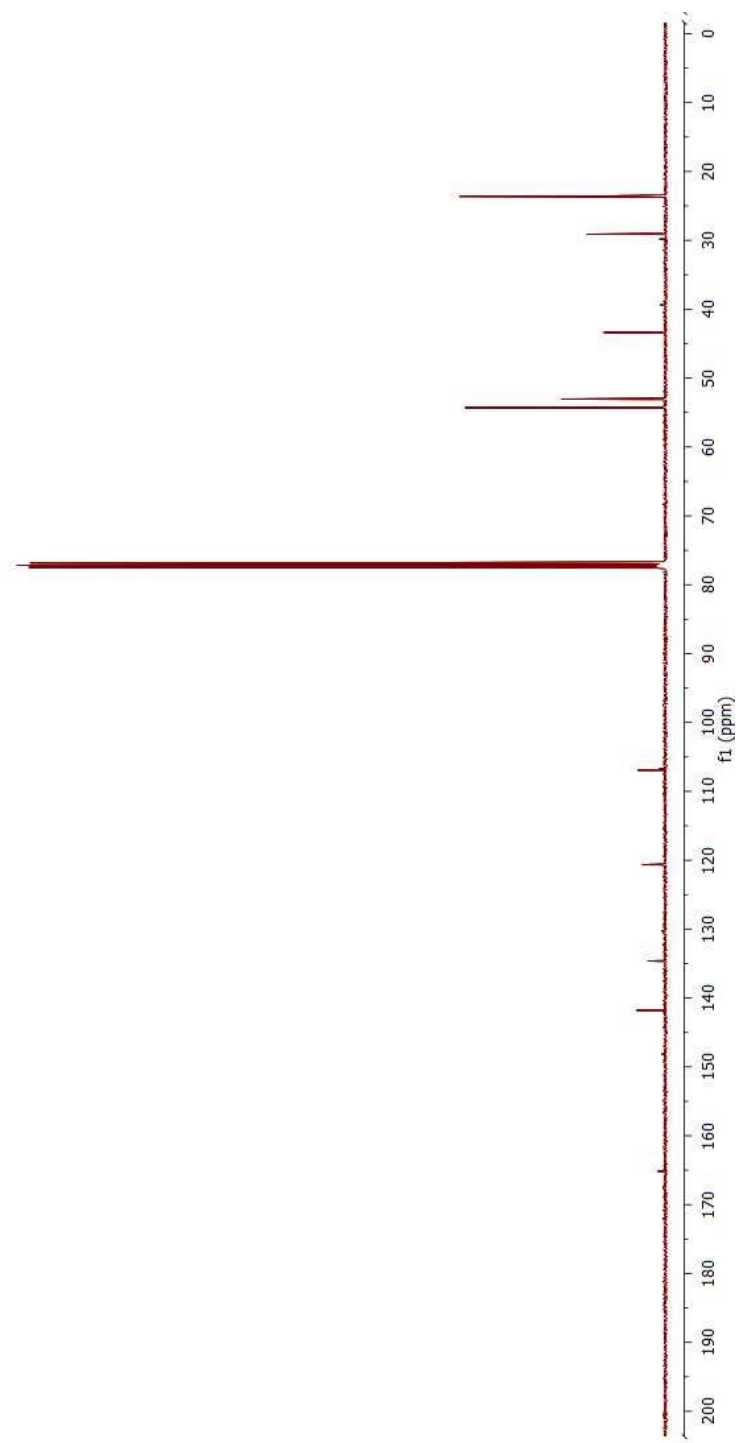
^{13}C NMR of compound **AM2,**

^{1}H NMR of compound **AM3**

14.25
21.08
60.43
171.15

132.6   132.2   131.8   131.4
f1 (ppm)

200 190 180 170 160 150 140 130 120 110 100 90   80   70   60   50   40   30   20   10   0
                                        f1 (ppm)
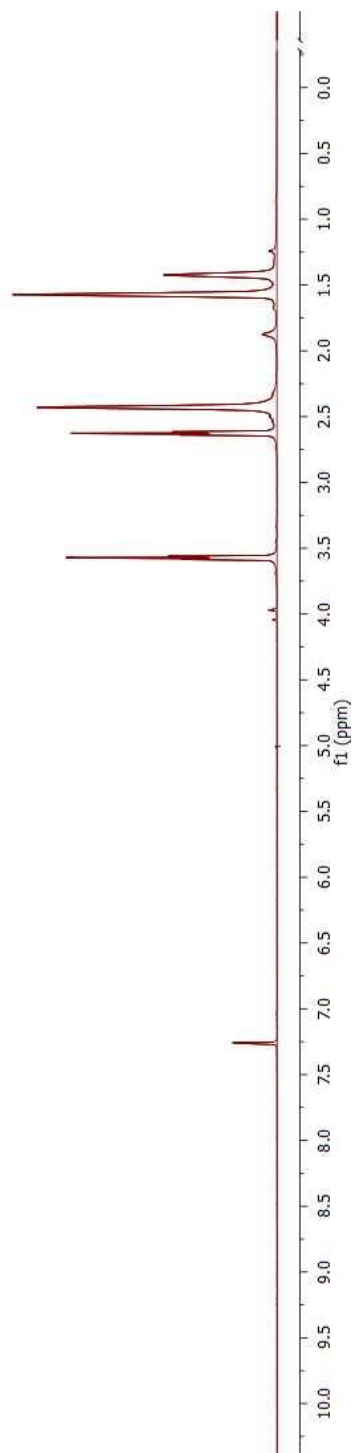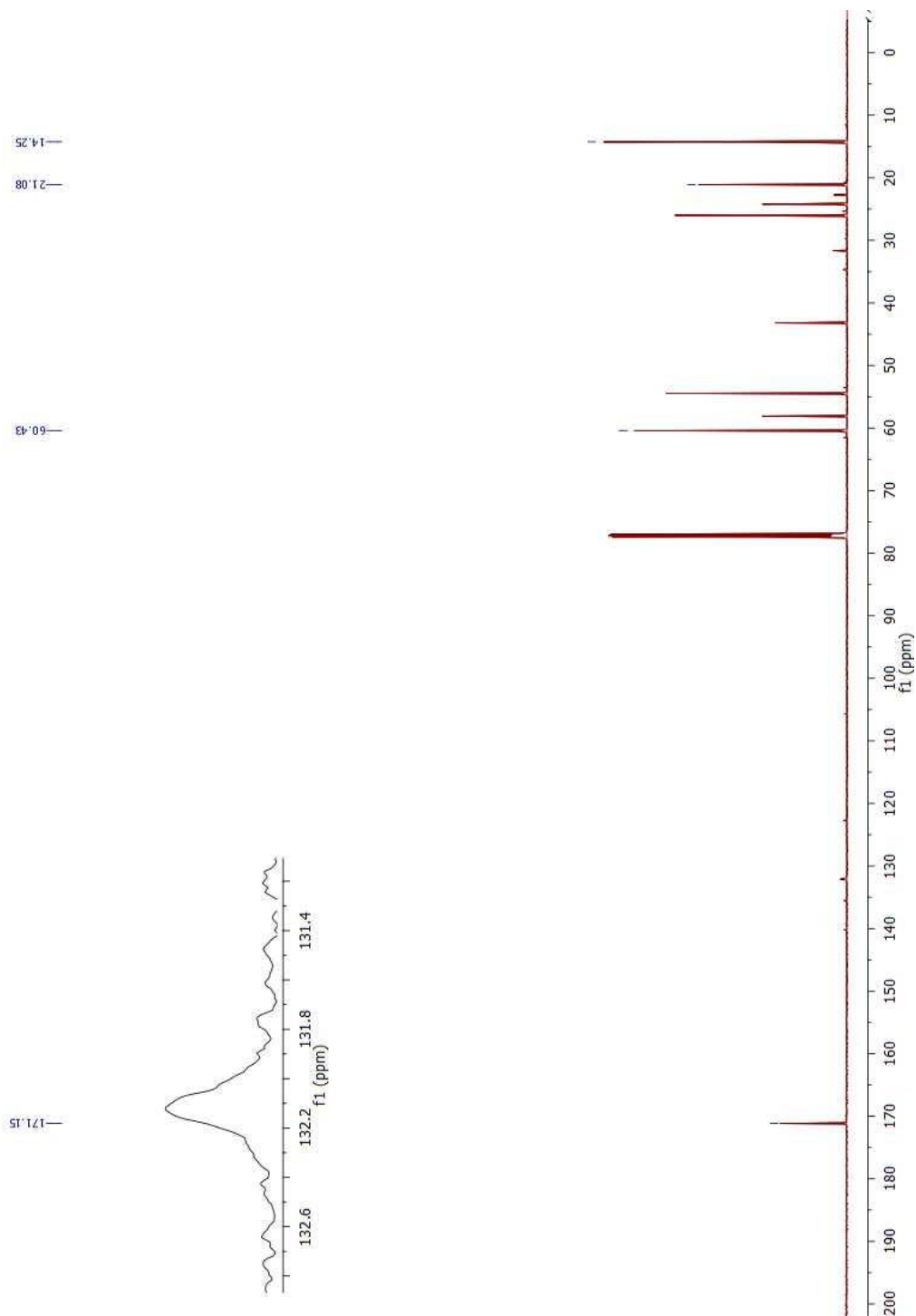
^{13}C NMR of compound **AM3**,
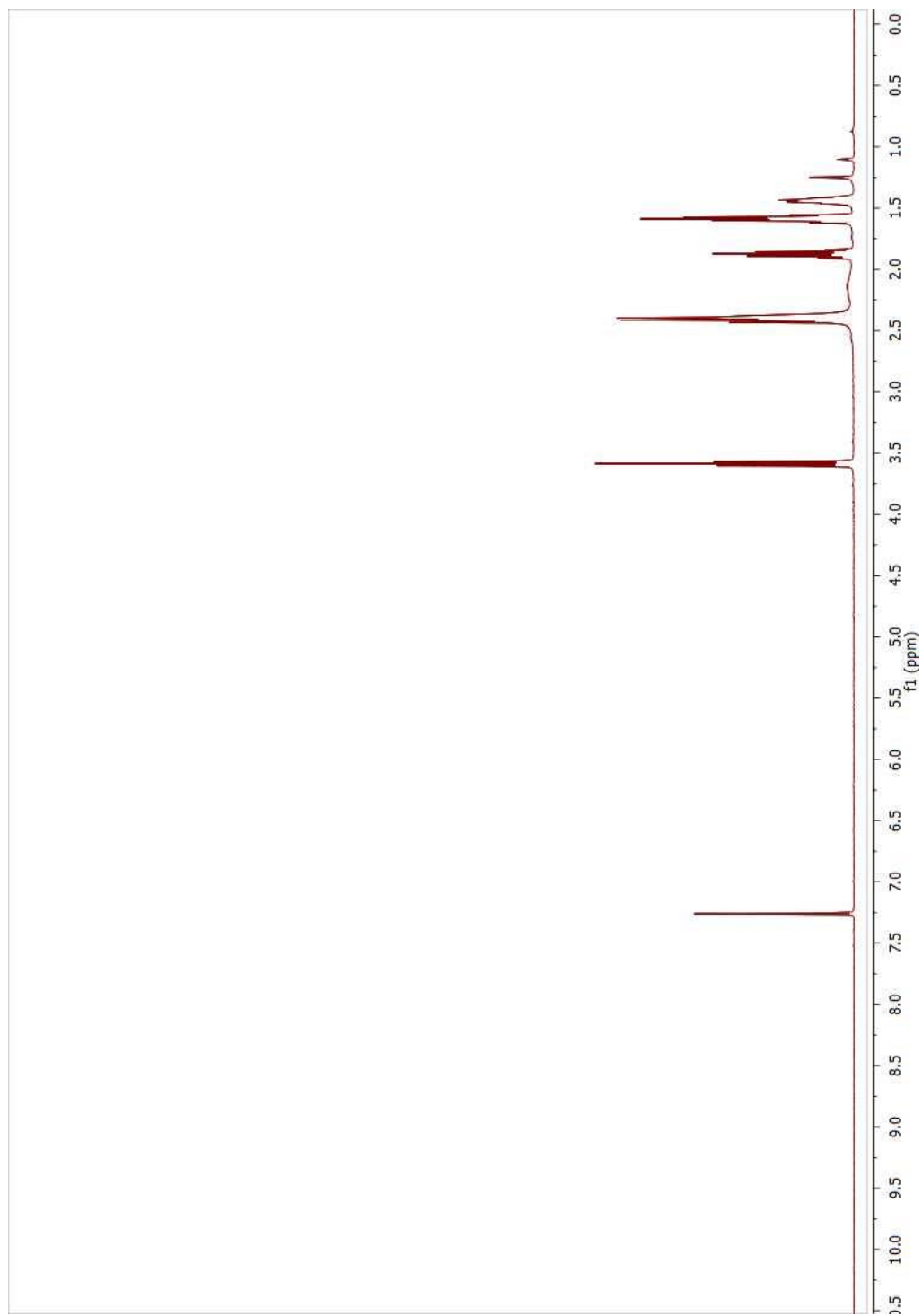
^{1}H NMR of compound **AM4**
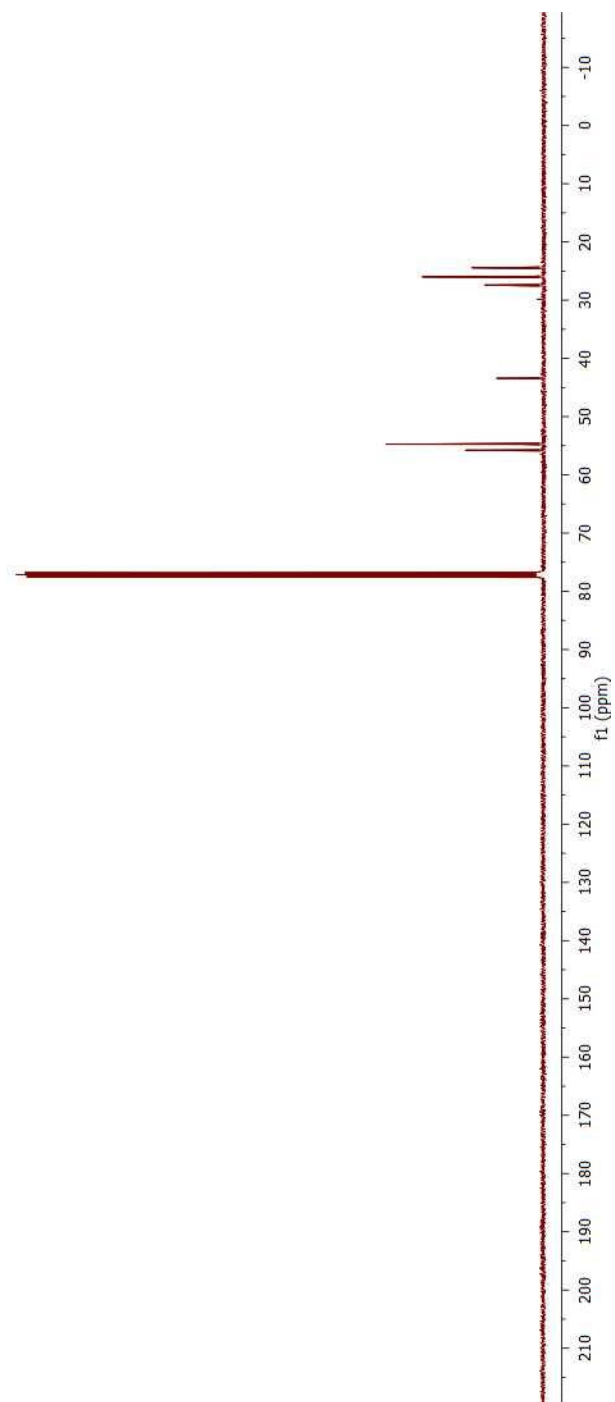
^{13}C NMR of compound **AM4**,
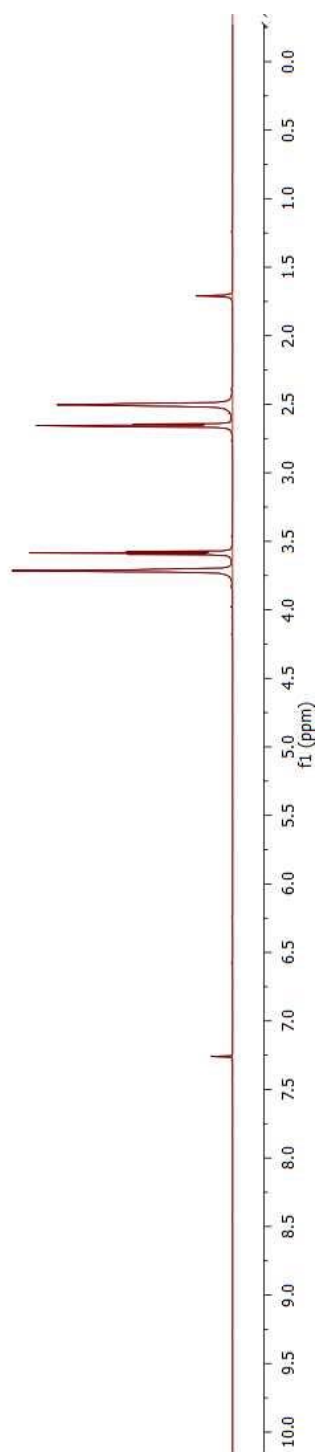
^{1}H NMR of compound **AM5**

^{13}C NMR of compound **AM5**,

ss
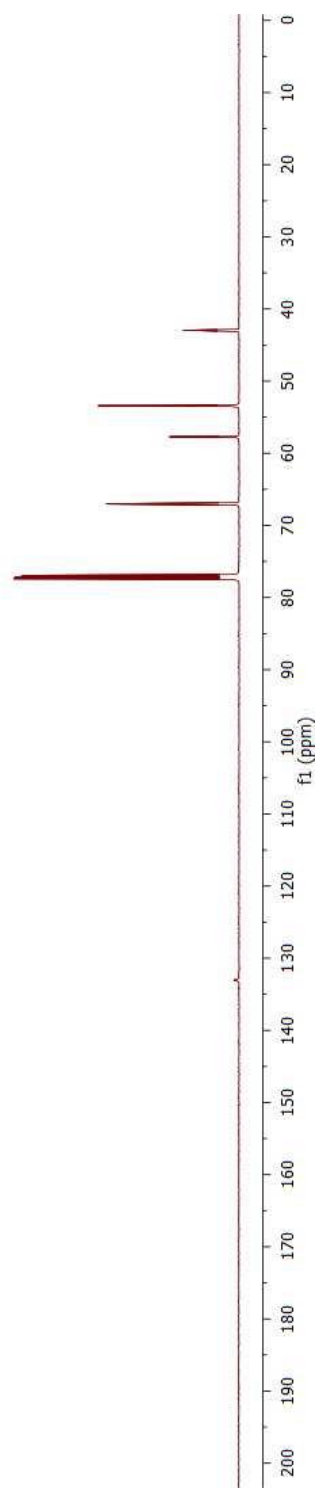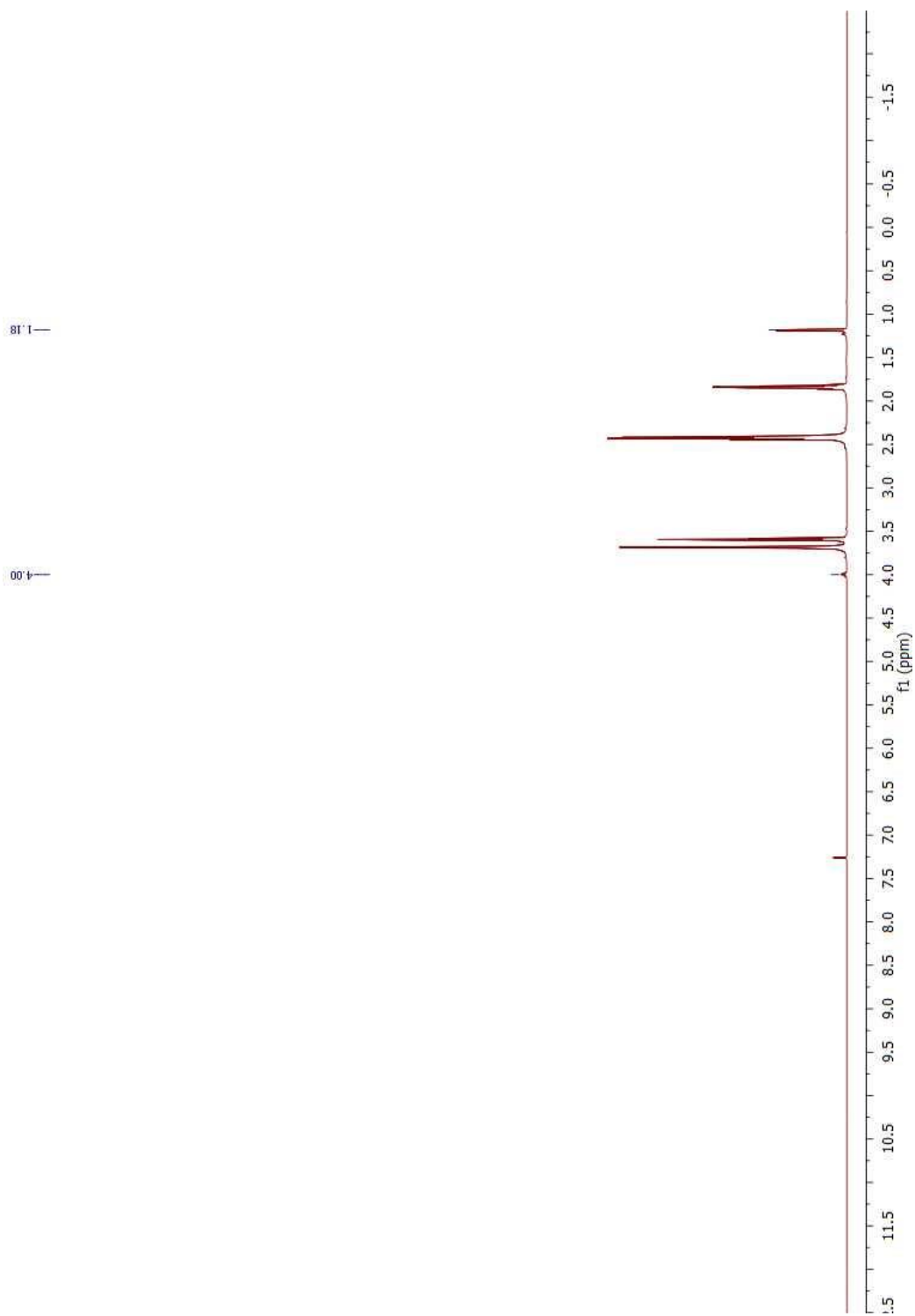
^{1}H NMR of compound **AM6,**
traces of isopropanol is labelled
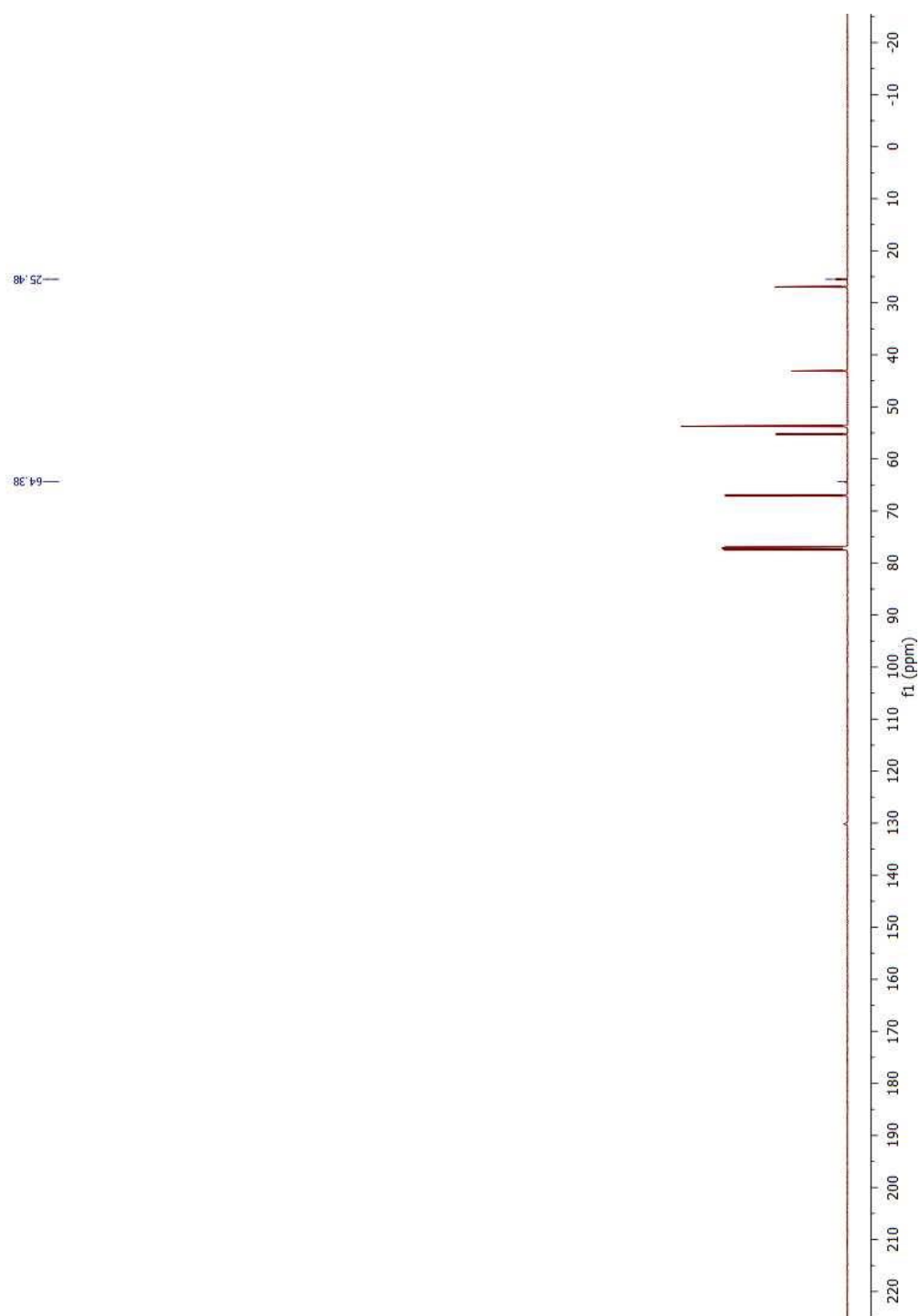
^{13}C NMR of compound **AM6**, traces of isopropanol is labelled.
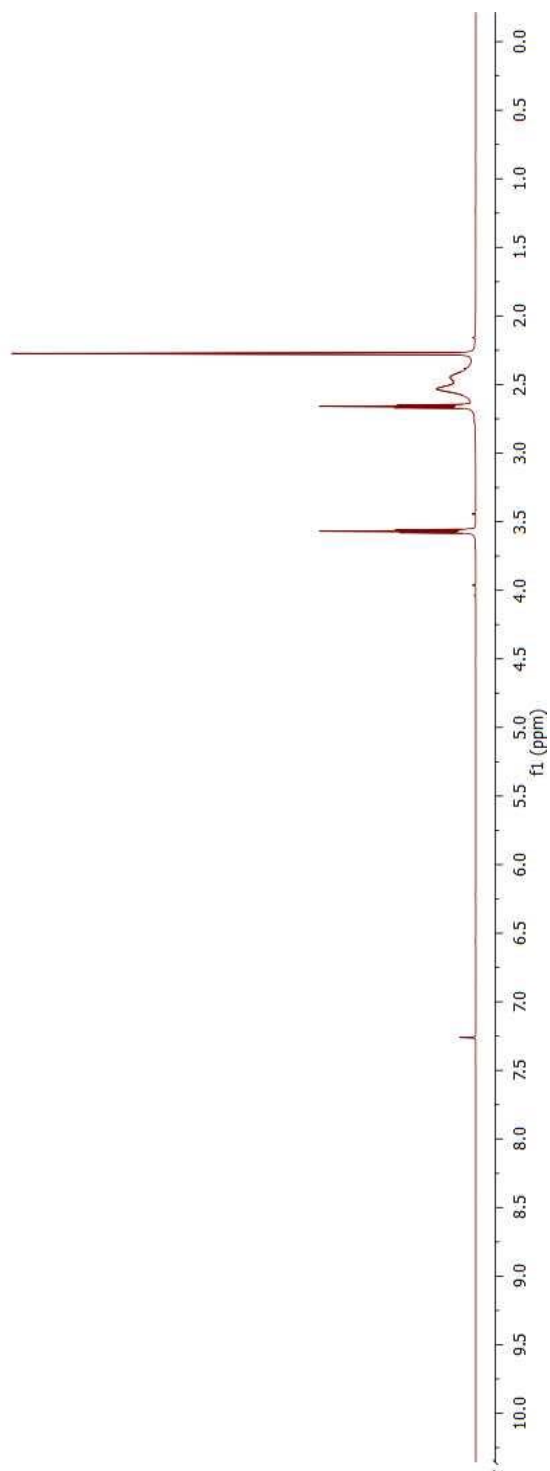
^{1}H NMR of compound **AM7**
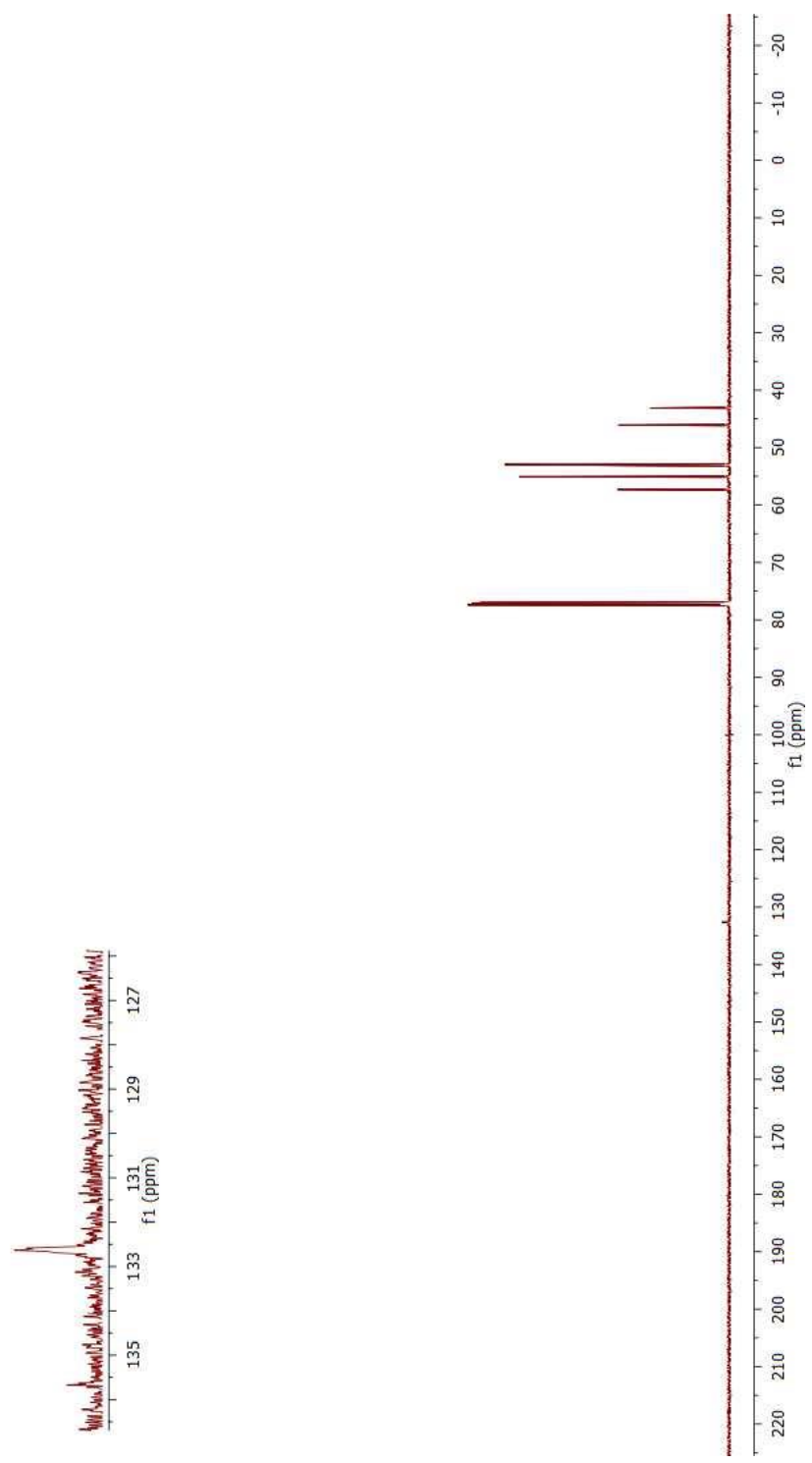
^{13}C NMR of compound **AM7**,
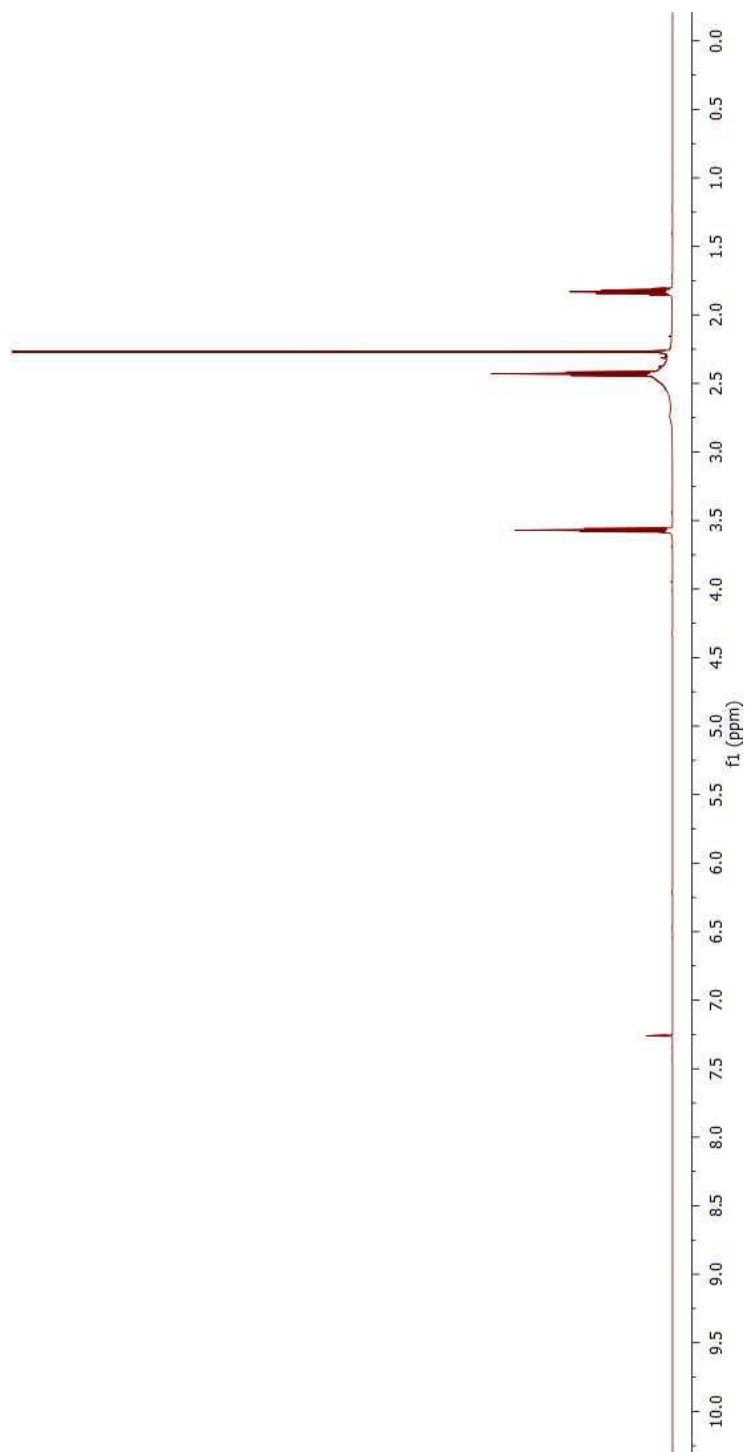
^{1}H NMR of compound **AM8**
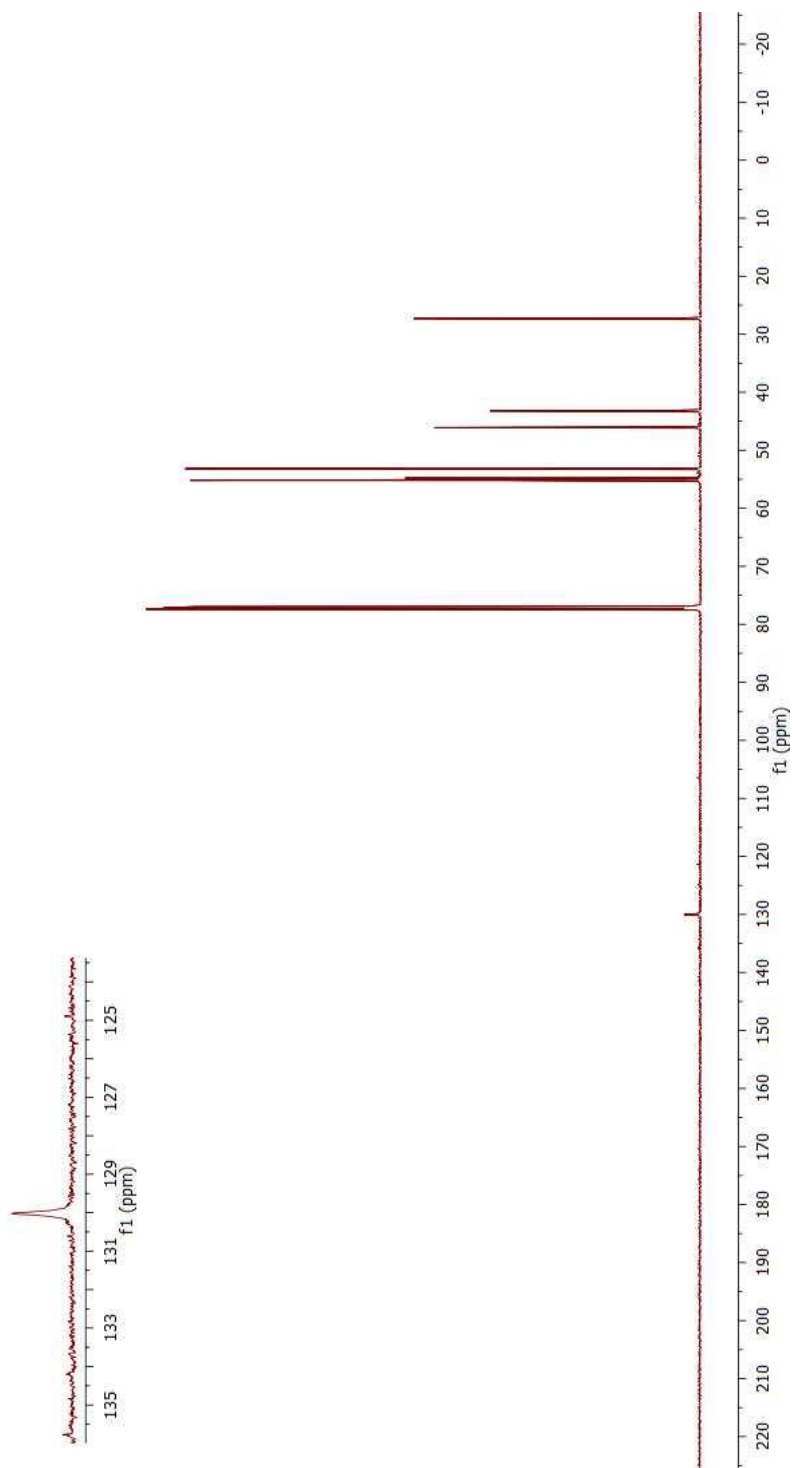
^{13}C NMR of compound **AM8**
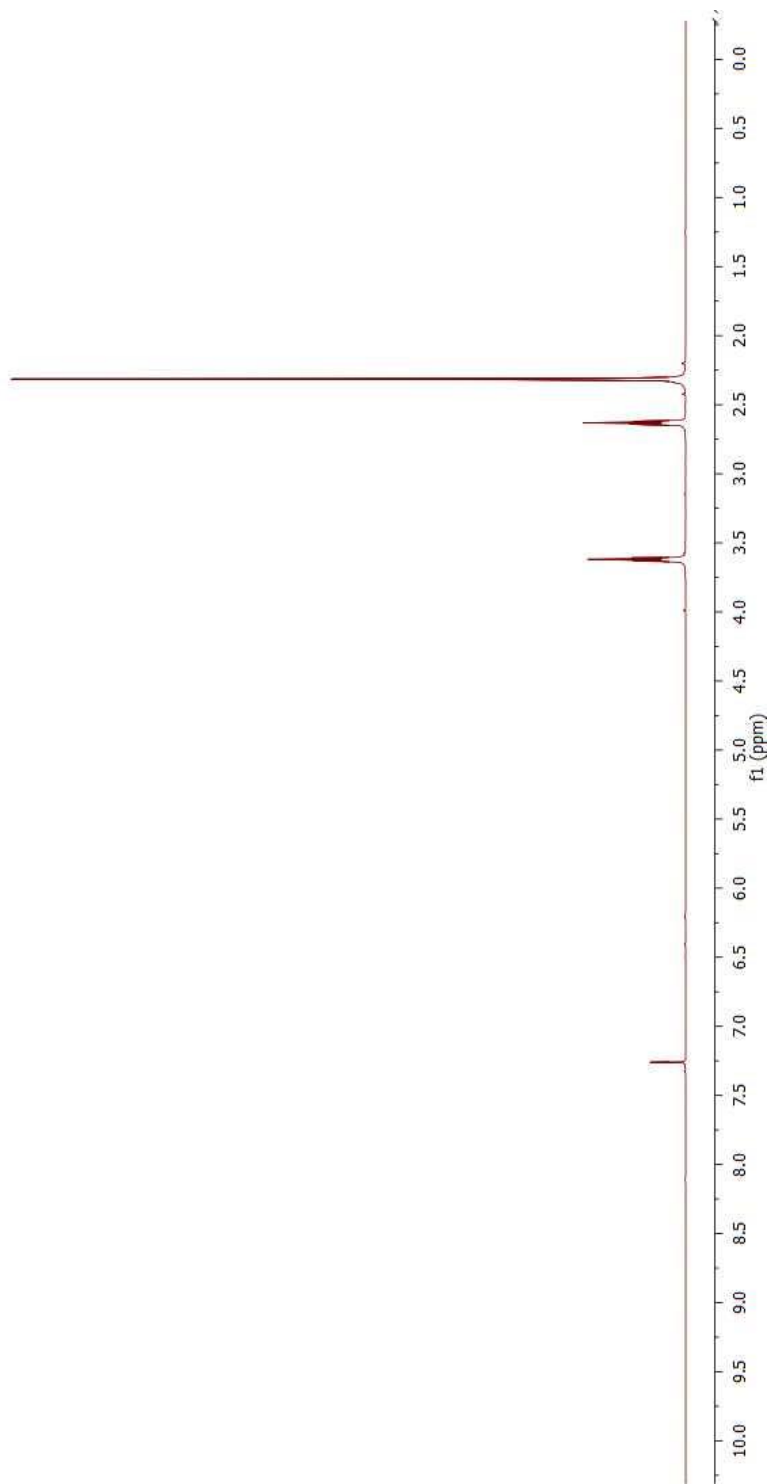
^{1}H NMR of compound **AM9**
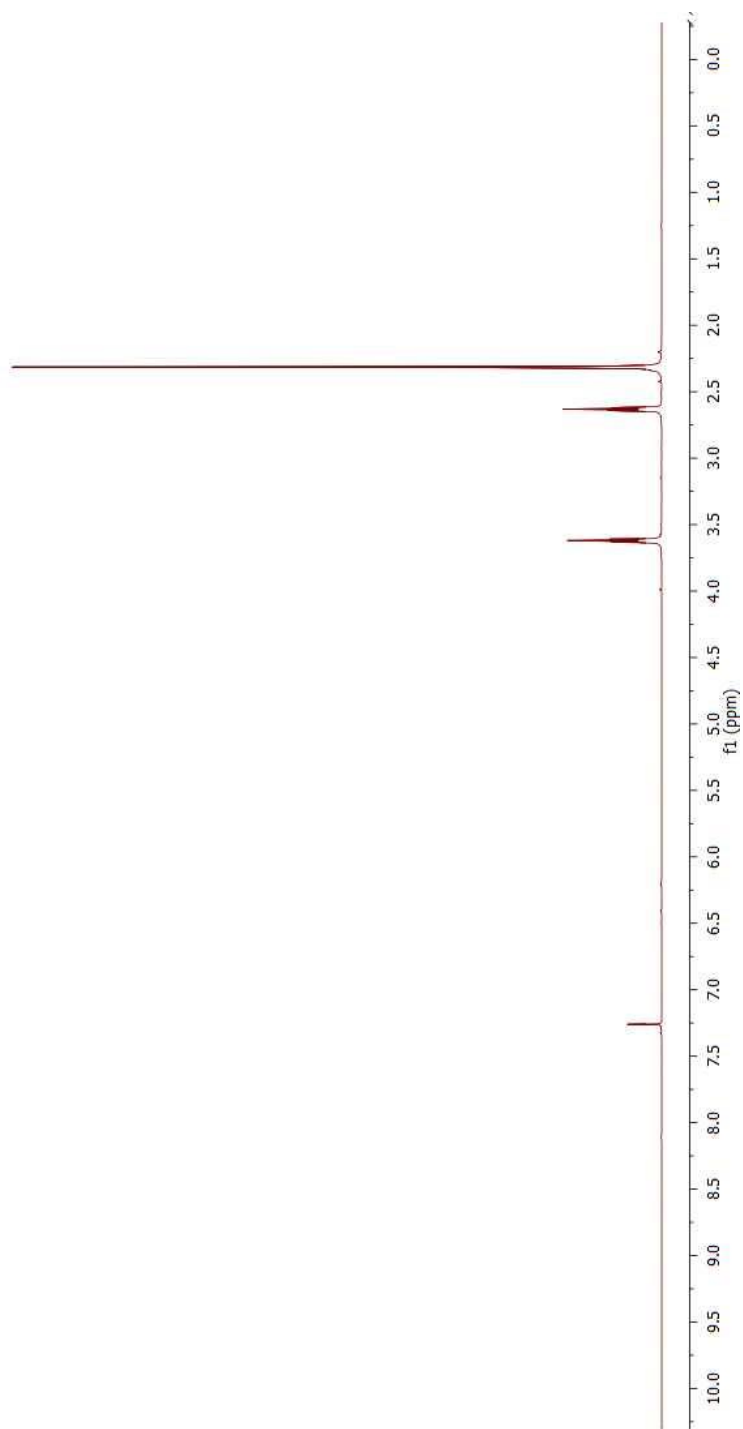
^{13}C NMR of compound **AM9**
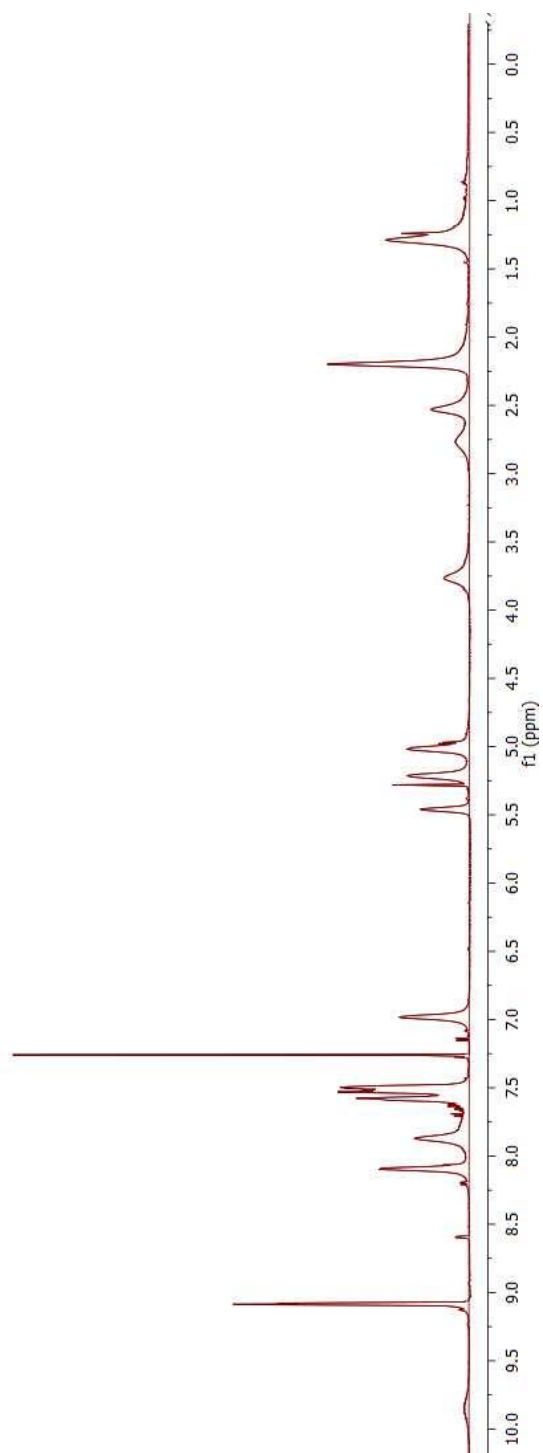
^{1}H NMR of compound **10**
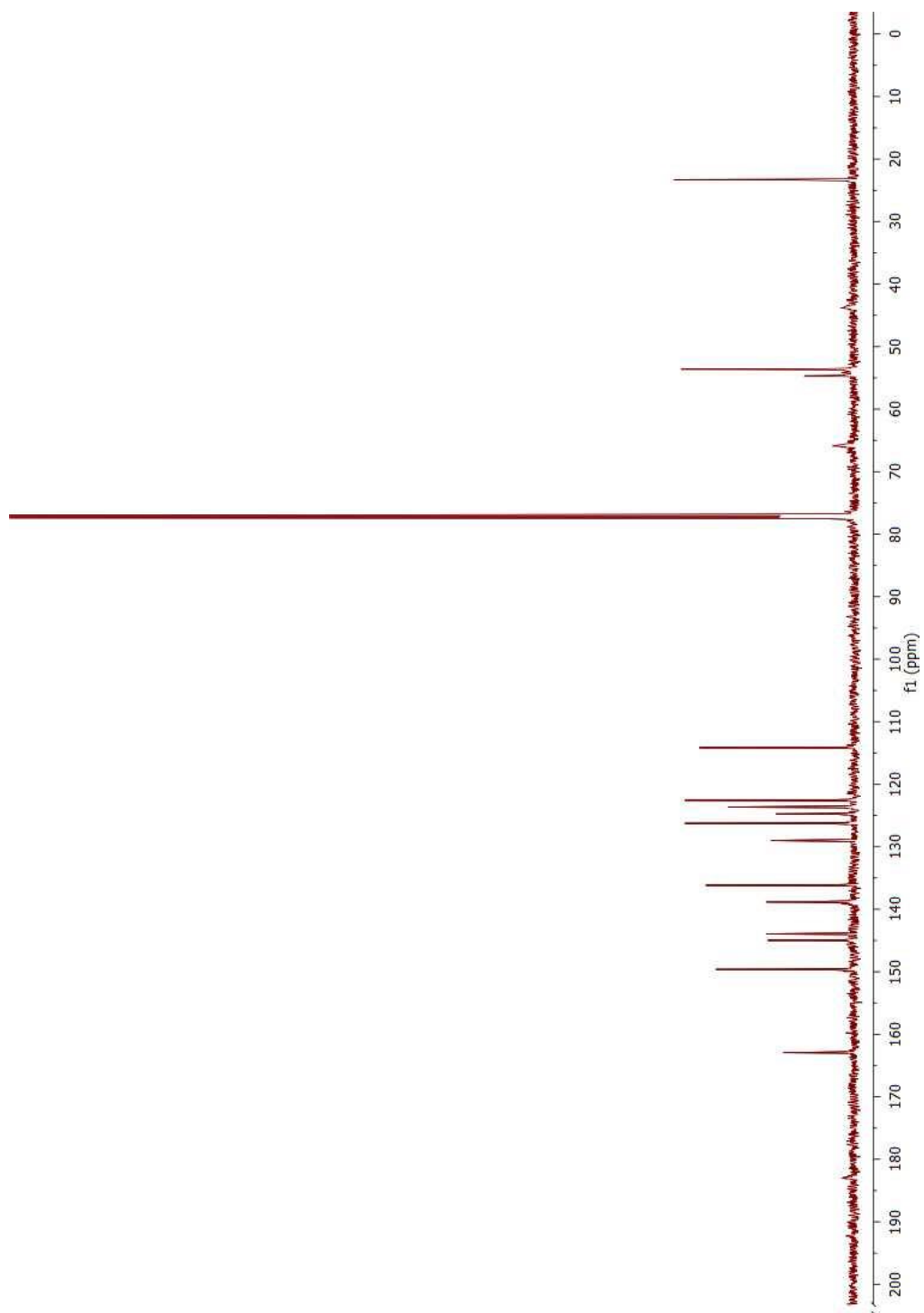
^{13}C NMR of compound **10**
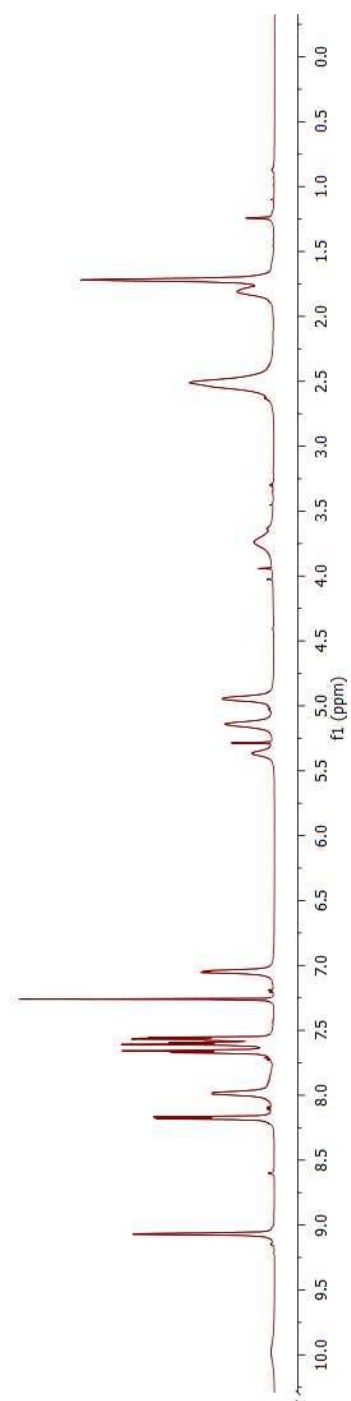
^{1}H NMR of compound **11**
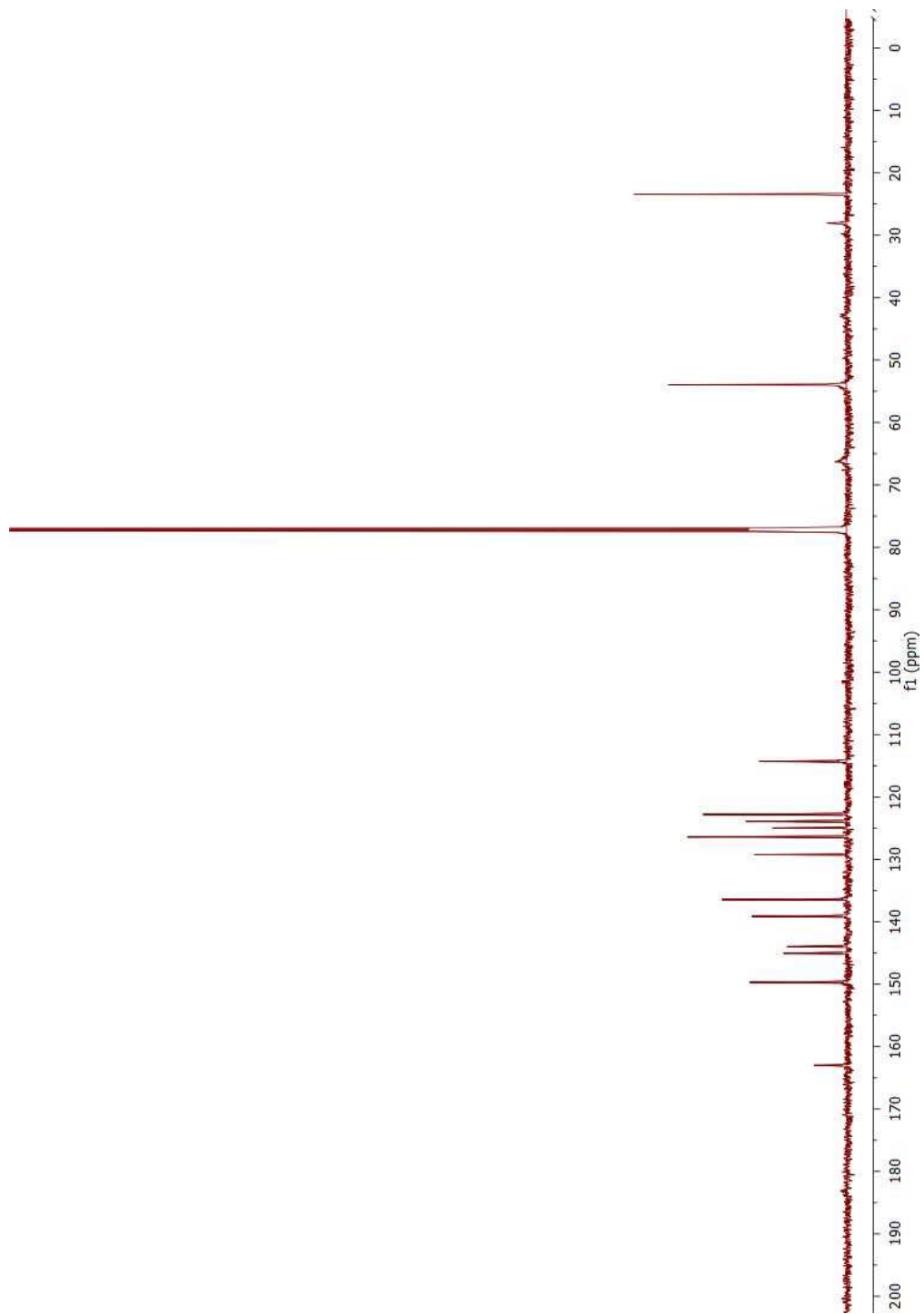
^{13}C NMR of compound **11**
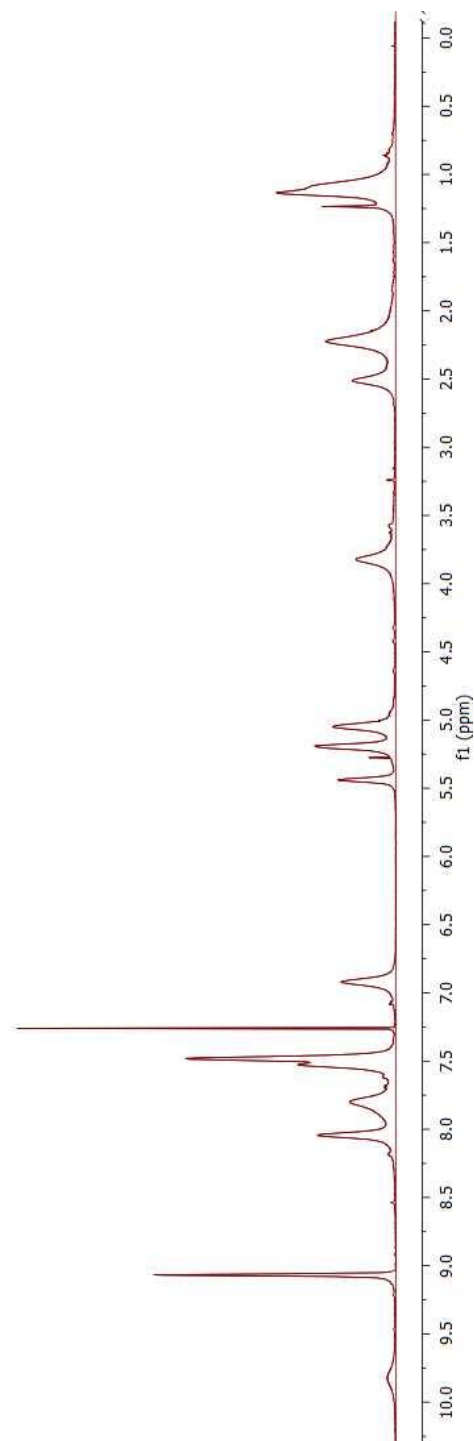
¹H NMR of compound **12**
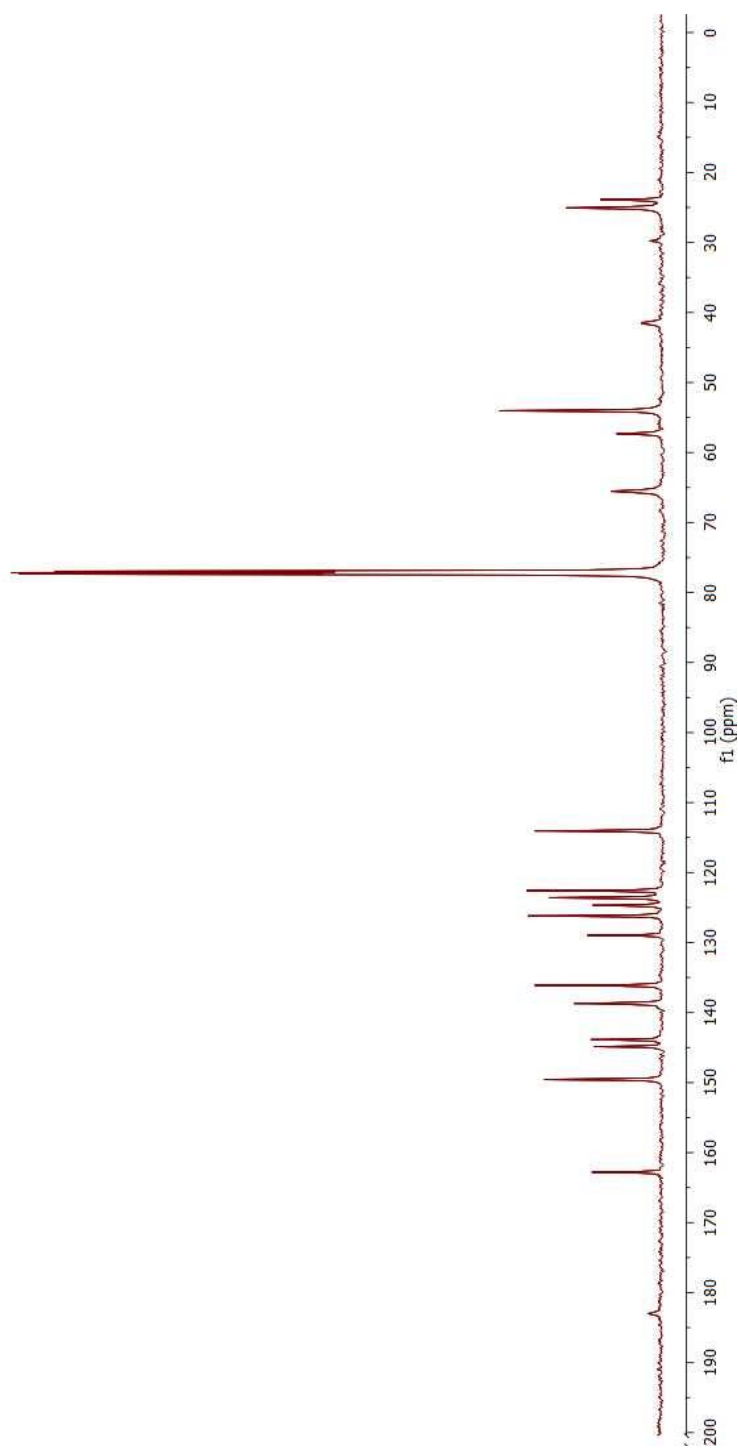
¹³C NMR of compound **12**
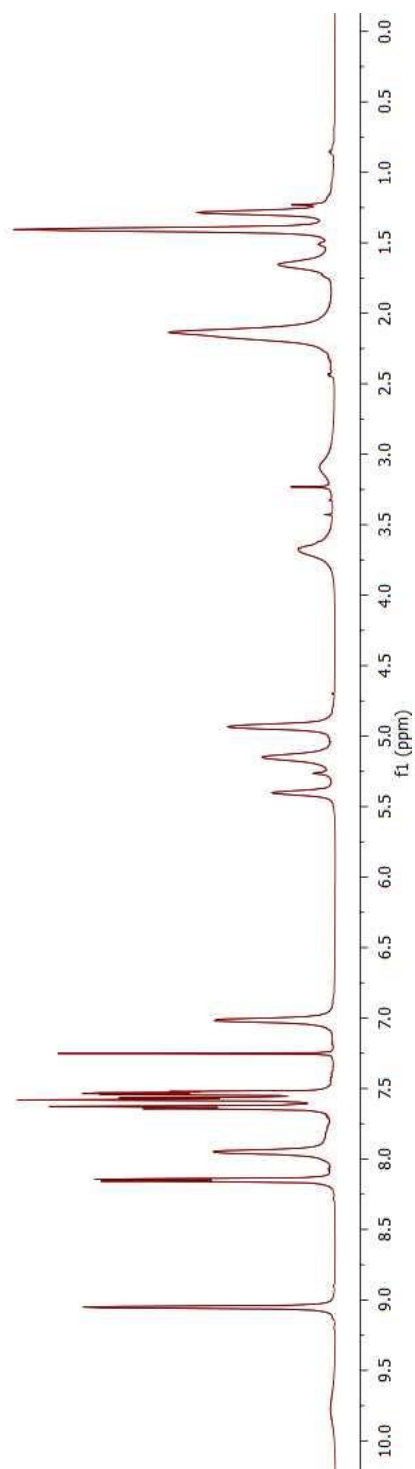
^{1}H NMR of compound **13**
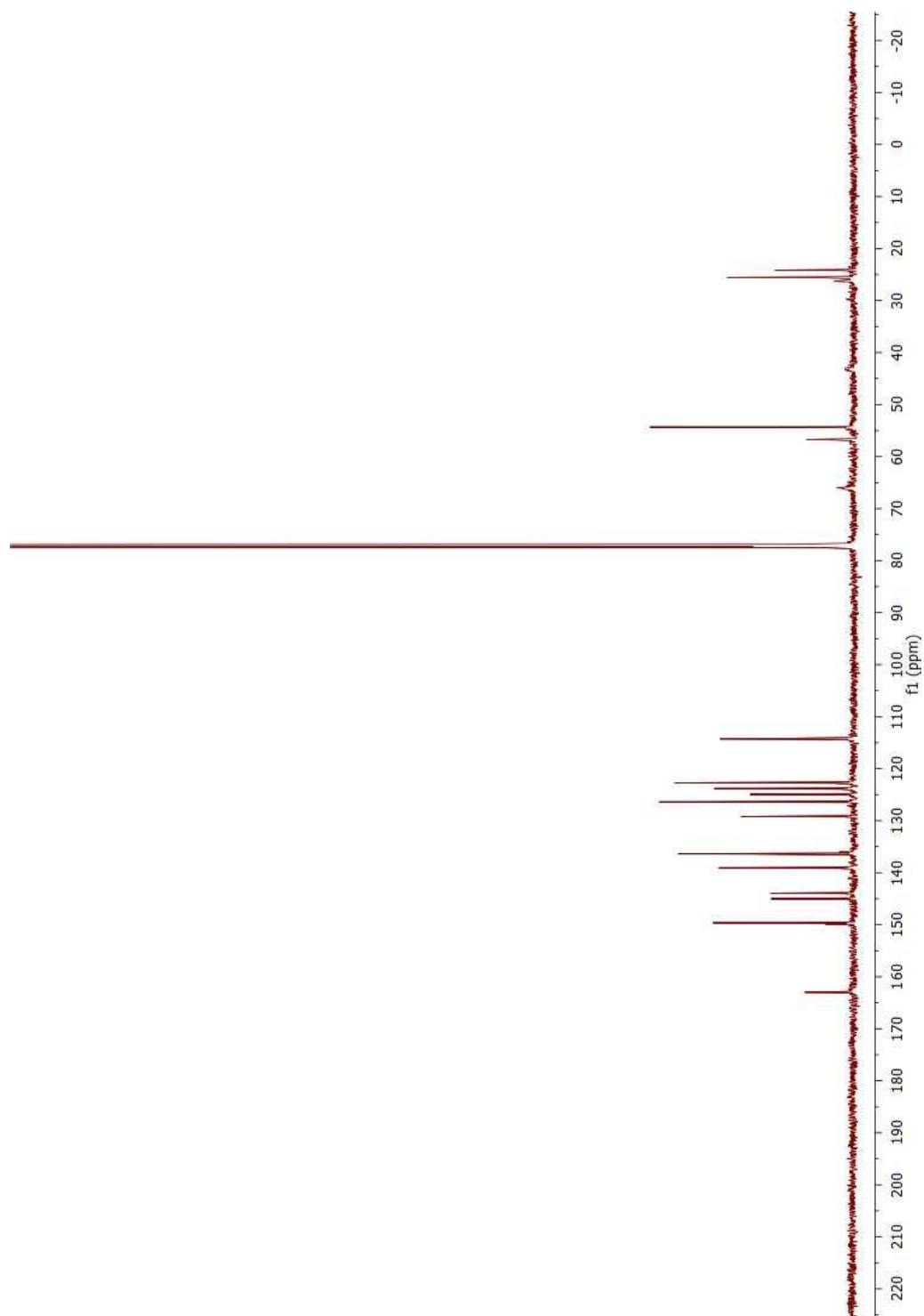
^{13}C NMR of compound **13**
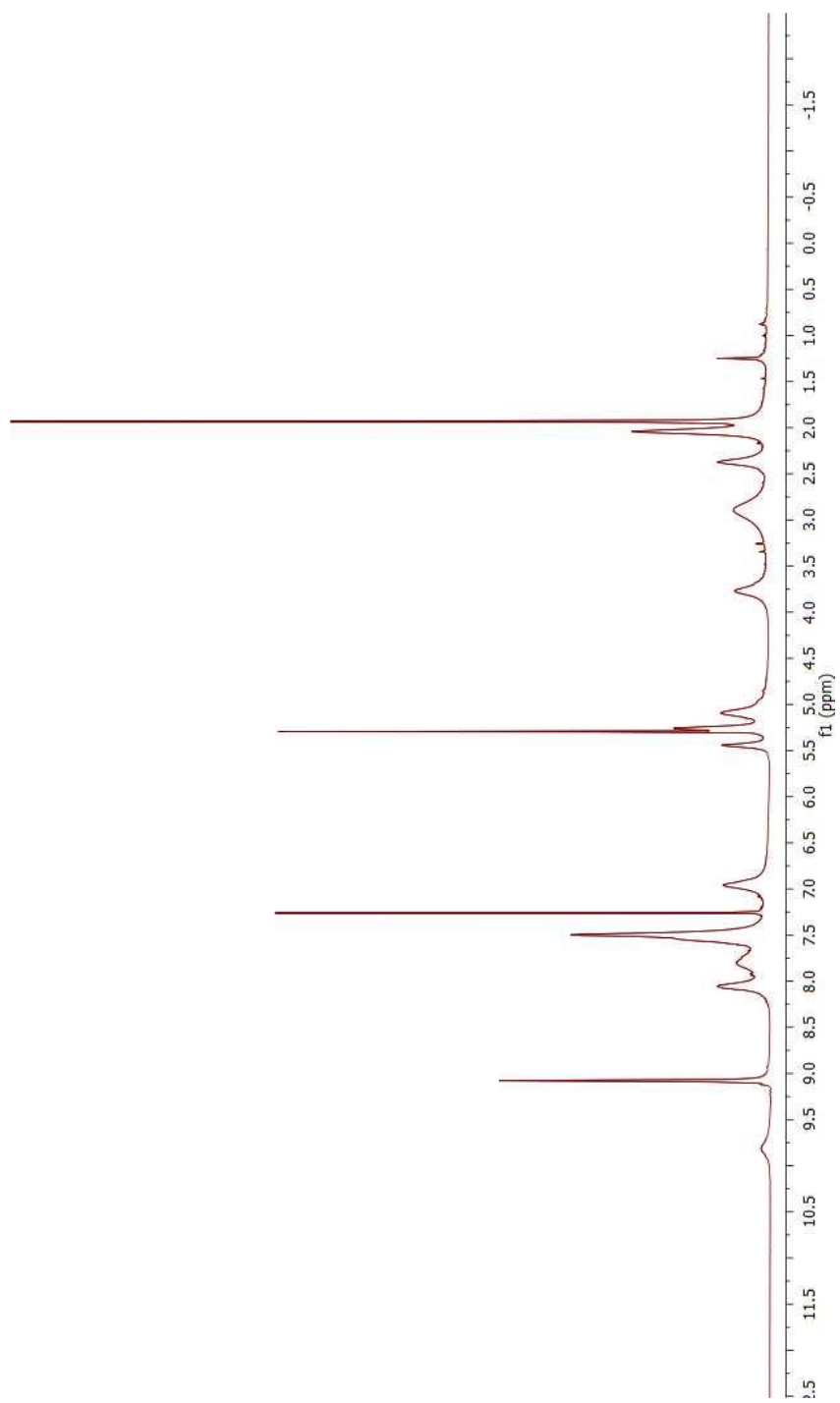
^{1}H NMR of compound **14**
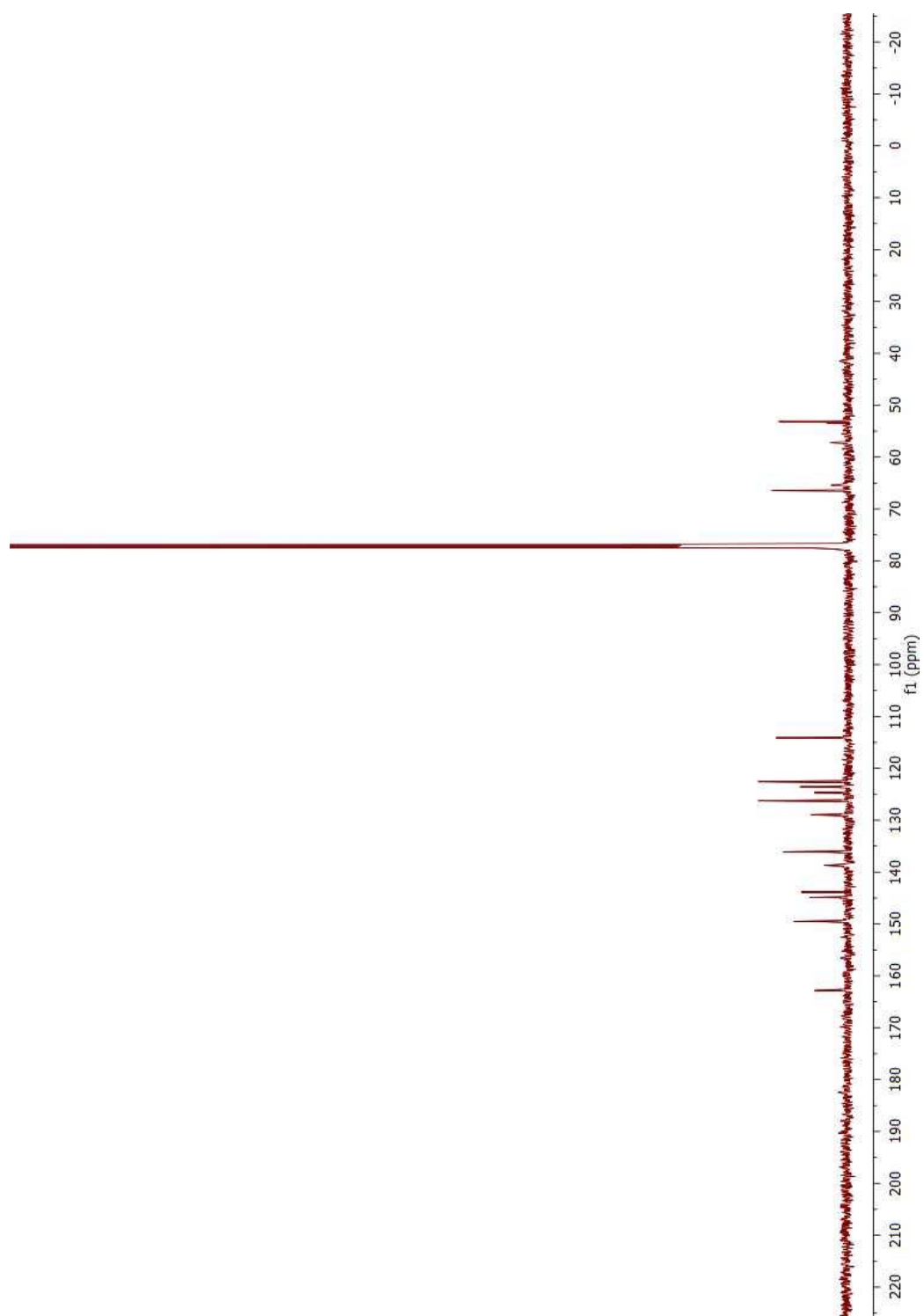
^{13}C NMR of compound **14**
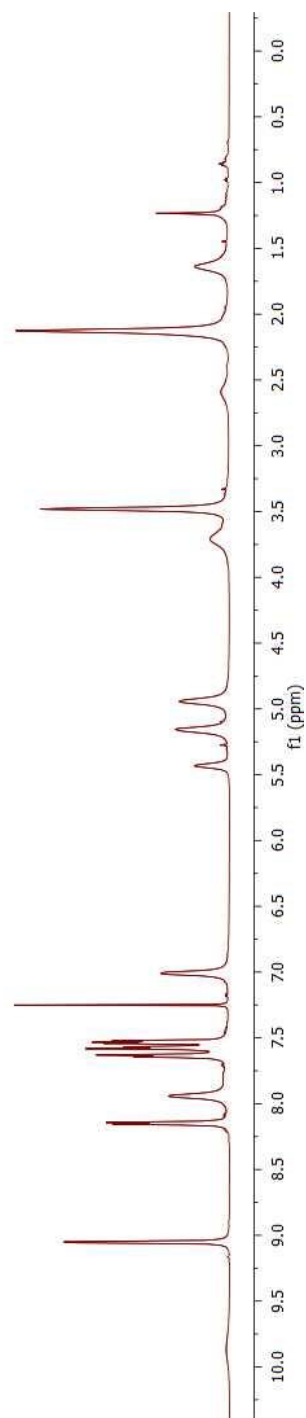
^{1}H NMR of compound **15**
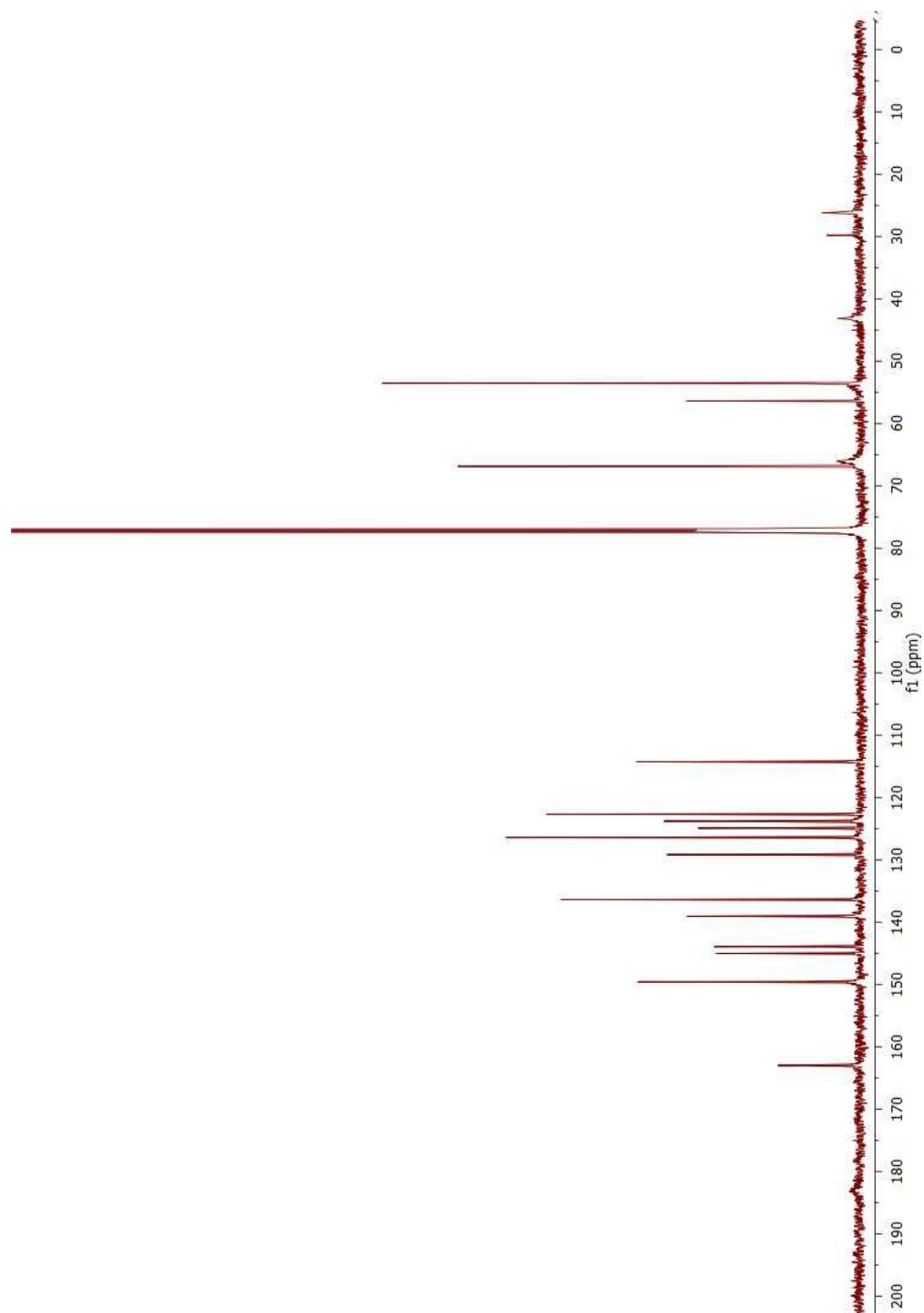
^{13}C NMR of compound **15**
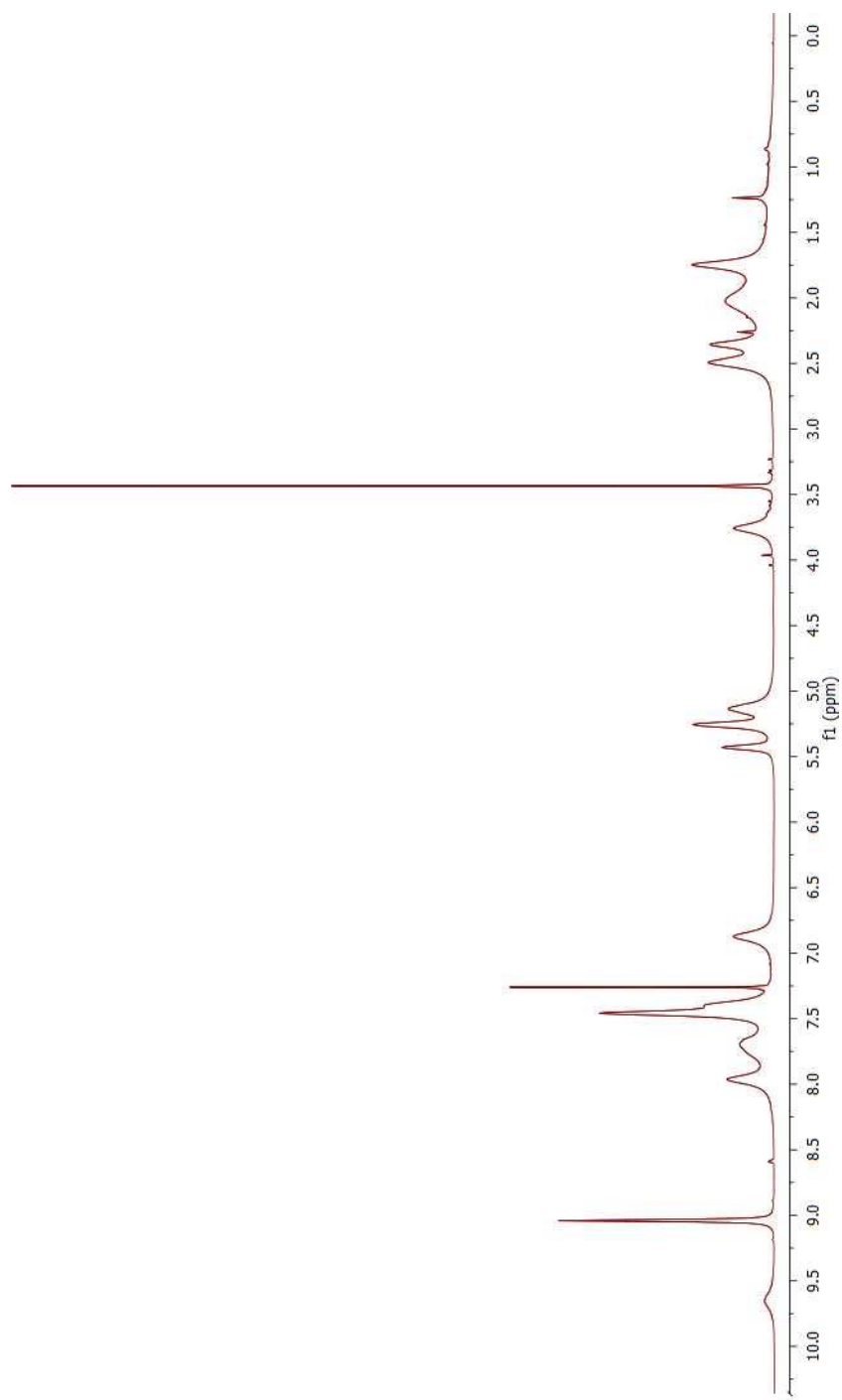
¹H NMR of compound **16**

^{13}C NMR of compound **16**

^{1}H NMR of compound **17**
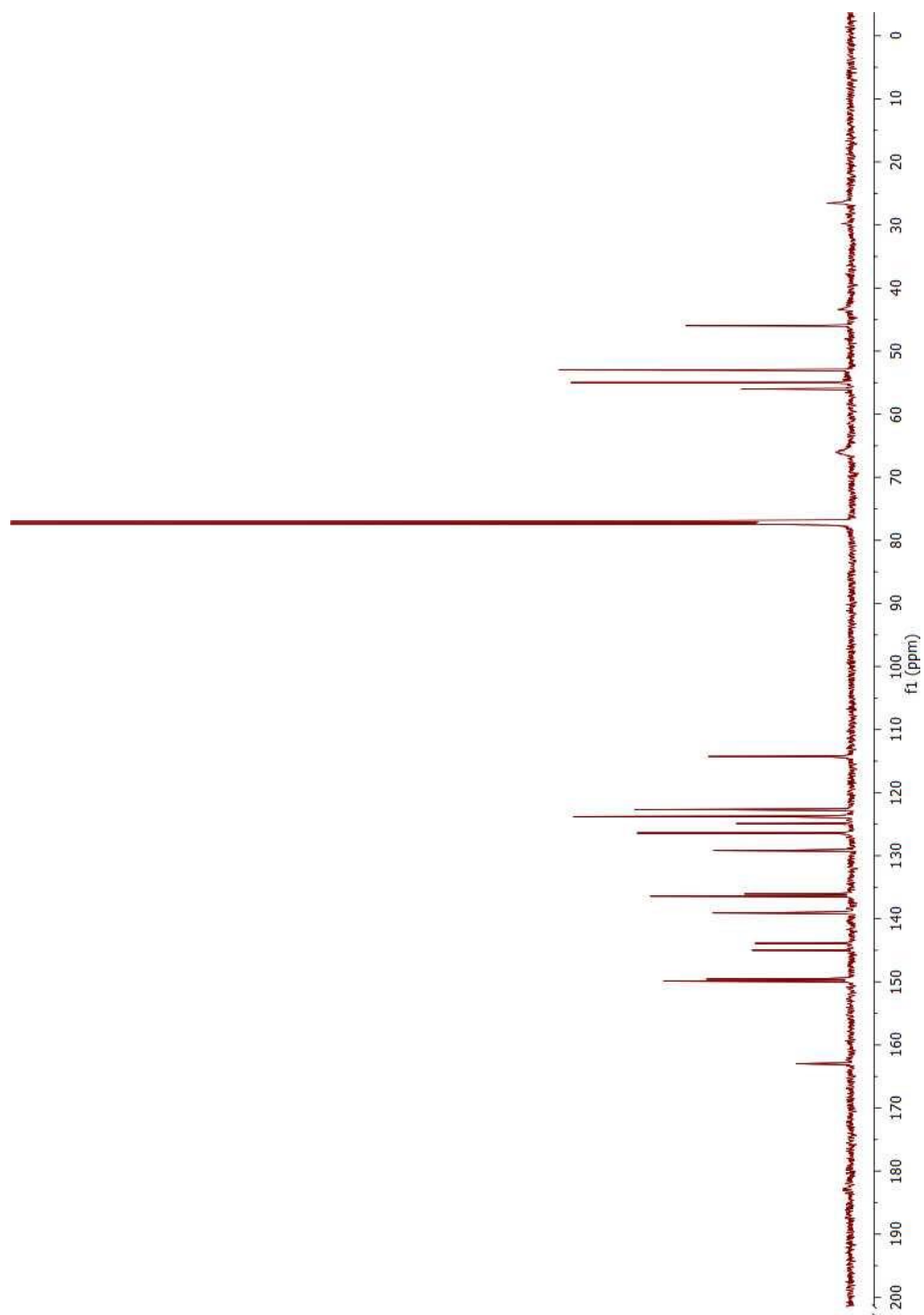
^{13}C NMR of compound **17**

^{1}H NMR of compound **18**

APPENDIX B: IR SPECTRA OF THE SELECTED COMPOUNDS
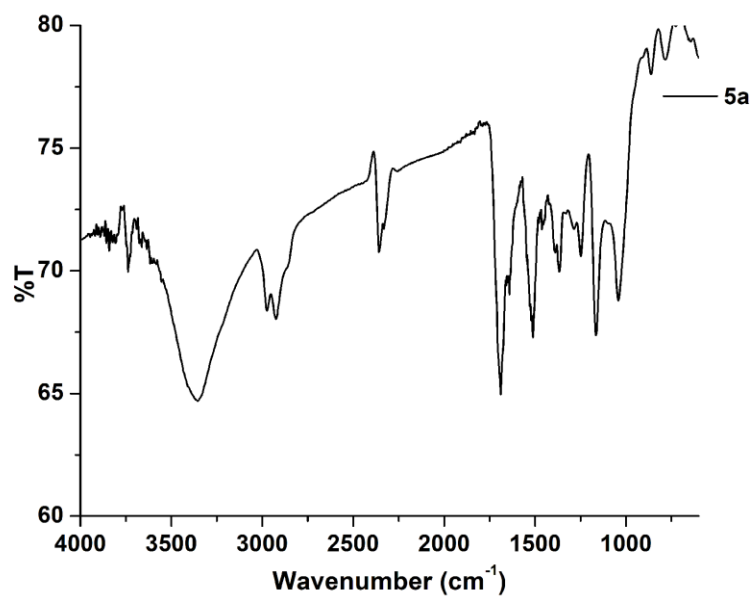
IR Spectrum of **5a**
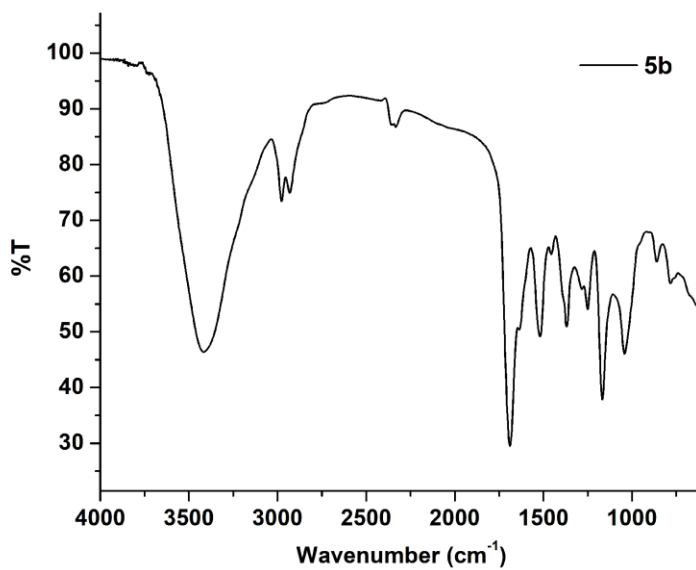
IR Spectrum of **5b**
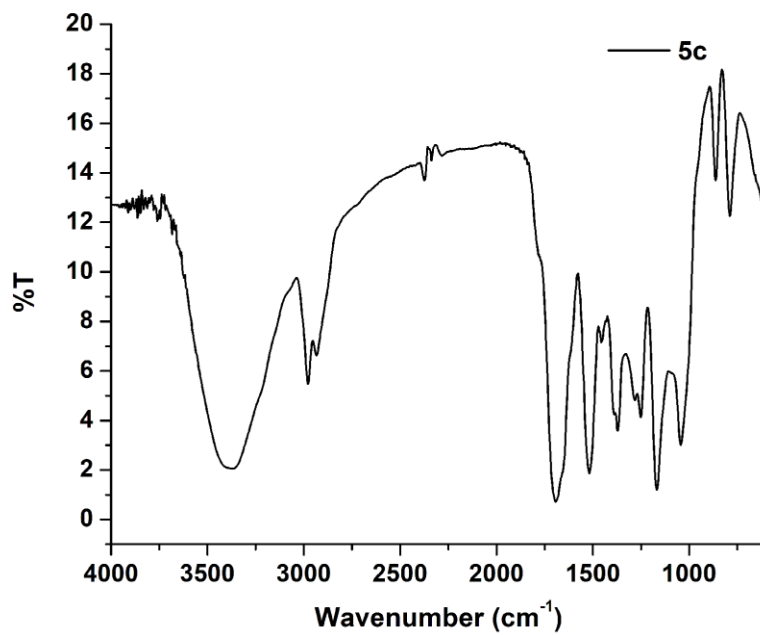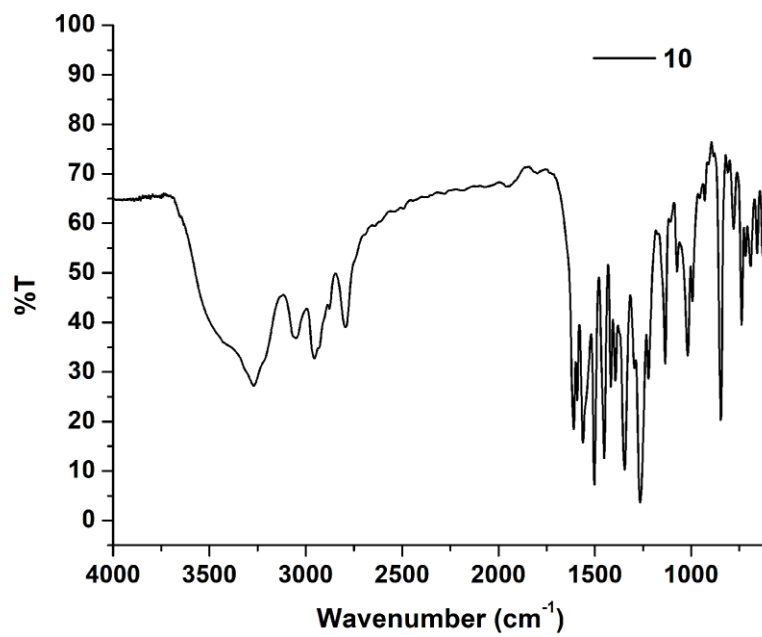
IR Spectrum of **5c**
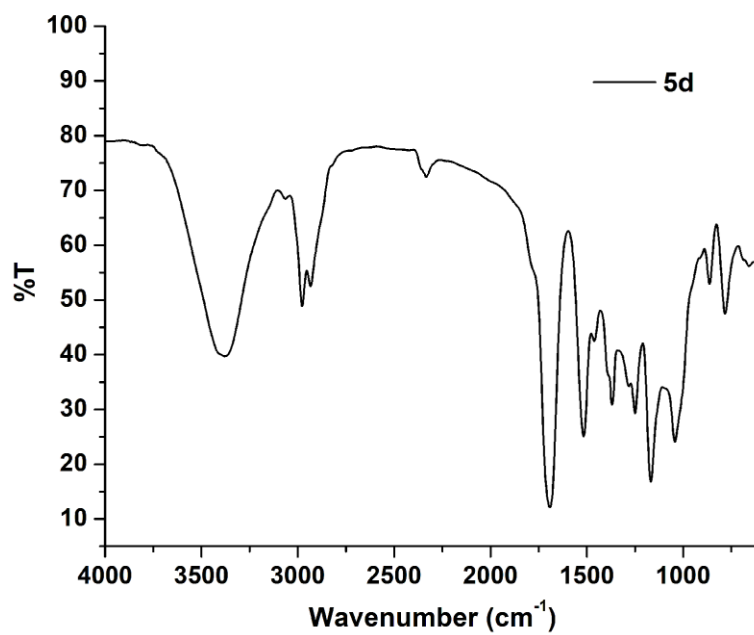
IR Spectrum of **5d**
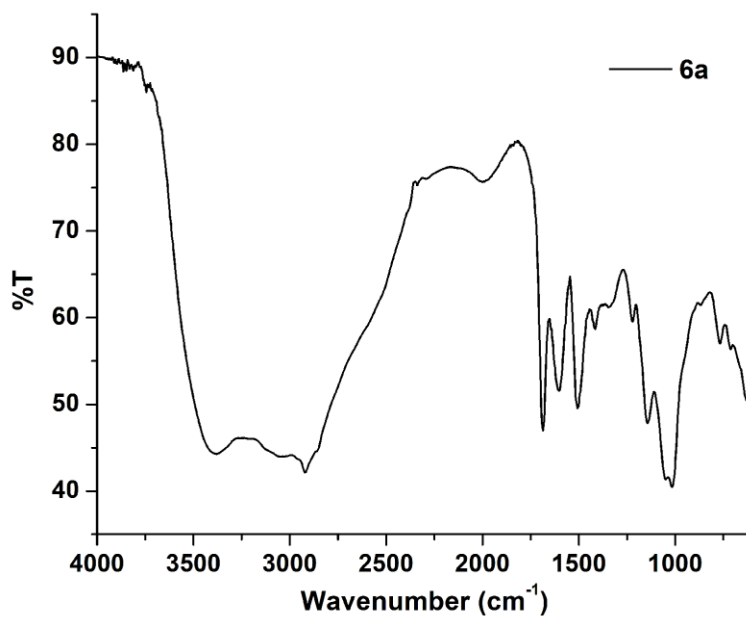
IR Spectrum of **6a**
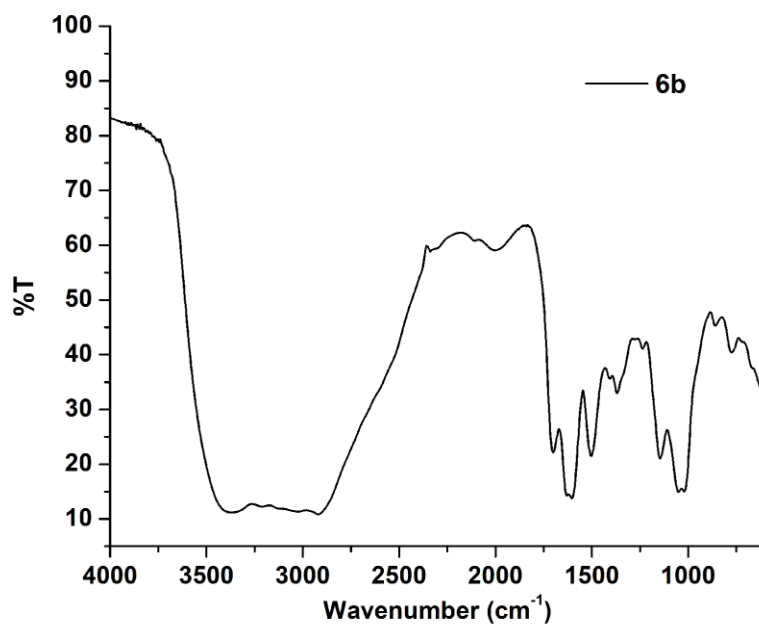
IR Spectrum of **6b**
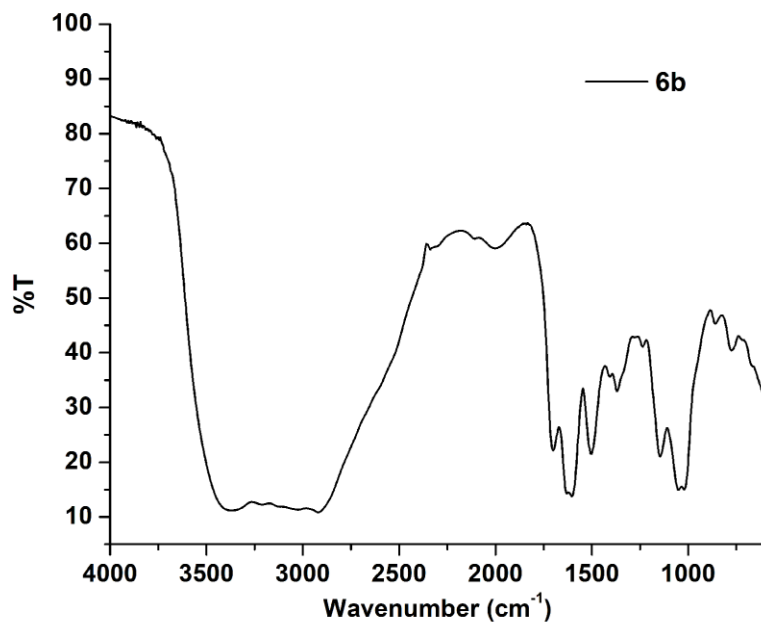
IR Spectrum of **6c**
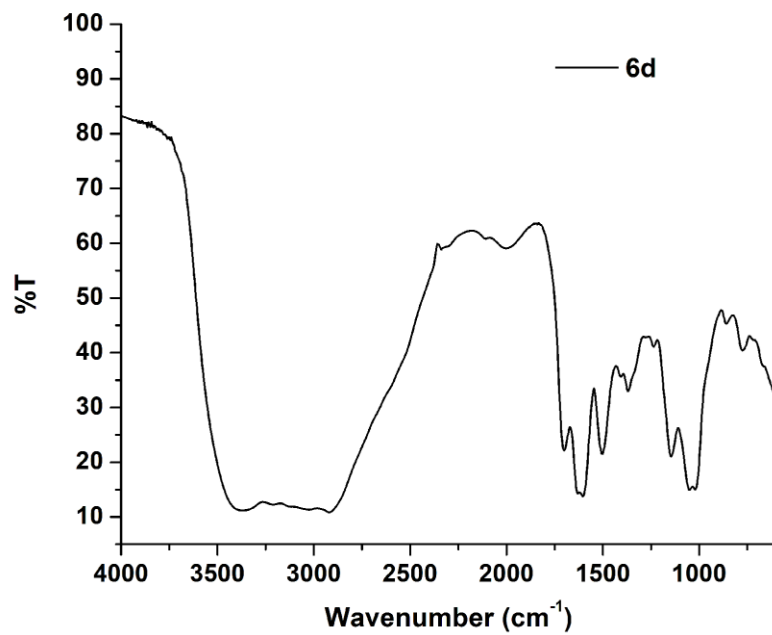
IR Spectrum of **6d**
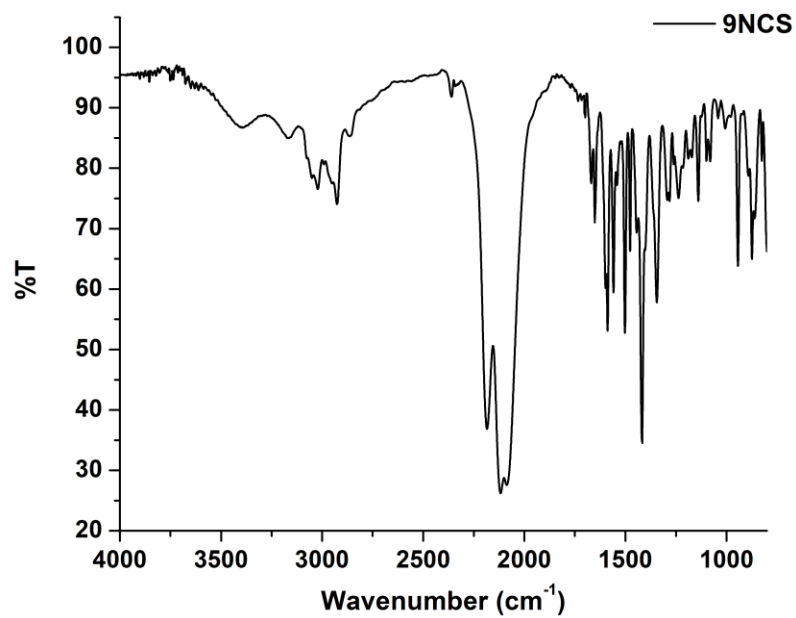
IR Spectrum of **9NCS**
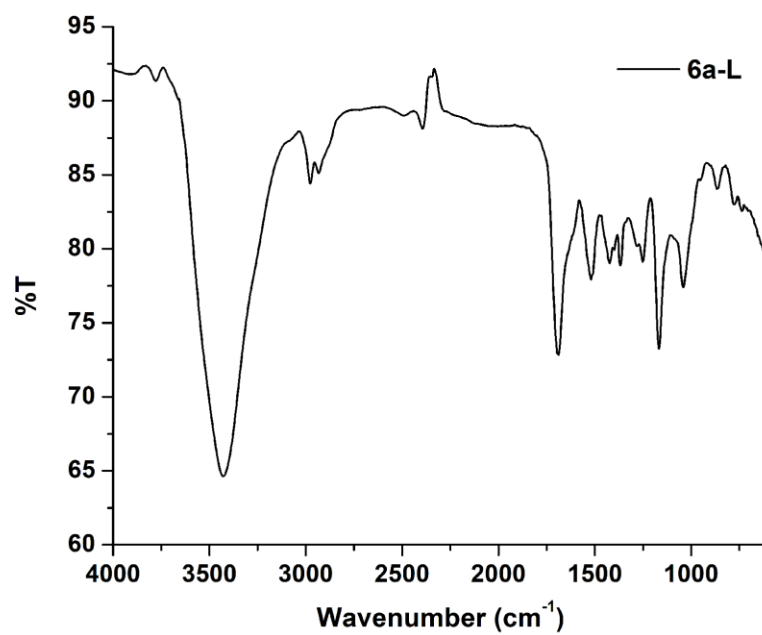
IR Spectrum of **6a-L**
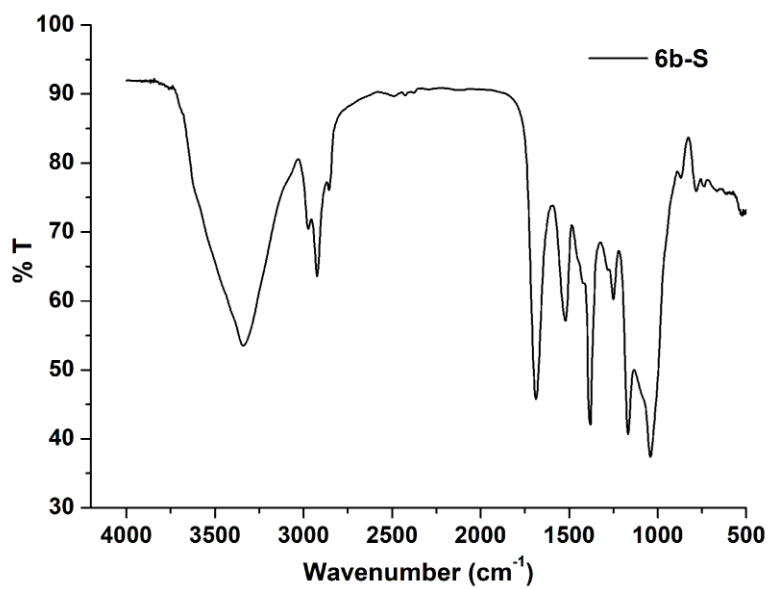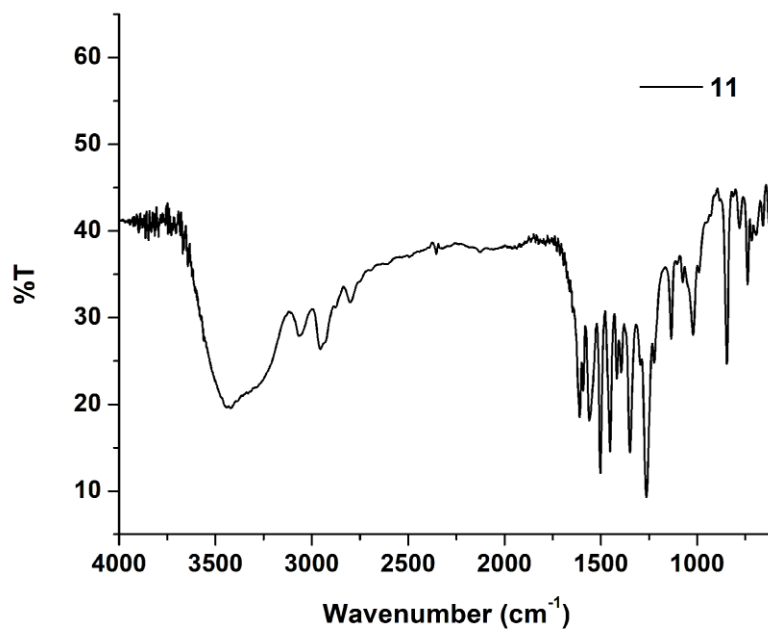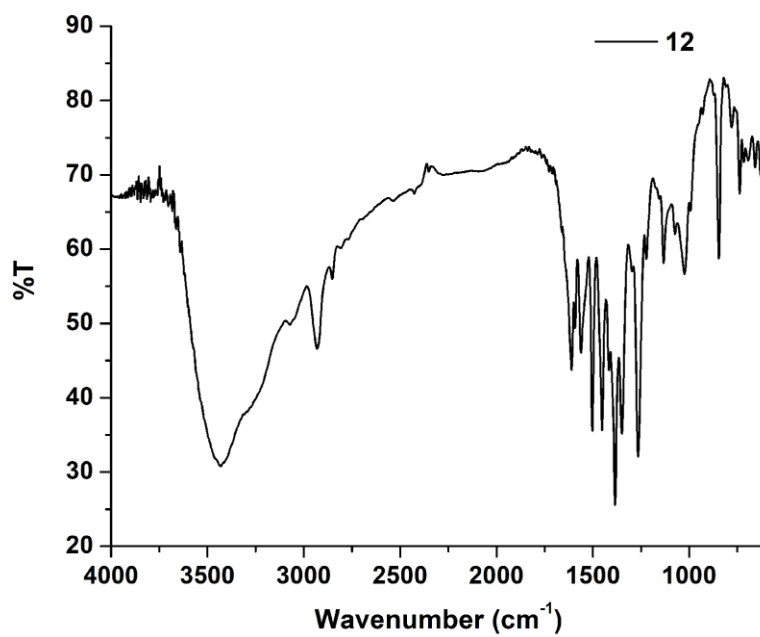
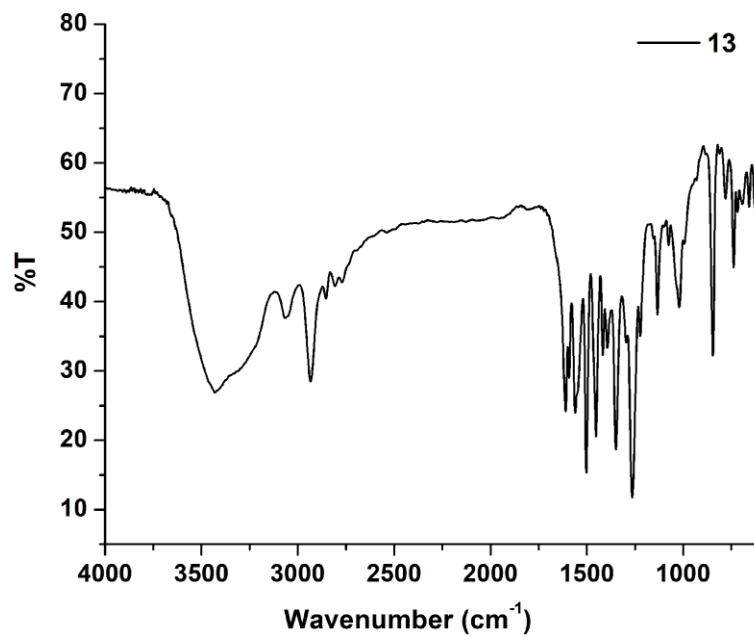IR Spectrum of **6b-S**
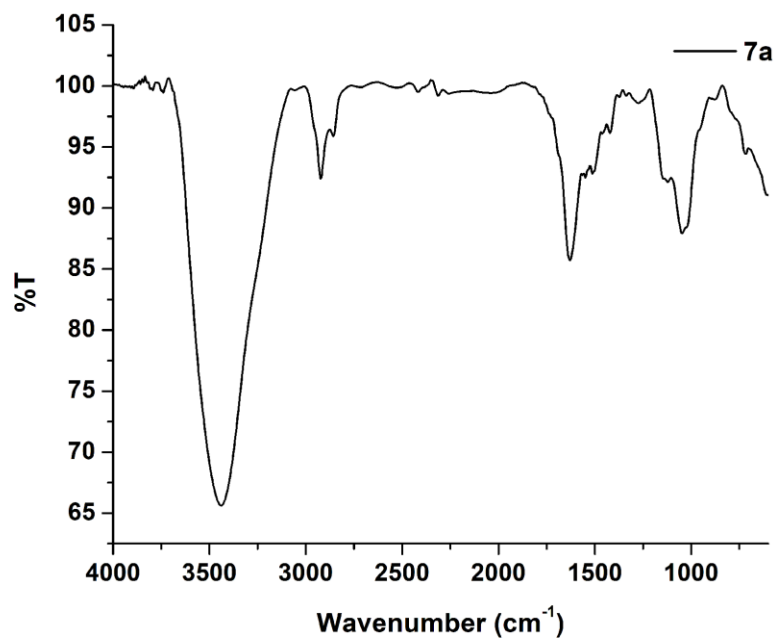
IR Spectrum of **7a**
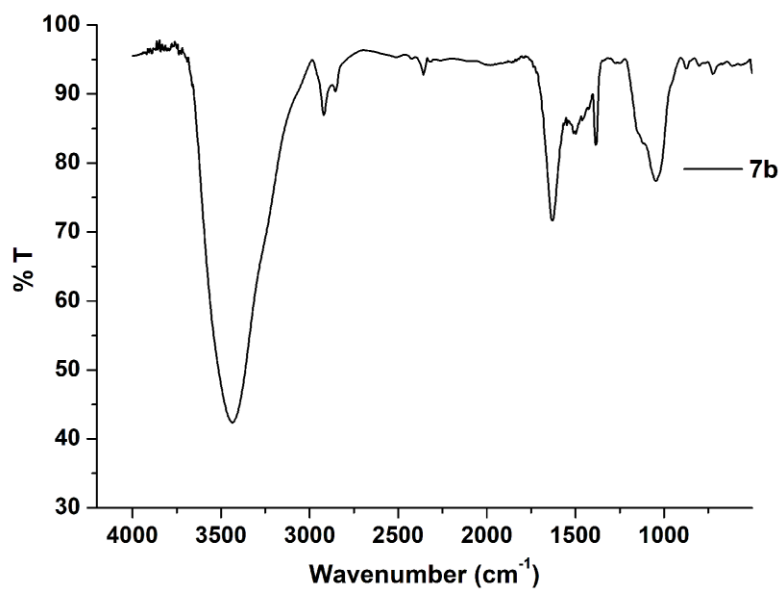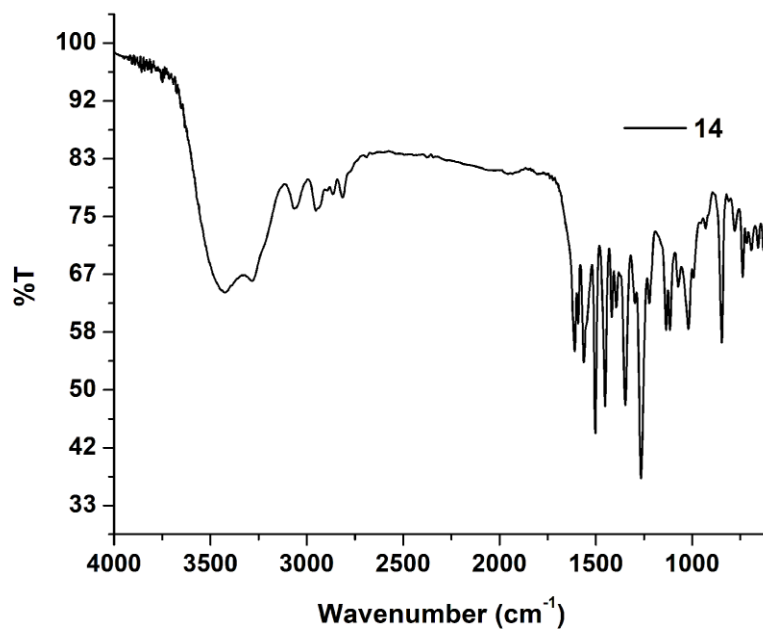
IR Spectrum of **7b**
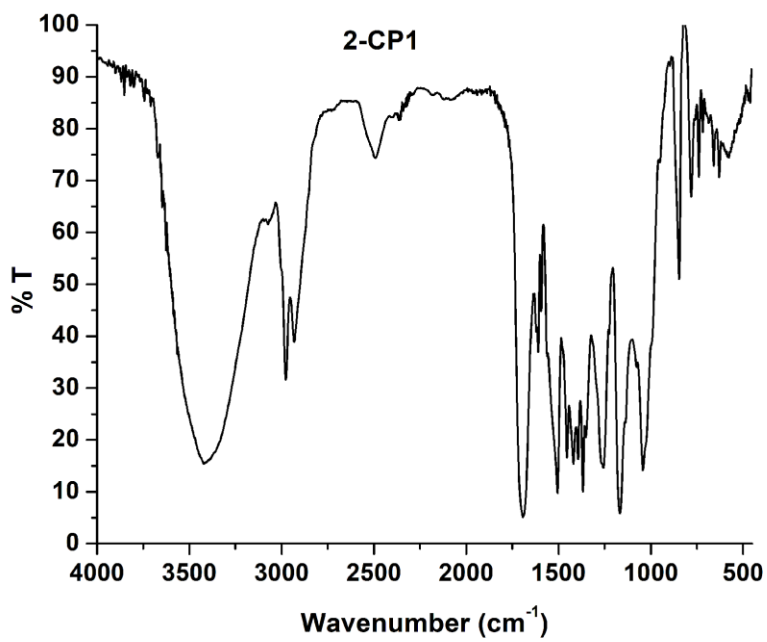
IR spectrum of **2-CP1**
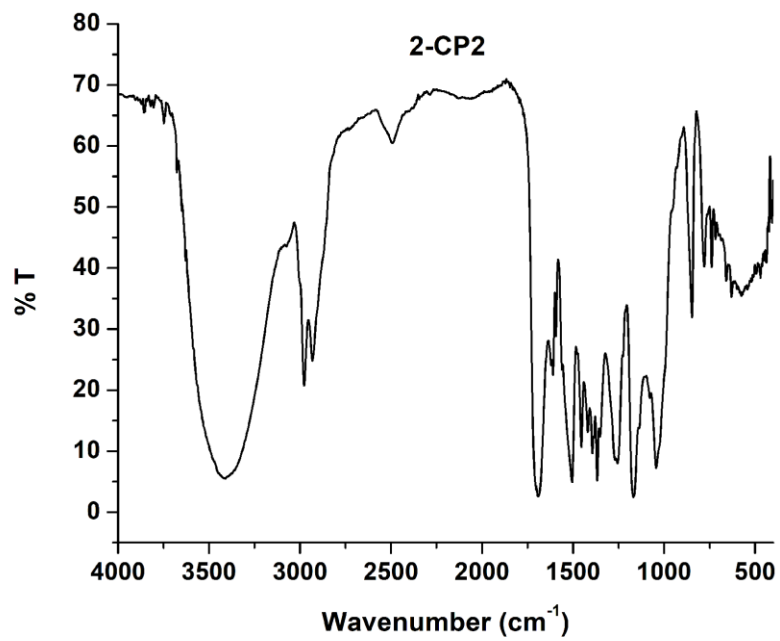
IR spectrum of **2-CP2**.
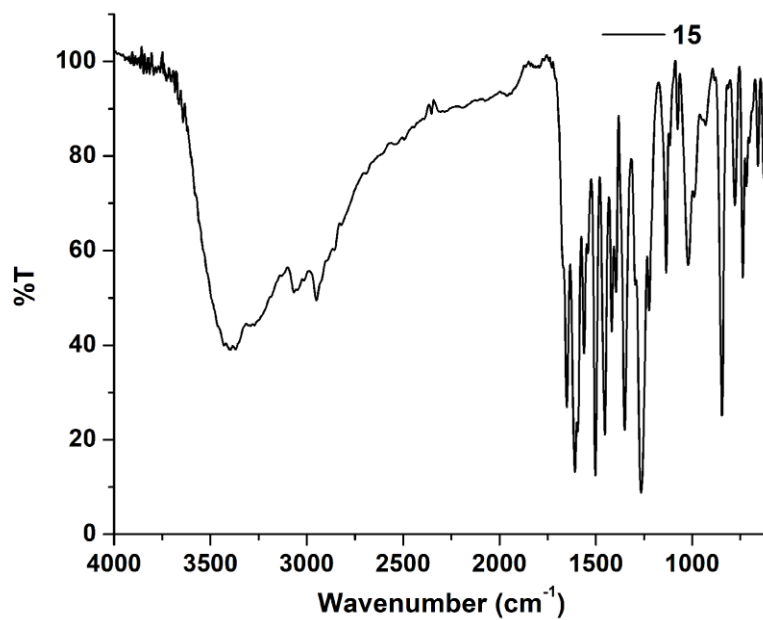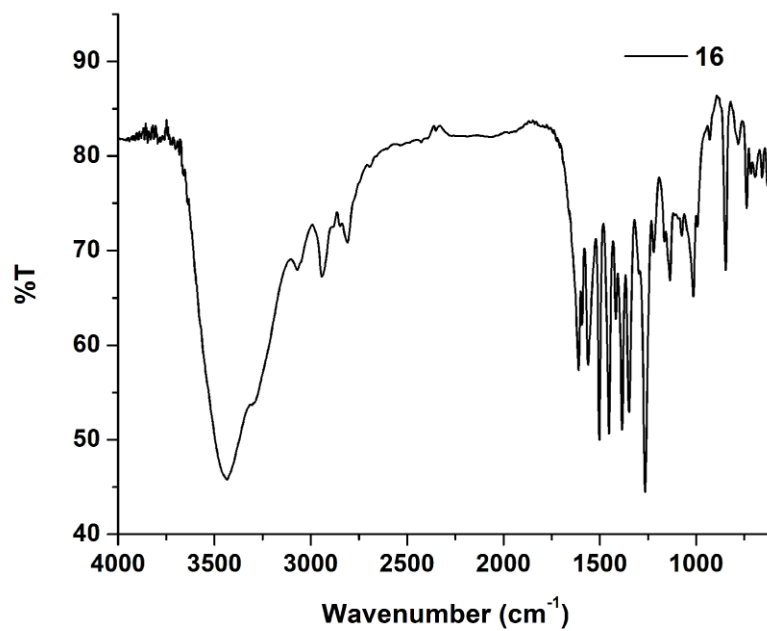

IR spectrum of **10**

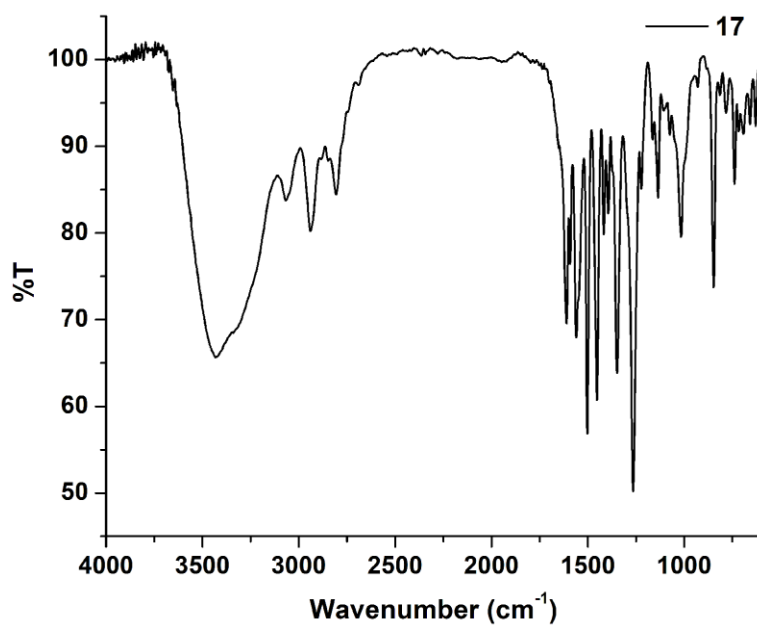IR spectrum of **11**

IR spectrum of **12**

IR spectrum of **13**

IR spectrum of **14**

IR spectrum of **15**

IR spectrum of **16**

IR spectrum of **17**
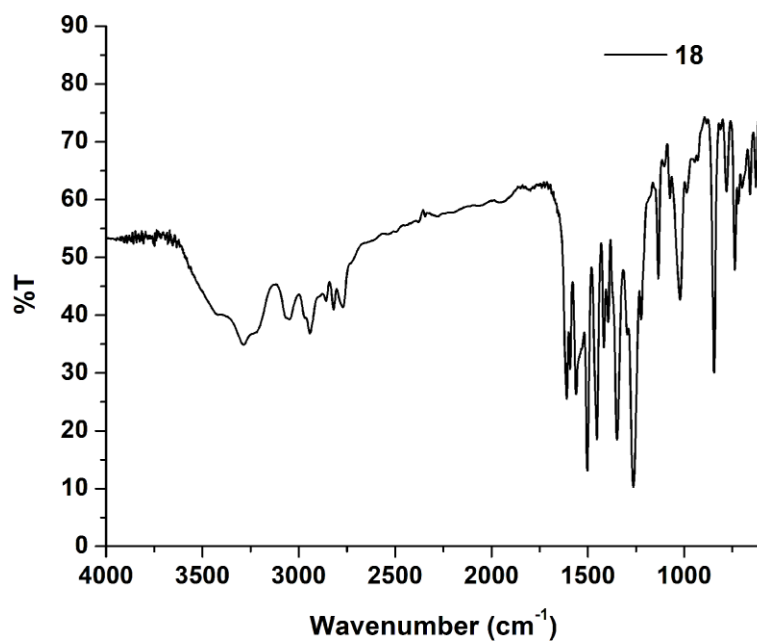
IR spectrum of **18**

APPENDIX C: CD MELTING AND CD TITRATIONS SPECTRA OF **G1** AND **G2** WITH

2-CLIP-PHEN DERIVATIVES
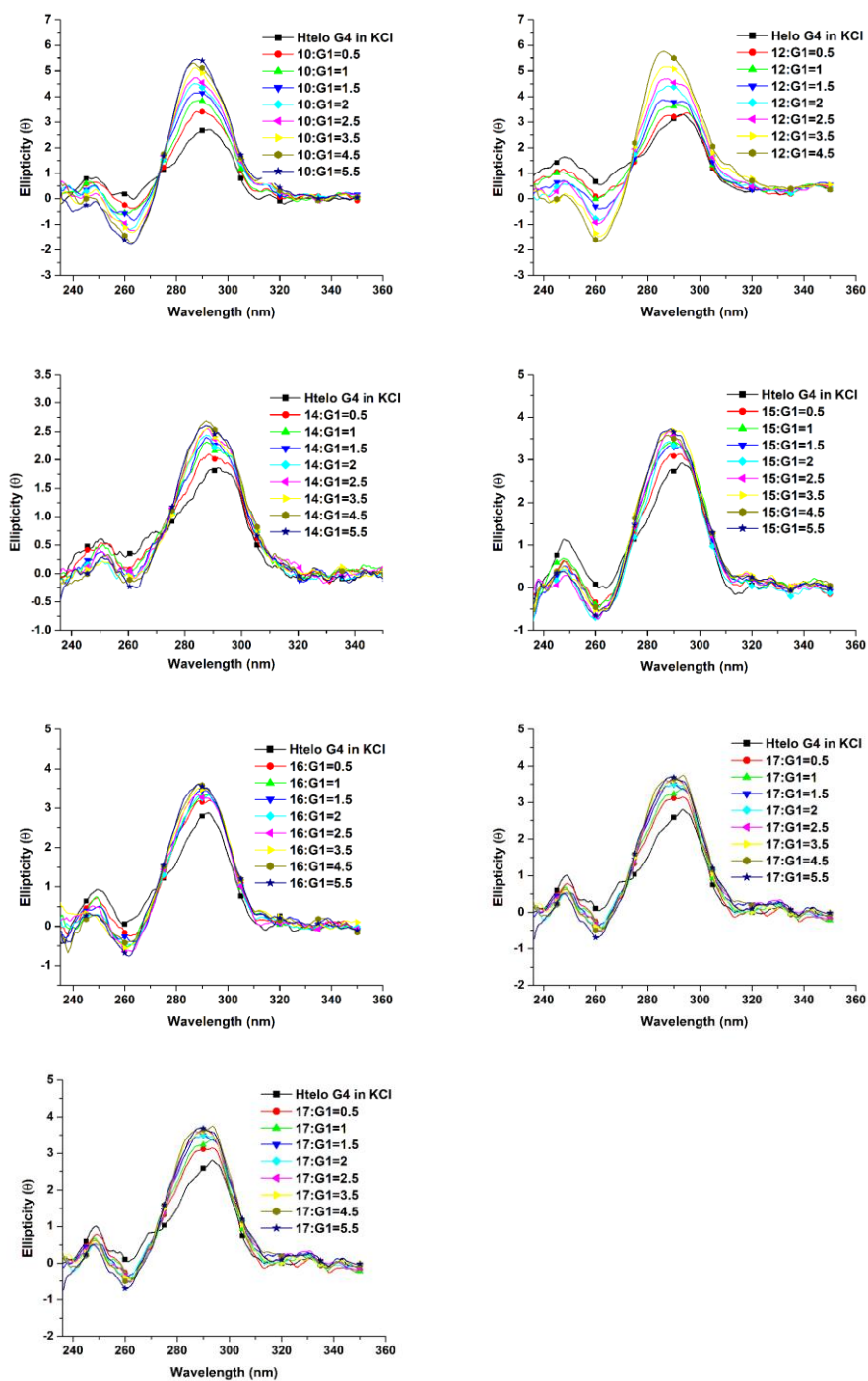
Figure C.1. CD Titrations curves for **G1** with 2-Clip-phen derivatives.

Figure C.2. CD Titrations curves for **G2** with 2-Clip-phen derivatives (**10**-**15**).
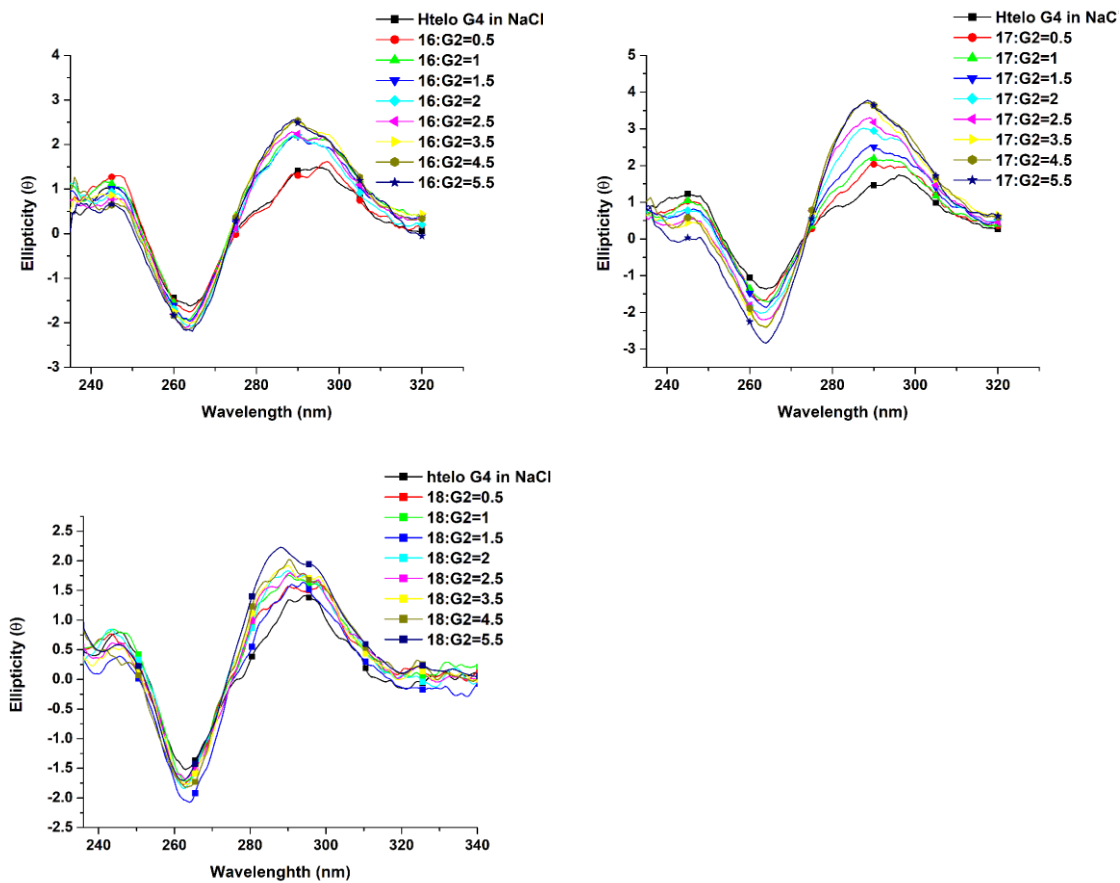
Figure C.3. CD Titrations curves for **G2** with 2-Clip-phen derivatives (**16**-**18**).

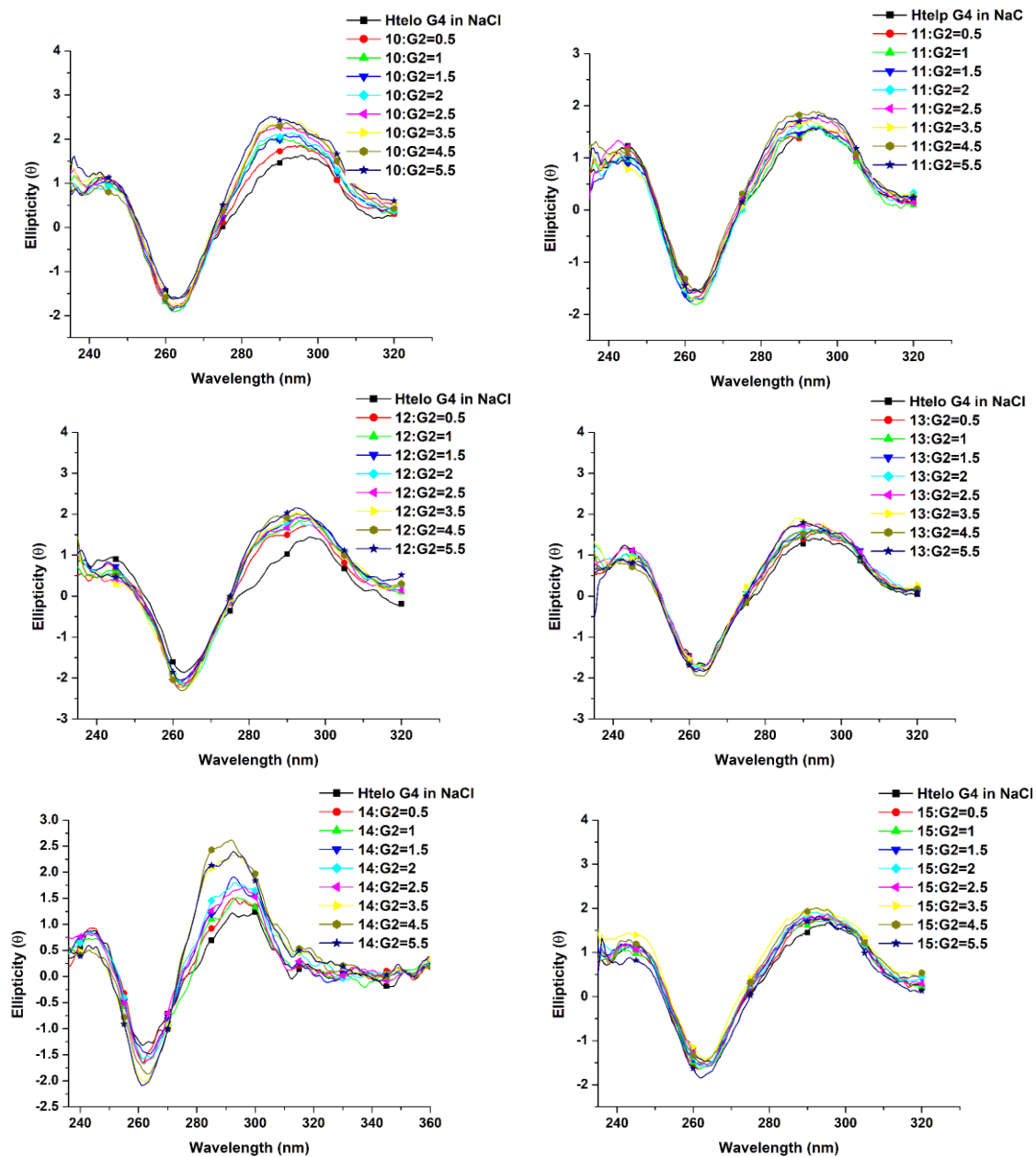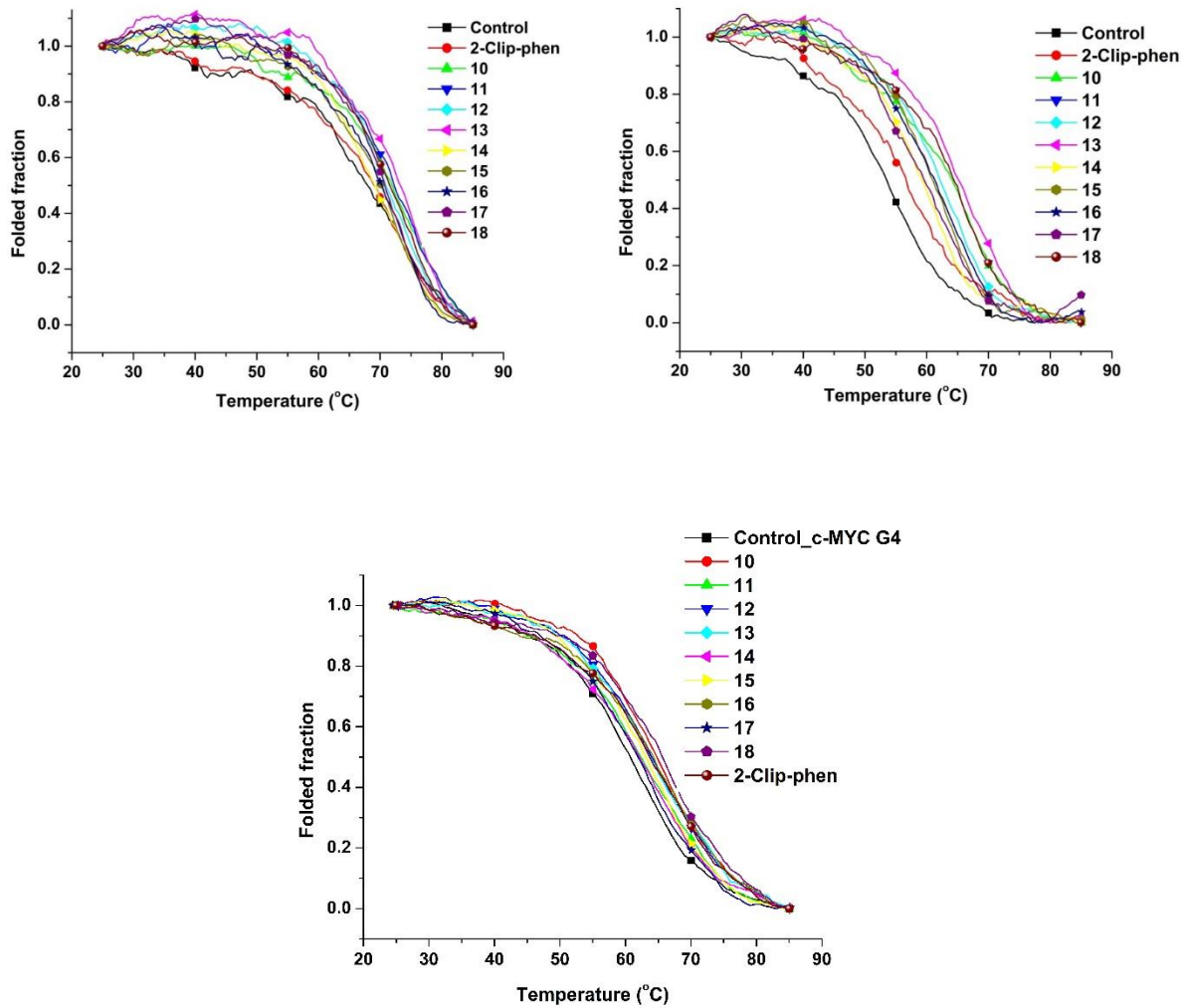*Figure C.4.* CD melting curves of **G1** (top-left), (**G2**) (top-right) and (**G3**) (bottom right) with 2-Clip-phen derivatives. The temperature at 0.5 folded fraction is called the melting temperature.
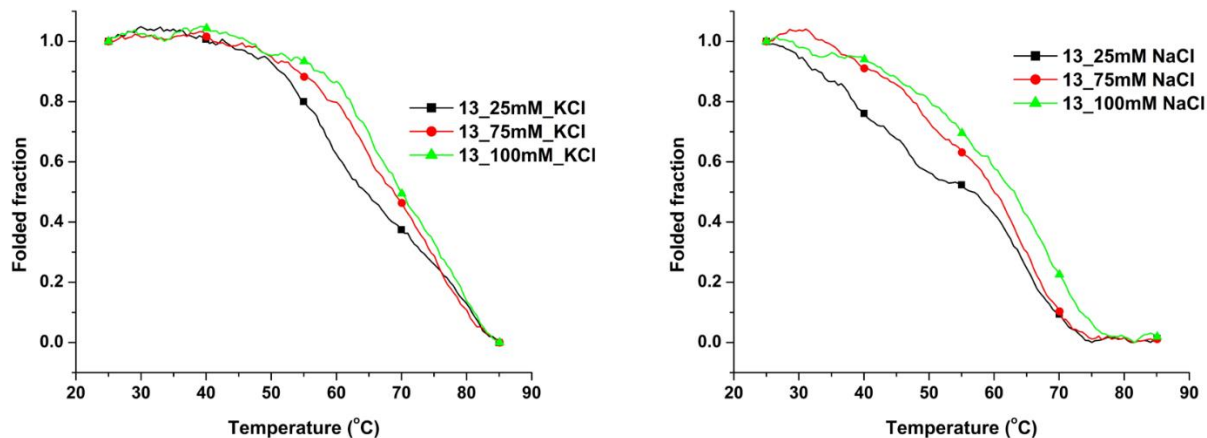
Figure C.5. Stabilizing effect of the compound **13** on the **G1** (left) and (**G2**) under different salt conditions. The two transitions are clearly visible at low salt concentration (25 mM, biphasic curves).
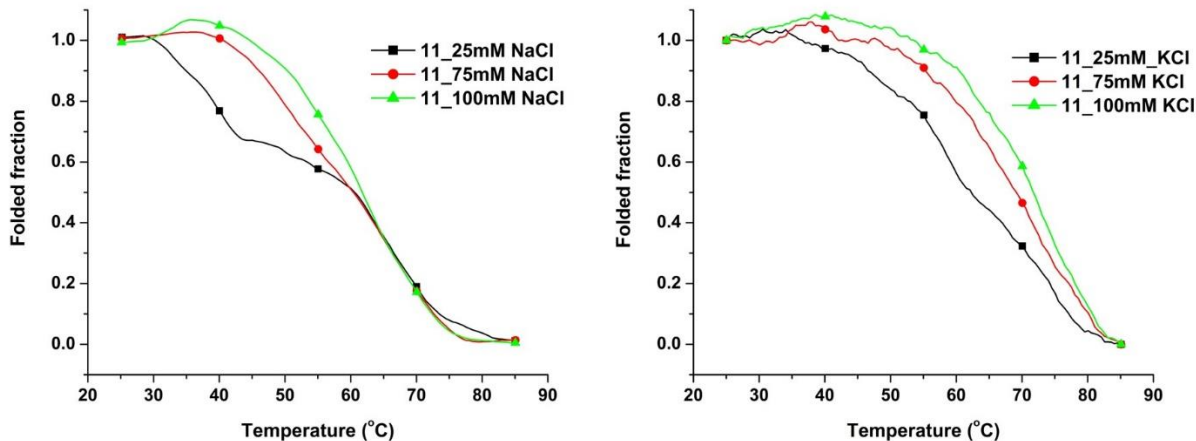
Figure C.6. Stabilizing effect of the compound **11** on the **G1** (left) and (**G2**) under different salt conditions. The two transitions are clearly visible at low salt concentration (25 mM, biphasic curves).

Figure C.7. Effect of KCl and NaCl concentration on the melting temperature (Tm) of the **G1** (left) and **G2** (right) respectively.

APPENDIX D: MIC DETERMINATION OF NEOMYCIN AND NUCEOBASE-NEOMYCIN

CONJUGATES

E. coli (Gram-)

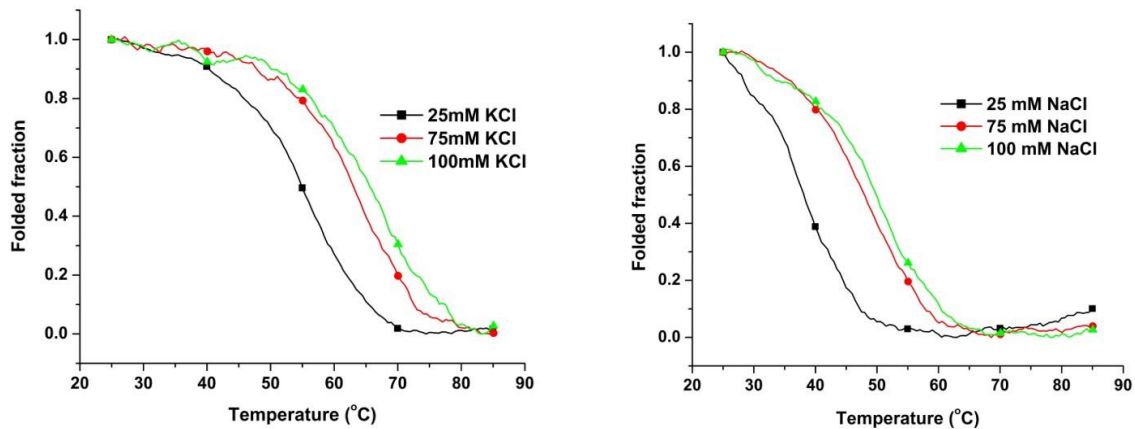| Conc (μM)	256	128	64	32	16	8	4	2	1	0.5	+ve Control	Emtpy
Neo	0.04	0.04	0.039	0.037	0.04	0.04	0.051	0.119	0.144	0.154		0.038
N-A	0.04	0.038	0.037	0.047	0.12	0.12	0.141	0.147	0.175	0.176	0.199	0.039
N-T	0.04	0.038	0.037	0.041	0.11	0.12	0.144	0.154	0.159	0.151	0.176	0.043
N-C	0.04	0.04	0.039	0.04	0.1	0.1	0.125	0.15	0.17	0.17	0.172	0.039
N-G	0.04	0.04	0.039	0.039	0.04	0.1	0.119	0.135	0.166	0.194		0.039

S. marcescens (Gram-)

Conc(μM)	512	256	128	64	32	16	8	4	2	1	+ve Control	Emtpy
Neo	0.04	0.039	0.04	0.038	0.04	0.04	0.037	0.039	0.226	0.352	0.761	0.038
N-A	0.05	0.039	0.038	0.167	0.24	0.35	0.542	0.731	0.754	0.783	0.745	0.045
N-T	0.04	0.039	0.037	0.139	0.19	0.32	0.49	0.722	0.754	0.757	0.738	0.039
N-C	0.04	0.041	0.039	0.04	0.16	0.22	0.278	0.45	0.63	0.78	0.73	0.039
N-G	0.06	0.047	0.042	0.041	0.04	0.13	0.249	0.351	0.479	0.655	0.8	0.039

B. subtilis (Gram+)

Conc (μM)	512	256	128	64	32	16	8	4	2	1	+ve Control	Emtpy
Neo	0.04	0.039	0.038	0.039	0.04	0.04	0.039	0.039	0.041	0.041	0.357	0.035
N-A	0.04	0.038	0.038	0.037	0.04	0.04	0.093	0.334	0.243	0.341	-	0.038
N-T	0.04	0.038	0.041	0.039	0.04	0.04	0.039	0.133	0.334	0.322	-	0.038
N-C	0.04	0.041	0.041	0.04	0.04	0.04	0.042	0.039	0.177	0.416	-	0.04
N-G	0.04	0.039	0.039	0.04	0.04	0.04	0.041	0.04	0.042	0.154	0.451	0.038

P. vulgaris (Gram-)

Conc (μM)	512	256	128	64	32	16	8	4	2	1	Positive Control	Emtpy
Neo	0.04	0.04	0.039	0.038	0.04	0.04	0.152	0.212	0.197	0.21	-	0.039
N-A	0.04	0.04	0.153	0.175	0.19	0.19	0.191	0.171	0.187	0.178	0.208	0.039
N-T	0.05	0.039	0.153	0.185	0.19	0.18	0.193	0.195	0.185	0.184	0.195	0.038
N-C	0.05	0.044	0.041	0.189	0.2	0.19	0.199	0.194	0.188	0.206	0.217	0.04
N-G	0.05	0.047	0.04	0.04	0.19	0.25	0.272	0.281	0.287	0.3	0.286	0.039

E. coli (resistant)

Conc (μM)	512	256	128	64	32	16	8	4	2	1	+ve Control	Emtpy
Kana	0.35	0.33	0.327	0.295	0.27	0.28	0.324	0.288	0.283	0.305	0.314	0.039
Neo	0.04	0.037	0.037	0.04	0.05	0.11	0.234	0.281	0.261	0.249	0.294	0.039
N-A	0.1	0.244	0.409	0.209	0.34	0.36	0.349	0.349	0.336	0.348	0.279	0.038
N-T	0.17	0.269	0.308	0.363	0.35	0.36	0.356	0.322	0.313	0.323	0.287	0.039
N-C	0.06	0.07	0.144	0.314	0.33	0.38	0.408	0.347	0.334	0.329	0.314	0.039
N-G	0.04	0.047	0.21	0.185	0.22	0.33	0.438	0.37	0.349	0.347	0.306	0.039

Figure D.1. MIC determination: The OD$_{600}$ values for each bacterial strain in the presence of neomycin and compound **6a-6d** at various concentrations.