UNIVERSITÀ
DEGLI STUDI
DI BRESCIA

DOTTORATO DI RICERCA IN

PRECISION MEDICINE

Settore scientifico disciplinare CHIM/08 CHIMICA FARMACEUTICA

CICLO XXXIV

TARGETING G-QUADRUPLEX DNA: SYNTHESIS OF NEW LIGANDS GUIDED BY ESI MASS SPECTROMETRY AND COMPUTATIONAL METHODS

DOTTORANDO

ALBERTO ONGARO

RELATORE

PROF.SSA ALESSANDRA GIANONCELLI

Summary of the chapters

Preface

This doctoral thesis represents the research work conducted during these last three years of doctoral studies in the field of the development of novel ligands for the targeting of the telomeric nucleic acid in its canonical double strand form and in its non-canonical G-quadruplex arrangement. The dissertation is composed of five main parts. The first part (Chapter 1) is an articulated introduction about the G-quadruplex, describing in the detail its function and structure with a particular focus on its different topologies and on the methodologies and techniques commonly used to study this particular fold; then the most important ligands for the telomeric G-quadruplex were presented also discussing their synthesis and their ability to stabilize the nucleic acid. The second, the third and the fourth parts (Chapters 2-4) are reporting, in the “paper collection” format, the investigations which I carried out on the development of new ligands to target the double strand and the G-quadruplex DNA. The three chapters were composed starting from three different articles in which I figure as the first name author or as one of the main contributors, (Chapter 2: Ongaro *et al.*, Chemistry Select, 2019; Chapter 3: Ribaudo *et al.*, Molbank, 2020; Chapter 4: Ongaro *et al.*, ChemMedChem, 2021); then additional *in silico* studies were included to each of the articles in order to harmonize the previous experiments and to add additional relevant data in the form of molecular docking and molecular dynamics studies. A general, overall conclusion, was then appended in the last chapter (Chapter 5).

Abstract (italian language)

La ricerca di nuove molecole in grado di legare il DNA a doppio filamento rappresenta un importante obbiettivo per il campo della chimica farmaceutica e potrebbe potenzialmente consentire di espandere il set di farmaci usati come cura per diverse patologie tra cui il cancro. Similarmente, l'identificazione di nuovi leganti selettivi per la struttura del DNA di tipo quadruplex può rappresentare un'interessante strategia alternativa per bersagliare l'acido nucleico nella sua forma non canonica. Inoltre i quadruplex sono presenti in specifiche regioni dei cromosomi chiamati telomeri, il cui mantenimento risulta cruciale per la vasta maggioranza dei tumori (J. W. Shay, W. E. Wright, *Nature Reviews Genetics*, 2019). Negli ultimi 30 anni, diversi ligandi sono stati sviluppati ed investigati per la loro capacità di stabilizzare il quadruplex; inoltre l'evoluzione nello studio strutturale ha portato alla scoperta di importanti tecniche analitiche ed investigative in grado di consentire lo studio sia della struttura stessa del quadruplex, che è ad oggi risaputo poter assumere diverse topologie, sia dell'interazione con esso di specifici ligandi. In questo contesto, la spettrometria di massa ESI in condizioni native è emersa come un potente strumento per analizzare importanti parametri del *binding* come la forza dell'interazione e la sua stechiometria. Inoltre l'uso di strumenti computazionali come il *molecular docking* e la *molecular dynamics* ha consentito, con un crescente grado di confidenza, la predizione della struttura finale dei complessi (Ribaudo *et al.*, *J Med Chem*, 2021). In questa tesi di dottorato verrà presentato il lavoro di ricerca da me condotto durante questi tre anni per lo sviluppo di nuovi ligandi capaci di interagire con il DNA telomerico, sia nella sua forma a doppio filamento e ancor più nella sua forma non-canonica di tipo quadruplex. Nello specifico, i tre lavori presentati comprendono la sintesi di diversi ligandi, la loro caratterizzazione e lo studio dell'interazione con l'acido nucleico attraverso tecniche spettroscopiche e spettrometriche ed indagini di tipo computazionale. Nei primi due studi verrà discusso lo studio di ligandi mono- e di-sostituiti del DNA a doppio filamento e di tipo quadruplex, basati sulla struttura di tipo antracene-propargilammina. In particolare, un set di 6 ligandi sintetizzati attraverso la reazione di *A3 coupling,* hanno dimostrato una buona abilità di interagire con l'acido nucleico ma senza selettività per la struttura quadruplex (Ongaro *et al.*, *Chemistry Select*, 2019). Per questo motivo la struttura è stata oggetto di ottimizzazione attraverso l'aggiunta di una seconda catena laterale, aumentando sia l'affinità che la selettività per il DNA di tipo quadruplex (Ribaudo *et al.*, *Molbank*, 2020). Nel terzo studio riportato (Ongaro *et al.*, *ChemMedChem*, 2021), la discussione si è invece focalizzata sui ligandi a struttura di tipo coniugato antrachinone-amminoacido, nello specifico 7 coniugati sono stati preparati usando l'efficace reazione di cicloaddizione CuAAC. Questi ultimi hanno dimostrato una elevata affinità e selettività per il DNA di tipo quadruplex ed inoltre l'abilità di indurre una variazione di topologia dello stesso.

Abstract (english language)

The search for novel molecules able to bind to the double strand DNA is an important topic in the medicinal chemistry field and could help therefore to expand the set of drugs used against various diseases including cancer. Similarly, the identification of new selective quadruplex ligands, represents an appealing, alternative strategy, to target the nucleic acid when folded in this non-canonical form. Moreover, the quadruplexes can be found in a specific region of the chromosomes called telomeres, which maintenance is crucial for the vast majority of tumors (J. W. Shay, W. E. Wright, *Nature Reviews Genetics*, 2019). In these last 30 years, various scaffolds were developed and investigated for the stabilization of the quadruplex motif; also the evolution in the study of its structure led to the discovery of important techniques to investigate either the quadruplex arrangement itself, which is nowadays known to assume different forms, namely topologies, and also the ligand binding event. In this context, the native ESI mass spectrometry emerged as a compelling tool to evaluate relevant parameters as the interaction strength and the binding stoichiometry. Furthermore, the use of computational tools such as molecular docking and molecular dynamics, allowed the prediction of the final structure of the complexes with an increasing degree of confidence (Ribaudo *et al.*, *J Med Chem*, 2021). In this doctoral thesis it will be presented the research work that I have conducted during these three last years on the development of new ligands able to target the telomeric DNA, either in its double strand form and even more in its non-canonical quadruplex fold. In the specific, the three works presented included the synthesis of different ligands, their characterization and the study of the interaction with the nucleic acid through spectrometric, spectroscopic techniques and computational investigations. In the first two studies, the investigation on mono- and di-substituted double strand DNA and quadruplex ligands based on the anthracene-propargylamine scaffold is discussed. In particular, a first set of 6 ligands, synthesized through the A3 coupling, demonstrated a good ability to interact with the nucleic acid but didn't show a selectivity for the quadruplex fold (Ongaro *et al.*, *Chemistry Select*, 2019). For this reason, the scaffold was optimized with the addition of a second side-chain, increasing both the affinity and the selectivity for the quadruplex DNA (Ribaudo *et al.*, *Molbank*, 2020). In the third presented study (Ongaro *et al.*, *ChemMedChem*, 2021), the discussion is focus instead on the anthraquinone-amino acid conjugate scaffold, prepared leveraging the CuAAC reaction which allowed an efficient generation of a set of 7 conjugates. The ligands demonstrated an elevated affinity and selectivity for the quadruplex DNA and the ability to induce a variation of the fold topology.

1 Introduction

1.1 The G-quadruplex

The canonical form of DNA is known as B-form double helix, this arrangement is formed by two single stranded DNA filaments held together by Watson-Crick-type base-pairs (Figure 1.1, a). Structures that disobey the Watson-Crick canon are described as non-canonical and provide for less common base pairing patterns. Guanine-rich sequences are known to disobey the rule and to fold into the non-canonical arrangement called G-quadruplex (G4), which compared to the double strand form of the DNA (dsDNA) represents a small part of the whole DNA content of a cell (Figure 1.1, b). Current estimates however suggest that human cells contain at least 716'000 DNA G4s in a volume (for a HeLa cell nucleus) of 0.22 pL, corresponding to a cellular G4 concentration of 6 µM, which appears as a quite remarkable presence.[1–3]

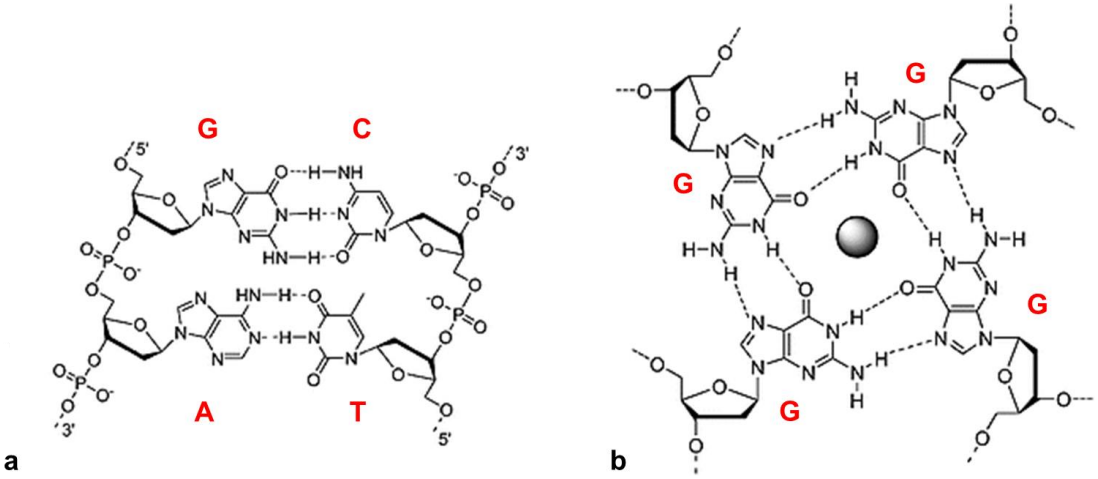


Figure 1.1. Canonical double strand DNA (a) and non-canonical G4 DNA (b). The dashed lines represent the H-bonds between the nucleobases.

G4s are present in the promotorial region of several oncogenes such as c-myc, vegf, bcl2, k-ras and c-kit and therefore the induction of these structures represents an interesting strategy for the selective inhibition of their transcription.[4–6] More in the detail, the oncogene suppression when induced by the G4 presence, represents an obstacle for the action of the RNA polymerase enzyme and of transcription factors, mostly leading to the downstream silencing of the gene which was observed for various oncogenes (Figure 1.2).[7–12] An up-regulation of the genes however is also possible as the G4 can interfere with transcription factor binding.[13]

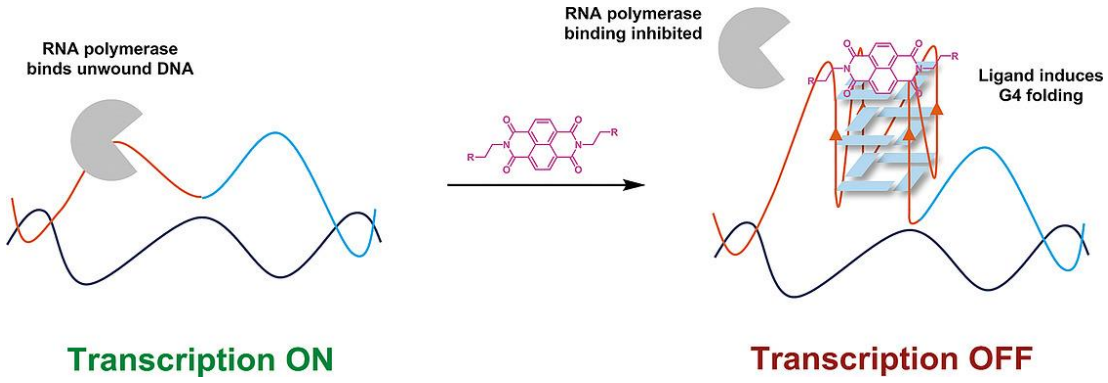


Figure 1.2. RNA polymerase transcription of dsDNA (left) and transcription inhibition due to ligand induced G4 folding (right).

G4s can be also retrieved in a particular region of the chromosomes which is called the telomeric region due to the presence of particular sequences of repeating nucleotides.[14] The telomeres have the role to regulate through their length the life cycle of the cells. In fact, most normal cells, cycle and divide a limited number of times,[15] in particular at every cell cycle the telomeres are shortened, until reaching a critical length called the Hayflick limit, when the cells cease proliferation permanently and enter a state called replicative senescence in which they are still alive and functional but do not divide. The telomerase is an enzyme, which has the role of adding telomere repeat segments to the 3' end of telomeric DNA, therefore preserving the cell from reaching the abovementioned senescence state. Specifically, telomerase is a reverse transcriptase enzyme that carries its own RNA molecule, which is used as a template for the telomeres elongation. Interestingly, telomerase is almost absent in non-immortalized cells but it is expressed at significant levels in most spontaneously immortalized cells, including human cancer cells.[16–18] Moreover, the reduction of the telomerase activity demonstrated to promote various phenomena including uncapping processes, chromosomal fusions, and senescence, which would eventually trigger DNA alteration and apoptosis.[19,20]

In addition to that, also the RNA nucleic acid is able to form G4s, which are reported to be even more stable than their DNA counterparts.[21] The telomeres, despite their heterochromatic state, are transcribed giving rise to long noncoding RNAs (lncRNA) called TERRA (telomeric repeat-containing RNA) which are known to play crucial roles in telomere biology, including regulation of telomerase activity and heterochromatin formation at the chromosome ends.[22]

Moreover, the presence of G4s was demonstrated in different bacteria and also in RNA or DNA viruses, including human immunodeficiency virus (HIV), herpes simplex virus (HSV), human papillomavirus (HPV), Epstein-Barr virus (EBV), Hepatitis C virus (HCV) and also in coronaviruses including the severe acute respiratory syndrome-coronavirus-2 (SARS-CoV-2).[23–26]

It was also demonstrated that the RNA-G4 motifs are key elements in regulating the life cycle of viruses and for this reason are extensively studied in the virology research field.[26]

Furthermore, different proteins were reported to interact *in vitro* with G4 structures, in particular helicases such as the Fanconi anemia group J protein (FancJ), the Bloom syndrome protein (BLM), the Werner syndrome protein (WRN), and the petite integration frequency 1 (Pif1) are able to unwind G4 motifs formed by various sequences. Genomic instability and increases in risk of cancer are usually associated with mutations impairing the function of these kind of proteins.[27]

Therefore, from the discovery of the existence of these peculiar DNA and RNA arrangements, G4s have been identified as promising therapeutic targets and the study of their stabilization became an appealing research field in medicinal chemistry.[28–33]

1.2 The structure of G4

1.2.1 Relevant interactions

The total free energy can be decomposed in different free energy contributions (*e.g.* coulombic forces, hydrogen bonding, hydration, van der Waals terms), which can themselves be decomposed in entropic and enthalpic contributions. Analysis of G4 melting and calorimetric data revealed that G4 formation is enthalpically driven.[34,35] The G4 motifs are constructed via pi-pi stacking of *quasi* co-planar quartets of guanines, usually triple overlapped single layers constitute the most common G4 structures but double or quadruple layers have also been observed.[36–38] The factors contributing to the stabilization of G4s and yielding a particular folding topology or mixture of topologies are: stacking interactions, hydrogen bonding, solvation, and cation binding. Cation coordination is indeed absolutely required to form G4s stabilizing the G-tetrad stacks. Moreover, sufficient ionic strength is required to compensate electrostatic repulsion between the phosphate oxygens of four strands in G4s, instead of two for duplexes (loops may be considered as unfolded single strands depending on the structure). Similarly to the secondary structures of other nucleic acid arrangements, G4 stabilization relies in part on stacking of aromatic bases, and more precisely of guanines from consecutive tetrads, molecular mechanics simulations predicted the relative stability order among guanine stacks to be (from 5' to 3'): syn/anti > anti/anti > anti/syn > syn/syn.[39] More in the detail, the guanine tetrad structure is maintained by Hoogsteen-type base-pair hydrogen bonds (H-bonds) between N1H and O6 lone pairs, and between N2H and N7.[40] In consequence of the formation of the G4 motifs, as H-bond acceptor and donor groups are likely hydrogen-bonding to water for unfolded of unassociated strands, there is probably only a small enthalpy change, but the release of water molecules to the bulk medium may induce an entropy gain.[41] A combination of high-resolution variable-temperature scanning tunnelling microscopy (STM) and density functional theory (DFT) calculations suggested that the stabilization of a G-tetrad induced by the network of eight hydrogen bonds is higher than the sum of four individual G-G pairs. This cooperativity initially ascribed to pi assistance, more likely arises from the charge separation that is associated to charge transfer between pairs of guanine, namely the donor-acceptor orbital interactions in the sigma-electron system. This results in an interruption of the pi-electron system, and an enhancement of both the positive charges on the H atoms of H-bond donor groups and negative charges of N and O atoms of acceptor groups.[42,43]

1.2.2 Coordination of ions

The coordination to the guanine O6 of various mono, di and even tri-valent cations in the central cavity formed by the G-quartets is crucial for G4 stabilization. The cations coordinated in the G4s are involved in electrostatic and donor-acceptor orbital interactions with the lone pairs of guanine O6 producing tight M+-O coordination bonds,[43] moreover the stabilization is also increased by electronic repulsion of the O6 lone pairs.[44] In solution, potassium is typically hexa-hydrated, and hence the coordination per se by guanine O6 instead of water molecules shall not provide a large enthalpy change. Since Na^+ and K^+ are the physiologically relevant monovalent ions, the majority of the studies focused on their role and influence on G4 structure and stability, however more complete works have determined specific patterns for mono and di-valent cations, by compiling a number of studies, is possible to build the following general stabilization trend: $Sr^{2+} > K^+ > Ca^{2+} > NH_4^+$, Na^+, $Rb^+ > Mg^{2+} > Li^+ \geq Cs^+$.[45,46] Generally, the bigger K^+ cations (1.33 Å) are specifically coordinated in between tetrads, in a nearly octahedral way, or alternatively within loops above tetrads, satisfying the usual hexacoordinate stereochemistry of these cations. Positioning of bases outside of the plane of the tetrad may be necessary to comply with this geometry, to an extent balanced by stacking and hydrogen bonding energies. Smaller Na^+ (0.95 Å) cations can be coordinated within the plane of tetrads, and can occupy a range of positions owing to lower steric constraints, hence reducing the electrostatic repulsions. Incidentally, in the first reported crystal structure of a small molecule (daunomycin) bound to a G4, the Na^+ cations are all coordinated in the plane of the tetrads.[47] In conclusion it is worth to point out that the chelation of K^+ form is generally considered to be more biologically relevant due to the higher intracellular concentration of potassium (~140 mM) than that of sodium (5–15 mM).[48]

1.2.3 Different G4 topologies

Different topologies are possible for G4s depending on the orientation, in terms of 5' to 3' polarity, of the guanine tracts; moreover these arrangements are acquired by the nucleic acids depending on their sequence and also in consequence of specific environmental conditions. For both intramolecular and intermolecular sequences, parallel, antiparallel and hybrid arrangements have been reported. The parallel topology, possesses four strands of the same polarity and is prevalent in overcrowded solution conditions as inside the cells,[49–53] in addition to that, molecular crowding conditions has been found to induce the transition from hybrid-type to parallel topology and to alter the competition between G4 and duplex formation.[52,53] The antiparallel G4 presents two strands in a way, and the two others in the opposite. The hybrid topology shows three strands sharing the same polarity, and the fourth one the opposite and is therefore sometimes referred to as "3+1" G4.

Furthermore, the glycosidic bond angle of the guanines constituting the tetrads can adopt an *anti* or *syn* geometry depending on the relative strand orientation: parallel topologies contain almost exclusively anti guanines,[54–56] whereas a mixture of *syn* and *anti* is observed for antiparallel and hybrid structures.[57] Consequently, parallel G4s generally contain *anti/anti* stacks, antiparallel present *anti/syn* and *syn/anti*, and hybrid ones can have all of the above (Figure 1.3).

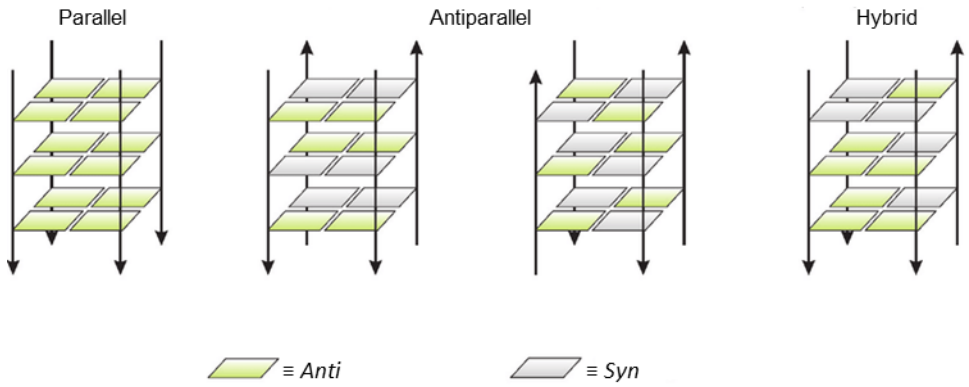


Figure 1.3. Various patterns of syn or anti glycosidic bonds of the stacked guanines forming all the different G4 topologies.

The part of the nucleic acid filaments connecting the G-tracts are referred to as loops, which can be differentiated in three types called lateral, diagonal, and propeller loops. The propeller loop connects adjacent parallel guanine chains, the lateral connects adjacent anti-parallel tracts, and the diagonal loops join antiparallel guanine tracts on the G4 diagonal. Therefore, the parallel topology presents three propeller loops, the hybrid one shows two lateral loops and one propeller loop, and the anti-parallel topology has either two lateral and one diagonal loops, or three lateral loops. Furthermore, the hybrid topology possesses two forms called hybrid-1 and hybrid-2, which mainly differ for the "sense" of the folded sequence and therefore also for the order of the guanine main chains and loops. In addition, the anti-parallel topology presents two forms, named the chair and the basket type, with different guanine main chains order and loop types.

Until recently, little was known about the relationship between G4 topologies and their biological functions. Few years ago the correlation between the thermal stability of the topological structures of G4 and the rate of DNA synthesis was studied in the detail by Sugimoto and co-workers, concluding that the extent of inhibition of DNA synthesis by the Klenow fragment of DNA polymerase I depends on the topology of G4. This clearly revealed the importance of the G4 topology, paving the way for a more conscious targeting of this particular fold.[58]

See the figure below (Figure 1.4) for a schematic representation of the different G4 topologies and the different types of loops.

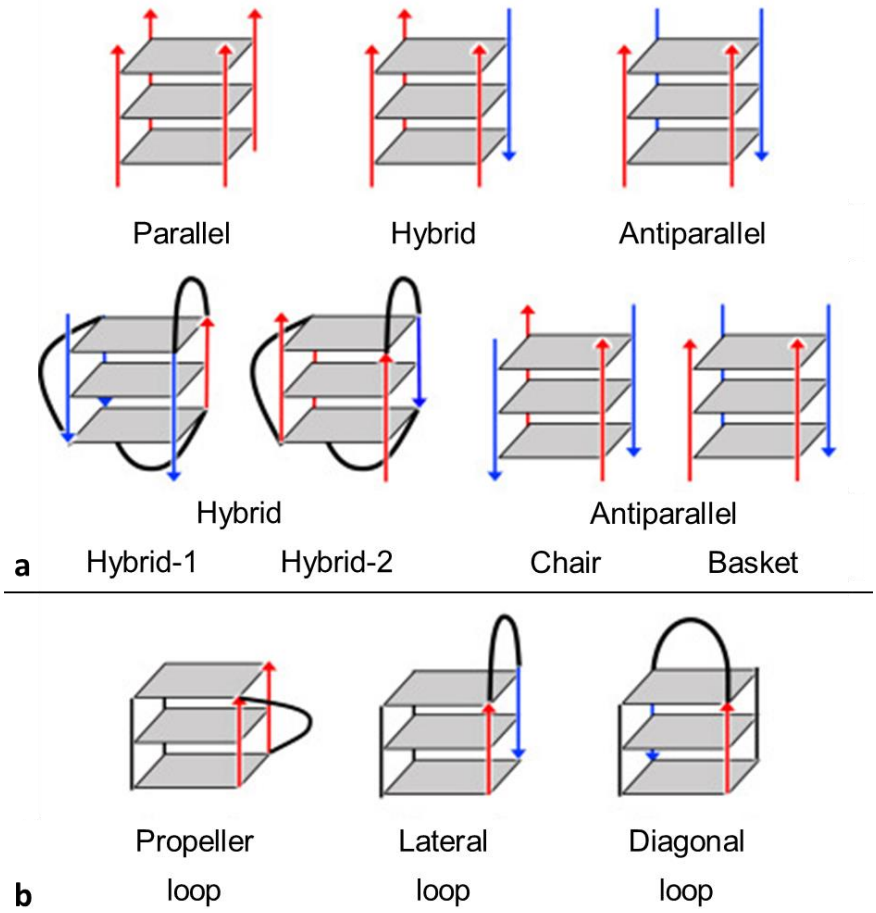


Figure 1.4. Representation of all the possible topologies for the G4 arrangement (a) and of the different types of loops linking the G-quartets (b).

1.3 Telomeric G4

Telomeres contain short tandem repeats—sequences of basepairs that are repeated numerous times. The sequence of the repeat unit is TTAGGG in mammalian cells and human telomeres contain around 10–15 kilobasepairs per chromosome end at birth. Other telomeric sequences as for example TTGGGG, TTTTGGGG, TTTAGGG and TTTTAGGG, can be found in telomeres of species like Tetrahymena, Oxytricha, Arabidopsis, Chrorella and Chlamydomona, respectively.[59] Four repeats of the human telomeric sequence, $(TTAGGG)_n$ are able to fold into a monomeric (single unit) G4 structure, moreover X-ray crystallographic and NMR studies have shown that the same sequence can adopt various topologies including the intramolecular antiparallel (chair and basket forms), the hybrid (hybrid-1 and hybrid-2) and the parallel forms.[60–63] Furthermore, depending on different parameters, namely the exact sequence, type of ions, solvent conditions, oligonucleotide concentration and other factors, it was demonstrated the possibility of a conversion between parallel, hybrid and antiparallel topologies.[64–66]

For example, a 22mer human telomeric sequence $[AG_3(T_2AG_3)_3]$ was reported to form in Na^+ solution a basket-type antiparallel G4 structure,[67] whereas in K^+ solution the hybrid-1 and hybrid-2 are primarily formed.[60,63,68,69] The same sequence instead, in a crystal grown in the presence of K^+ ions and/or in molecular crowding conditions, demonstrated to adopt a completely different parallel G4 motif.[70–72] The telomeric sequence $[G_3(T_2AG_3)_3T]$ was reported to adopt in K^+ solution an antiparallel basket type G4 motif composed by two G-tetrad layers.[37,67] However, nowadays it is commonly accepted that the predominant topology in K^+ solution for telomeric G4 is the hybrid both in the hybrid-1 and in the hybrid-2 forms, which are in equilibrium, moreover evidences in the literature indicate the 3′-flanking sequence determines the position of this equilibrium.[62,63,73] Very recently, it has also been reported that the human telomeric sequence can form hybrid-type G4s in living human cells.[74] An illustrative example of the different topologies adopted by a telomeric sequence is reported in the figure below (Figure 1.5).

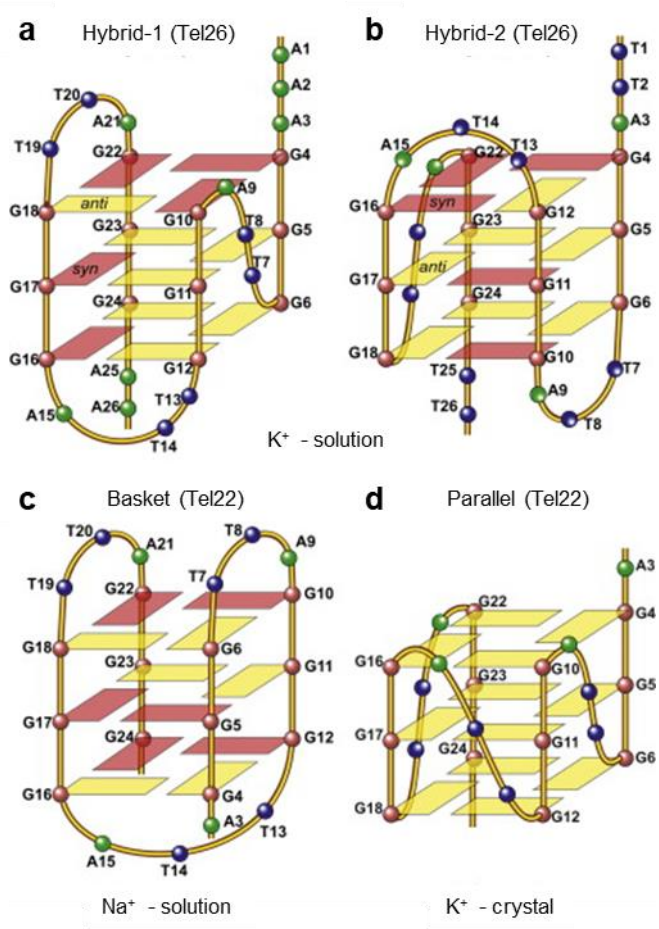


Figure 1.5. The four main topologies adopted by a telomeric sequence in different conditions, namely in the presence of K^+ ions in solution or in a crystal or in the presence of Na^+ in solution.[69]

Lastly it is noteworthy to briefly cite some particular cases of more exotic or complex G4 structures namely G-triplex (G3), monomeric intermolecular or multimeric intermolecular arrangements.

The formation of G3 structures was recently observed from oligonucleotides containing three tandem G-tracts. The removal of the tetranucleotide 5′-TTGG-3′ from the 3′ end of the thrombin binding aptamer of 5′-GGTTGGTGTGGTTGG-3′ (TBAG4) promotes the formation of an intramolecular G3 structure (TBAG3). Moreover, human telomeric DNAs with three tandem G-tracts demonstrated to fold into G3 structures (htG3).[75–80]

The monomeric intermolecular G4 arrangements, can be formed by the intermolecular association of four DNA strands having only one G-tract (for example: 5′-TTAGGG-3′, 5′-TTGGGG-3′, 5′-TGGGGT-3′, 5′-TTGGGGT-3′, and 5′-TGGGT-3)[81–83] or by the dimerization of sequences that contain two tandem G-tracts.[84–86] Even though these arrangements are generally considered less biological relevant than the intramolecular G4s, they have been proposed to occur *in vivo* in recombination, telomere pairing and other phenomena.[2,87]

On the other hand, multimeric G4 can be formed by consecutive G4s on the same DNA filament which are also called G-wires or less commonly by G4s coming from different sequences. The G-wires structures, are considered the most physiologically relevant as the single-stranded sequences of telomeric DNA can potentially fold into oligomers containing up to ten consecutive G4s connected by TTA spacers.[2,88–92] Their existence was for the first time demonstrated by Sugimoto et al in 2006 then mass-spectrometry assays conducted by Collie and co-workers confirmed the discovery,[93] furthermore their structure was also elucidated by atomic force microscopy (AFM) imaging.[94] As depicted in the following figure (Figure 1.6), these intriguing structures can belong to three different models based on the type of interaction, which happen between them. In the non-interacting "beads-on-a-string" model the G4s are separated and behave as single units, on the other hand in the interacting models the G4s can interact each other through the external G-tetrads (parallel topology) or through the loops (hybrid-type topologies).[88]

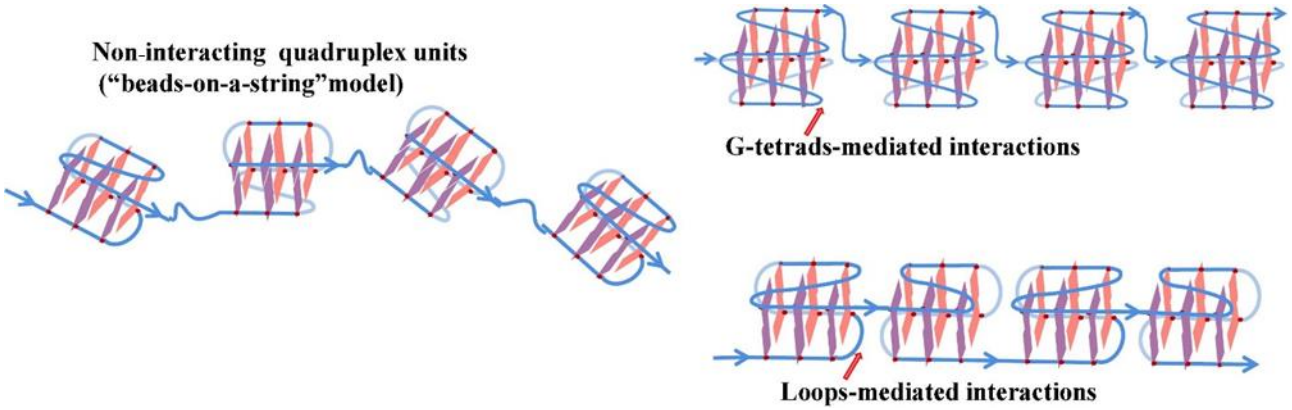


Figure 1.6. Three possible models of multimeric telomeric quadruplexes composed by non-interacting or by interacting G4 units.[88] The G-tetrads-mediated interactions arise between two external tetrads of two consecutive G4s; when the loops are present, additional interactions can be formed between the loops of two consecutive G4s.

1.4 Methods to study G4 and the G4-ligand interaction

1.4.1 X-Ray crystallography

X-Ray crystallography is a powerful technique, which is able to determine the molecular structure of a crystal starting from the position of the single atoms. It provides access to the atomic-scale resolution of small molecules but also of supramolecular entities such as proteins and nucleic acids. Its use in the G4 field gives important informations on their structure/topology, and on the position of cations, surrounding waters, and also of ligands. Through this method, different groups have determined the structure of many structures of G4, mainly as tetramers or dimers.[82,83,95,96] Furthermore, the vast majority of the crystallized G4s are parallel probably due to an induction of topology driven by the crystal packing forces, no examples of antiparallel or hybrid motifs were reported up to date. Even if its use appears as limiting for the study of G4 structures in solution conditions, the parallel structures obtained can be still valuable as they are probably close to the structures arising in crowded conditions which are found inside the cells, moreover the structures provided are most of the times better defined than the NMR solved ones which are limited by internal dynamics or the presence of mixtures of conformations in equilibrium.

1.4.2 Nuclear Magnetic Resonance (NMR) spectroscopy

High-field NMR is broadly used to study G4 structures and their topology in an aqueous solution and at room temperature,[57,97,98] this technique permitted to solve many nucleic acids's structures including the human telomeric DNA[37,60–63,67,73,99–107] and RNA sequences (TERRA),[108–110] telomeres from other species,[37,45,111–117] and human oncogenes.[118–121] As previously said, the presence of internal dynamics or of mixture of conformations can limit the power of NMR to fully solve some structure, however crucial informations are still obtainable. The assignment of the peaks is usually performed by through-space NOE experiments,[114] often helped by base modifications or by site specific enrichment with ^{15}N and ^{13}C.[122]

The formation of G4 can be assessed by typical signals as the ones arising from the guanine imino protons H1 in the 10-12 ppm range, canonical Watson-Crick H-bond interactions instead give rise to signals at 13-14 ppm.[123,124] Furthermore, compared to Watson-Crick protons, the protons of the central tetrads exchange more slowly with water, producing sharper peaks.[114,125] The number of peaks belonging to imino proton is correlated to the number of tetrads producing one peak per guanine, hence four peaks per tetrad.[126] A higher number of guanine peaks indicates a mixture of conformations which can arise in the NMR time-scale.[98] The topology determination is also possible to perform, in particular studying the connectivity between guanine H1 and H8 and the J-couplings through tetrad hydrogen bonds.[97]

In addition to that, with the NMR technique, is also possible to study efficiently the binding event of a ligand to a G4 structure. In particular, the determination of the binding stoichiometry and of the binding constants is possible by monitoring the variation of specific signals, as the ones related to the imino protons, upon titration of the ligand to a solution of the G4 or viceversa.[127] Then methods like the Scatchard or Job's plot are generally used to interpolate the data as it is possible with the other spectroscopic techniques.[102,117,127]

1.4.3 Electronic circular dichroism spectroscopy

One of the most used and powerful low-resolution technique to study the structure of nucleic acids and in particular of their non-canonical arrangements is the circular dichroism (CD).[128–131] CD relies on the difference in absorption between the left- and right-handed circular polarized light by chiral compounds or by achiral compounds placed in a chiral surroundings, which is the case for example of a ligand bound to a receptor. The G4 motif is typically characterized by specific bands of absorption in the UV (210-300 nm) which primarily arise from electronic transition (π-π*) of the guanines involved in the G-quartets. More in the detail the sign (+ or -) and the intensities of the bands depends on the relative orientation of the stacked guanines (anti/syn, syn/anti or anti/anti, *syn/syn*). Thanks to this peculiarity of the G4 arrangement, through the CD analysis, is possible to obtain important information on its structure, most of the times even assigning the relative topology. The parallel topology possesses a typical positive band at 260 nm and a negative band at 240 nm. The antiparallel is characterized by a positive band at 290 nm, a negative one at 260 nm and a shoulder between 260 and 270 nm. The hybrid 1 and 2 topologies are not distinguishable; however, the hybrid types generally show a positive band at 290 nm and a shoulder at 260-280 nm. A general profile for the different topologies can be deduced by the figure below (Figure 1.7).

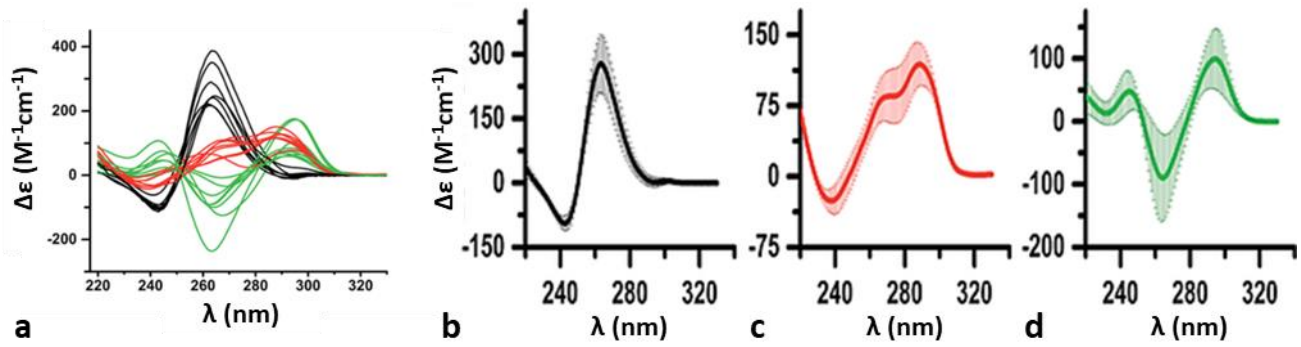


Figure 1.7. Clusterization of the CD spectra of parallel (black), hybrid (red) and antiparallel (green) G4 topologies. The graphs reported were retrieved from the work of Villar-Guerra and colleagues which obtained them from a curated library of circular dichroism spectra of 23 G4s of known structure.[132] In the first graph, the curves related to the different topologies were overlapped (a). In the graphs on the right the curves correspond to the three topologies, parallel (b), hybrid (c) and antiparallel (d), the curve in bold was obtained averaging the various curves registered.

CD is usually a fast and easy technique but has some drawbacks as it fails when in the presence of mixtures of structures, however issues are present with all the analytical techniques, and therefore only a combination of them can provide a clear information about a particular structure.[133] Other than the structural application, the CD technique possess a great value also in the study of the binding of a ligand to a G4 arrangement, offering through titration experiments, relevant information about the relative binding constants and, when present, about the induction of topology phenomenon. Moreover, CD is also used in the determination of a ligand induced thermal stabilization of a G4 motif in melting assays.

1.4.4 UV-Vis absorption and fluorescence spectroscopy

The UV spectrum is usually similar between different G4 topologies and for this reason the UV-Vis absorption spectroscopy is not used to determine the G4 fold. The difference in the UV absorption before and after melting named thermal difference spectrum (TDS) produces a "fingerprint" spectrum which can be used to distinguish different nucleic acids high ordered structures, including G4.[134] The TDS fingerprints of G4 usually show a minimum at 295 nm and local maxima at 240, 255 and 275 nm.[134] Different guanine-guanine stacks can modify the intensity of specific absorption bands permitting to discriminate parallel from hybrid and antiparallel topologies, thanks to the greater ΔA_{240} nm/ΔA_{295} nm ratio. Isothermal difference spectra (IDS) are calculated by subtraction of the UV spectra of a sample acquired in absence or presence of cations, at constant temperature, with this method is possible to study kinetics of folding and structural interconversion.[135] In addition to that, the stability of a G4 structure can be assessed by UV-melting monitoring the absorption at 295 nm, in fact the temperature increase promotes the G4 motif unfolding which translates into a hypochromism at 295 nm.[136,137]

Another possibility to study the G4 temperature driven folding/unfolding processes is given by a method which relies on the Förster resonance energy transfer phenomenon (FRET), which is a mechanism of energy transfer happening between two light-sensitive molecules also called FRET pair (for example dabcyl-FAM).[138] A donor chromophore which can be excited through irradiation at a specific wavelength can transfer energy to an acceptor chromophore by non-radiative dipole-dipole coupling, with an efficiency inversely proportional to the sixth power of their distance. The excited acceptor then releases its energy by fluorescence and the intensity is revealed by a fluorimeter instrument. This method can be used for example to detect the melting temperature (T_m) of an intermolecular G4 motif using two single strand DNA sequences tagged with the FRET acceptor or the donor. This technique is widely used also to determine the ability of a ligand to stabilize a G4 arrangement, usually first analyzing the DNA T_m itself and then measuring the difference in temperature (ΔT_m) when in the presence of a ligand.[139–141] From the ΔT_m values is also often possible to determine with a certain approximation the binding constants through simple equations.

1.4.5 Native mass spectrometry

The Native Mass Spectrometry (MS) is an analytical mass spectrometry technique, which conditions both in the sample solution and in the spectrometer, during the ionization and vaporization phases, are non-denaturing, therefore are considered not to influence the integrity of a supramolecular arrangement, and in our specific case of the DNA.
The electrospray ionization MS (ESI-MS) is the most widely used soft-ionization technique for the native MS method, and even if it has a higher propensity to denaturate, also the matrix-assisted laser desorption/ionization (MALDI) has been used for the scope.[109,142,143]

In the ESI-MS source the analyte dissolved in an aqueous solution, usually containing also a certain percentage of an organic cosolvent (usually methanol or isopropanol), is infused at atmospheric pressure then an electric field is applied and the analyte enters a heated capillary as charged droplets. Due to the collision with an inert gas (usually nitrogen) which provides the energy required for the solvent evaporation the droplets shrink until the repulsion between the charges at the surface increases and, at a certain droplet radius, this repulsion overcomes the cohesive force of the surface tension leading to the so-called Coulomb fission which gives rise the formation of desolvated ions in the gas phase which are then separated and revealed by their mass on charge ratio (m/z).[144] In these mild conditions, during the ESI process, minimal fragmentation occurs, and as a result noncovalent interactions are not altered.[145]

The native ESI-MS method, which first use date back to the 1990s,[146] is nowadays commonly applied to determine different parameters of nucleic acid samples as strand stoichiometry, folding kinetics and thermodynamics, number of ions retained in the arrangement and also to study the binding of ligands to nucleic acids. Moreover, for the study of nucleic acid-ligand interaction, the method demonstrated its reliability in different studies, in particular by the comparison with the fluorescence melting assay.[147–150]
Information on crucial ligand binding parameters can be obtained by the relative intensities of the peaks corresponding to the free ligand, free receptor and the ligand-receptor complex, moreover in the native MS method, the relative intensities of the signals corresponding to the relative species are assumed proportional to their concentration in the sample solution. An exemplificative spectrum is reported in the figure below (Figure 1.8).

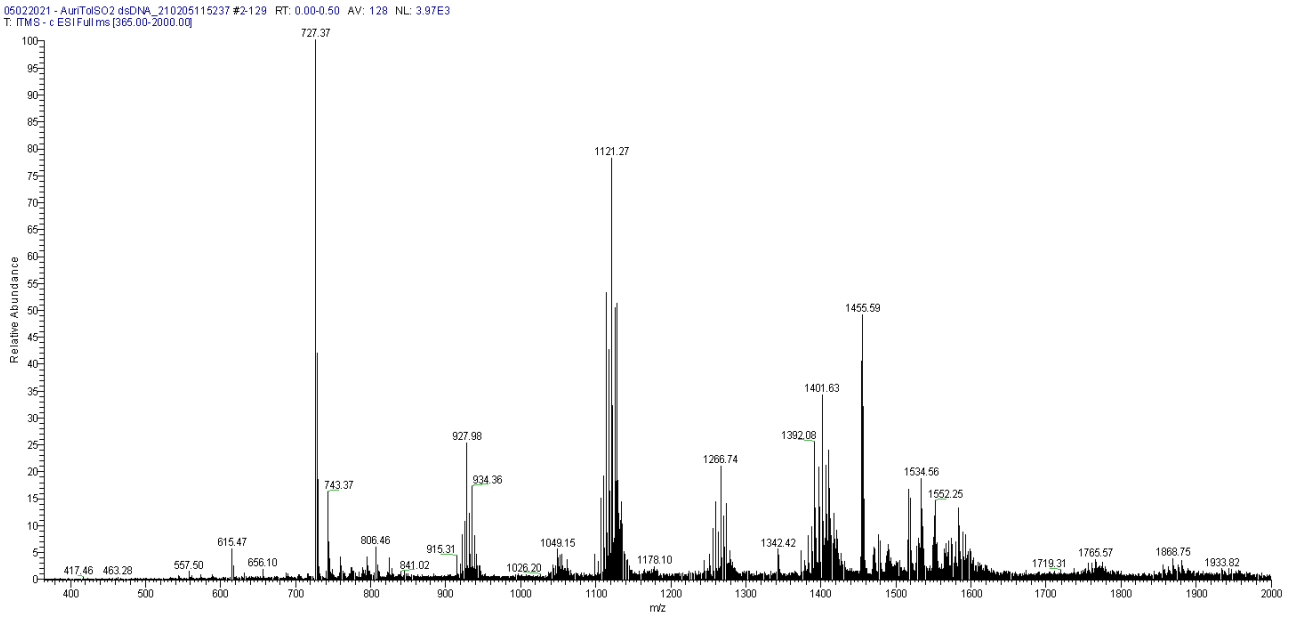


Figure 1.8. ESI-MS spectrum of an interaction assay between a potential ligand and G4 DNA. The signals at m/z = 927.98, 1121.27 and 1401.63 correspond to the unbound DNA (z = -12, z = -10 and z = -8, respectively), the signal at m/z 727.42 corresponds to the unbound ligand. The interaction peaks were detected at m/z = 1049.15, 1266.74 (ligand:DNA 2:1, z = -12 and z = -10, respectively).

The binding affinity (BA) represents an important semiquantitative parameter that determines the affinity of a ligand for a certain receptor, which in our case is the nucleic acid in its different possible arrangements. More in the detail, basing on the signal intensities retrieved from the MS experiment, the formula reported here below, weights the ration between the amount of the complexed receptor and the sum of the amounts of all the forms of the receptor (free and complexed).

$$BA = 100 \times \frac{\sum I_{GQ\ bound}}{\sum I_{GQ\ unbound}}$$

Another relevant parameter that can be obtained by this MS approach is the stability of the ligand-receptor complex. More in the detail, for this purpose, Collision Induced Dissociation (CID) experiments can be performed to reveal the relative gas-phase kinetic stability of the complexes.
In a common CID experiment a particular ion is selected, corresponding usually to a single entity or to a complex, then it is accelerated applying an electrical potential that increases its kinetic energy, which is converted in internal energy upon collision with neutral molecules (usually an inert gas). This produces either the fragmentation of a complex in its constituents or the breakage of molecular bonds. Furthermore, the relative intensities of the adduct and of the dissociation products are recorded at increasing collision energies (applied potential), from which the relative intensities I(%) are calculated through a specific formula (which is reported here below) and plotted against the applied potential. From the so obtained graph is possible to obtain a specific energy value $E_{COM}^{50\%}$ expressed in electronvolts. An exemplificative spectrum is reported in the figure below (Figure 1.9).

$$I(\%) = \frac{I_{complex}}{I_{complex} + I_{dissociation\ products}} \cdot 100$$

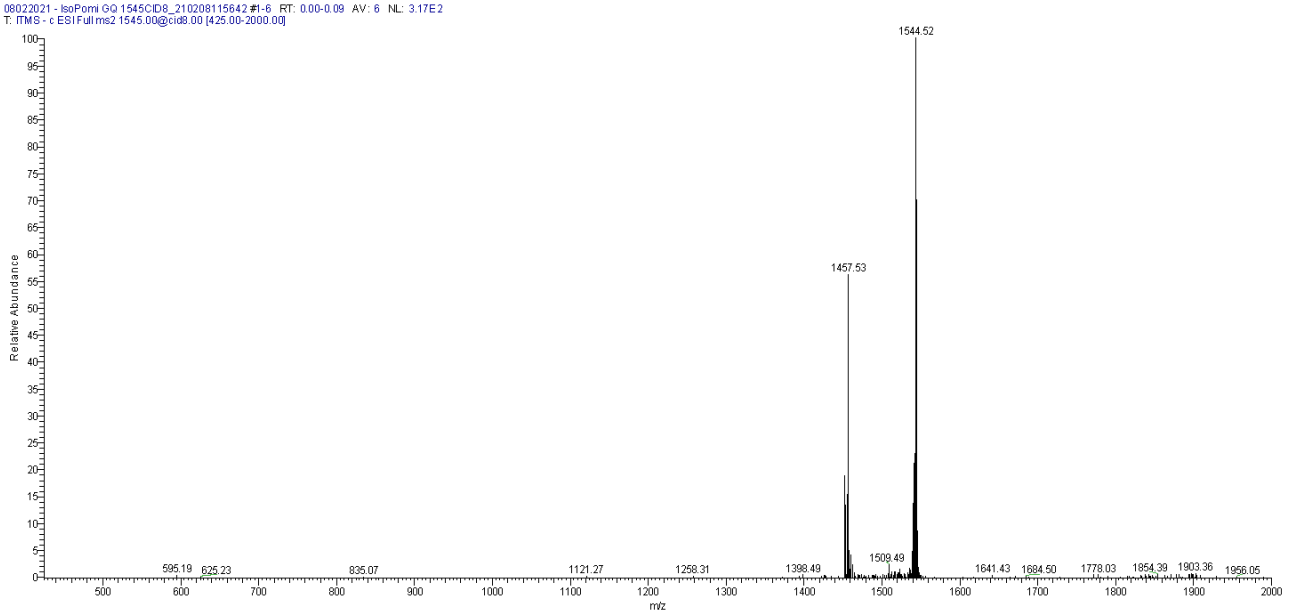


Figure 1.9. CID fragmentation spectrum of a G4 - ligand complex (normalized collision energy = 8). The signal of the complex is still present (m/z = 1544.52, z = -5), the peak at m/z= 1457.53 corresponds to the unbound G4 which starts to appear upon complex dissociation.

Another method to determine the stability of a complex can be performed using the native MS technique obtaining a thermal stabilization parameter T_m for the complex. More in the detail different spectra are acquired at increasing capillary temperatures (typically from 60 to 400 °C), the values are then plotted obtaining a result similar to FRET or CD melting assays.[151–154]

1.4.6 Miscellaneous methods

In the addition to the ones already reported, other methods are commonly employed to study the topology, stability, and formation of G4 motifs. For example, the isothermal calorimetry (ITC) and the surface plasmon resonance (SPR), can quantitatively determine the ligand-receptor binding event, giving its stoichiometry and the associated binding constants.[35,155–161] Furthermore a relevant biochemical assay is the telomere repeat amplification protocol (TRAP) which measures the elongation of a telomeric sequence subjected to specific concentrations of a ligand giving an indirect measure of its inhibitory effect as a consequence of its affinity for the G4 arrangement. The Taq polymerase stop assay reveals the arrest of the DNA synthesis by the thermostable enzyme DNA polymerase I, the quantification of the inhibitory effect of the ligand for the process is as for the TRAP assay an indirect proof of the affinity of a ligand for a nucleic acid. Other relevant techniques are: fluorescence melting,[86,139,162,163] single molecule FRET microscopy,[164] denaturing PAGE, electron

paramagnetic resonance (EPR),[165] ^{125}I-radioprobing,[166,167] fluorescence spectroscopy,[168,169] Raman spectroscopy[170,171] and temperature-jump relaxation experiments.[41,86,102,172,173]

1.4.7 Computational studies

In the past two decades, the use of computer-aided drug design revolutionized the approach for the identification of small molecules targeting different macromolecular assemblies and, more in general, the drug discovery process. Virtual screening (VS) methods are generally used for the simultaneous evaluation of thousands of compounds using structure-based (receptor-based) and ligand-based techniques.[174] To overcome the already cited issues generally present in the wet-lab experiments, in the past years, new *in silico* methodologies and prediction algorithms were developed, and nowadays represent an essential tool for a complete investigation of the structure and the dynamics of most of the nucleic acid arrangements like the G4 which is the focus of this thesis.[175–177]

In the structure-based studies, the structures deriving from X-ray crystallography and NMR spectroscopy are used as three dimensional templates to screen large libraries of compounds,[178,179] furthermore these studies are today the basis for the identification and rational optimization of G4 ligands.[180]

Ligand-based methodologies instead use specific 2D or 3D queries, based on structure similarity or 3D pharmacophore models, to screen databases and to pinpoint the best matching molecules.[181] The pharmacophore models are commonly generated through an automated analysis of the chemical structures and structural features of a compounds library with known inhibitory activities/binding properties, then the resulting 3D models are used as a query to screen a set of compounds, which generally correspond to in-house databases previously created by a particular research group/institution or to freely accessible or commercial online databases (*e.g.* ZINC, DrugBank).[182] This approach was for example used by Li and colleagues to identify novel potential G4 binders screening a library of nearly 10000 molecules with a pharmacophore model generated from 1,4-disubstituted anthraquinone derivatives.[183] On the other hand, in the ligand-based studies field, Alcaro *et al.* reported in 2011, the analysis of conformational properties and of the solvent-accessible surface area of known G4 ligands, these physico-chemical data were then used for the prediction of the ligands' binding activities deriving from the stacking interaction mode with the G4 arrangement.[184]

The study of the interaction between G4 arrangements and their potential ligands is in these days greatly enhanced by different computational methods, among them molecular docking and molecular dynamics (MD) represent the most commonly used strategies to investigate the event of binding between a receptor and a ligand. More in the detail, these kind of simulations relys on molecular mechanics (MM) methods, which are based on particular force fields (ff) that are constituted by

functional forms and parameter sets which regulate the energies arising from the interactions involved.

Molecular docking applies a search and score method. The search algorithm tests all the possible positions and orientations of a ligand for its binding to a receptor, meanwhile all the possible ligand conformations are explored considering as mobile its internal angles, dihedrals and improper dihedrals. The search algorithm can greatly differ among the considered software, the most used are: geometric matching, incremental construction, Monte Carlo, genetic, ff-based or their combination. Following the generation of multiple possible docking poses, each of them are evaluated through a scoring function, which can be either ff-based, empirical, knowledge-based or a combination of these. The score that is assigned is usually expressed in arbitrary units or in some cases it can be correlated to a free-energy value.[174] An increased level of performance can be achieved exploiting what is called a flexible ligand-flexible receptor approach, which can be pursued with different methods, namely the ensemble docking, the induced-fit docking (IF) or the flexible residues docking. The ensemble docking relies on the simultaneous usage of multiple receptor structures, usually obtained from different crystals or solved by NMR of the same receptor or as frames from a MD simulation, typically the bigger is the number of the available structures the higher is the accuracy of the method.[185] The induced-fit docking is usually based on a three steps technique: first the ligand is docked in a binding site with softened potentials than the surrounding residues are allowed to move increasing the interactions and decreasing the clashes and in conclusion the ligand is either scored in place or re-docked obtaining a final binding score, one of the most famous IF method is implemented in Schrödinger-Glide software.[186] The flexible residues docking, on the other hand, offers the possibility to move the residues contextually to the docking process increasing the accuracy but also the computational time, the method is for example implemented in the AutoDock suite. Among the flexible-receptor methods cited above, IF and flexible residues docking are usually developed for protein receptor in specific software, therefore for nucleic acid receptors the most valuable tool is the ensemble docking which doesn't require a high level of scripting but can be performed by the automated consecutive docking to different structures.

MD can be described as an ff application for a system that is allowed to evolve over time from the starting conditions. The whole process happens in a simulated box where specific variables are maintained as constant, these usually belong to a particular ensemble which can be either microcanonical (NVE), canonical (NVT) or isothermal-isobaric (NPT). The system simulated can be either only the receptor or the ligand alone or the complex formed by the ligand binding to a receptor, this is then usually placed in an explicit solvent box and depending on the timescale adopted for the simulation several phenomena can take place and being observed and measured. Relevant measurements can be obtained from MD simulations; the most significant is the root mean square deviation (RMSD) that can be used to assess the overall stability of the system. In particular, it can

be calculated either for only the receptor or to the ligand, in the latter case the values are usually fitted to the receptor RMSD to account for its movement during the simulation.

Other important parameters are the root mean square fluctuations (RMSF) and the radius of gyration (Rg), the first is generally related to the fluctuations of specific residues in the protein simulation context, the latter is a measurement of the compactness of a structure.

In conclusion, it is important to briefly describe the free energy calculations methods. In particular, the Molecular Mechanics Poisson Boltzmann Surface Area (MM/PBSA)[187–189] and the Molecular Mechanics Generalized Boltzmann Surface Area (MM/GBSA) are the most widely used continuum solvation free energy calculations methods applied to ligand-receptor structures deriving from molecular docking or MD calculations.[190] These methods take in account the MM-derived potential energies and the free energy variations involved due to the solvatation events occurring in consequence of the ligand binding to a receptor, in the presence of an implicit water model that can be either Poisson–Boltzmann in the case of MM/PBSA or generalized Born and surface area continuum solvation in the case of MM/GBSA.[190] Having cited these basic concepts about the MM computational methods a focus on their application in the particular field of the interaction of ligands with nucleic acids will be reported in the following.

1.4.7.1 Molecular docking

Molecular docking represents a powerful tool for the computational study of the binding modes of sets of molecules, which can be either under preliminary study prior their synthetic preparation, already synthesized or for drug repurposing. Compared to more sophisticated techniques as for example MD, it has the advantage of requiring relatively low computational resources. This allows to perform simultaneous screening of entire ligand libraries by VS methods. Nowadays different large libraries are currently available as: MayBridge (http://www.maybridge.com.), AnalytiCon (https://ac-discovery.com/screening-libraries/), ZINC (http://zinc.docking.org/), ChemDiv (http://www.chemdiv.com/services-menu/screening-libraries/), SPECS (http://www.specs.net), Mcule (https://mcule.com/database/), eMolecules (https://www.emolecules.com/info/plus/download-database), PubChem (https://pubchem.ncbi.nlm.nih.gov/), Life Chemicals (http://www.lifechemicals.com/), ChemBridge (http://www.chembridge.com/screeninglibraries/).
Different molecular docking software suitable for use with nucleic acid receptors are currently available, even if most of them were originally developed for protein receptors. A brief description of their functioning, together with some applications retrieved from the literature, are reported hereafter.

AutoDock and AutoDock Vina

AutoDock is one of the most widely used docking software and it is composed of two distributions, namely AutoDock 4 (AD4) and AutoDock Vina (Vina), respectively released in 2009 and 2010.[191,192] The biggest difference between them consists in the scoring function that is a semi-empirical, AMBER ff based for AD4 and fully empirical for Vina. In addition to that, Vina demonstrated to be sensibly faster than AD4 by two orders of magnitude.[192] Both AD4 and Vina can be used by command line or are also implemented in graphical user interface (GUI) software like AutoDockTools (ADT). The standalone software PyRx and Raccoon, that use the AD4 and Vina algorithms, possess virtual screening (VS) features. In particular, PyRx allows the simultaneous screening of library of ligands and Raccoon implements ensemble docking of multiple receptor conformations and ligand libraries automatic processing.

Alcaro *et al*. identified a psoralen derivative as a new G4 ligand through the docking screening of 2.7 million molecules of the ZINC database, in particular the molecules were subjected to an ensemble docking considering the 4 most representative quadruplex arrangements namely the parallel (PDB ID: 1KF1), antiparallel (PDB ID: 143D), the hybrid-1 (PDB ID: 2HY9) and the hybrid-2 (PDB ID: 2JPZ).[193] Cosconati *et al*. performed a structure based VS in tandem with NMR experiments finding six new G4 groove binding compounds. The docking was performed with AutoDock 4, screening 6000 compounds of the Life Chemicals database and using as a receptor the G4 formed by the sequence $[d(TGGGsGT)]_4$ (PDB ID: 1S45). Applying a filter based on the ligands binding energy and discarding the poses not showing peculiar interactions, 30 final hits emerged. NMR titration experiments were then used to eliminate 8 false positive and to find 6 real G4 groove binders which binding ability was also confirmed by ITC experiments.[178,194] Ranjan *et al.* reported the binding of amino sugar-intercalator conjugates with anti-parallel G4 (PDB ID: 156D) derived from *Oxytricha nova*, in particular four conjugates between neomycin and common intercalators with different surface areas were studied. Typically, neomycin can recognize G4 with moderate to high affinity and its benzoquinoquinoxaline (BQQ) conjugate BQQ-neomycin displayed the best binding to the G4 structure with an association constant (K_a) of $1.01 \pm 0.03 \times 10^7$ M^{-1}, which is nearly 100-fold higher than the binding of neomycin to the same G4. The molecular modeling part of the study, performed with AutoDock Vina, revealed for all the conjugates that molecule interacts with G4 in a multiple mode: the neomycin part is positioned in the wide groove while the linker extends the intercalating moieties more towards the thymine loop regions. Moreover, only for the smallest polycyclic ring, the anthraquinone, stacking interactions was observed.[195]

Schrödinger Glide

Glide (Grid-based ligand docking with energetics) is the docking engine of the Schrödinger suite, it relies on an exhaustive search algorithm starting with an initial rough positioning of the ligand in its binding site, followed by a scoring calculation and by a torsionally flexible energy optimization under the OPLS (Optimized Potentials for Liquid Simulations) ff, nowadays at its 4th version, namely OPLS4.[196] The best candidate poses are further refined via a Monte Carlo sampling followed by a last selection of the best docked pose using a model energy function that combines empirical and ff based terms.[197] Glide offers three different precision settings intended to be used for different docking approaches, from VS to complex structural studies, namely HTVS (High Throughput Virtual Screening), SP (Standard Precision) and XP (Extra Precision), listed in an order of increasing precision. Moreover, specific VS tools are embedded in Schrödinger suite like XGlide and Virtual Screening Workflow.

A structure-based VS was conducted by Artese *et al.* on a set of 31 k natural compounds.[198] The study included the development of a single pharmacophore hypothesis on already solved complex structures of ligand-telomeric G4 complexes. After the validation of the protocol, using specific decoy sets, a VS was conducted through the Glide HTVS docking protocol. This allowed obtaining 12 final hits for which the chemical scaffolds were already associated with G4 binding properties and antiproliferative effects. Kar and colleagues based a VS on a two-step Glide docking protocol consisting in a preliminary SP precision screening followed by a more time consuming XP precision re-docking of the top-scoring ligands.[199] More in the detail, the NMR solved structure of the human telomeric G4 sequence $TAG_3(T_2AG_3)_3$ (PDB ID: 2lD8) was used prior placement of two potassium ions between the three G-quartets. Potential binding sites were identified with the SiteMap application embedded in Schrödinger and were targeted in a VS of 14 k molecules included in the Maybridge database.[200] Two G4 ligands, showing selectivity for GC-rich DNA over dsDNA, were revealed by the screening and their ability to bind to the G4 arrangement was confirmed by fluorescence titrations. Wang *et al.* studied the ability of indenoisoquinoline topoisomerase inhibitors to bind and stabilize the c-myc promoter G4 downregulating c-myc. The docking protocol consisted in a Glide SP precision docking of 7-azaindenoisoquinoline molecule to a c-myc G4 NMR resolved structure (PDB ID: 2L7V). This approach produced binding poses that resembled those, previously obtained, of quindoline in the NMR structure of the 2:1 quindoline-G4 complex, in which a flanking DNA base from the 5′- or 3′-flanking segment was recruited to form a ligand-base plane stacking over the external tetrads (Figure 1.10).[201,202]

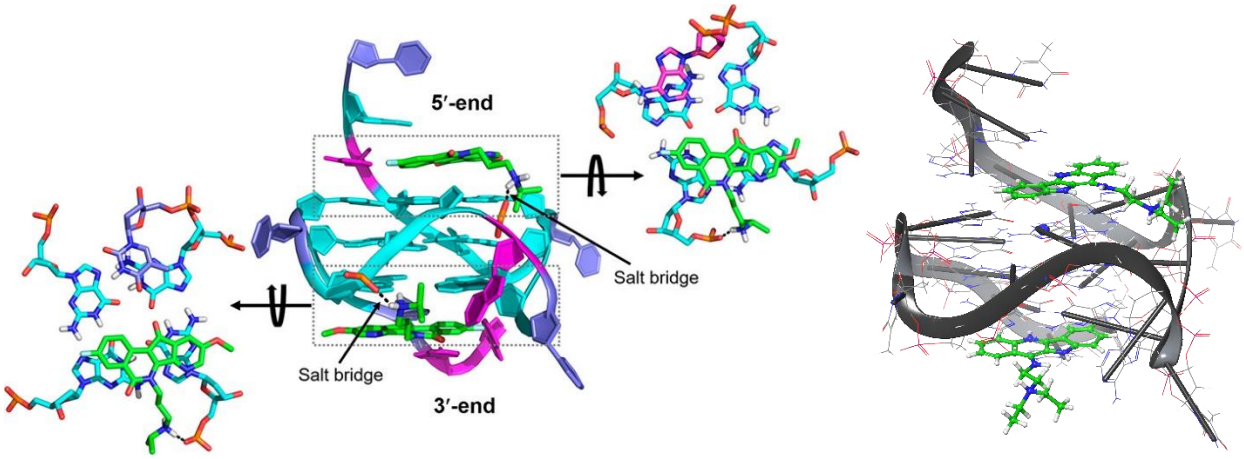


Figure 1.10. Model of the 2:1 complex of 7-azaindenoisoquinoline **5** and c-myc G4 obtained by Glide docking (left).[201] NMR resolved structure of quindoline c-myc G4 complex for comparison (right) (PDB ID: 2L7V).

UCSF DOCK

The docking software UCSF DOCK relies on an ff-based scoring function and on a rigid receptor-flexible ligand geometric matching (GM) algorithm. This last uses a sampling algorithm called anchor-and-grow which "builds" the ligands into the receptor's active site. More in the detail, the largest rigid scaffold of the ligand, namely the anchor, is identified, placed and oriented in the binding site, then the flexible portions of the ligand are systematically added to the anchor re-creating the whole molecule. Starting from DOCK 6, an AMBER based MD engine was implemented to allow receptor flexibility, therefore including rank ordering by energetic ensembles in the docking calculations.

Kang *et al.* identified three new c-myc G4 stabilizing ligands using a combined approach consisting in filtering 560 k compounds of the libraries ChemDiv and SPECS using a pharmacophore search algorithm, then the refined set of molecules was docked with DOCK (v5.4) to the TMPyP4 binding site of the NMR solved c-myc G4 complex (PDB ID: 2A5R). The grid-based flexible docking results were re-scored through the MMGB/SA scoring method. Three final compounds with different chemical scaffolds showed selective PCR-arresting effect, inhibition of c-myc transcription and decrease of promoter activity by binding to G4 in the promoter region without conformational change of a parallel G4.[203]

MolSoft's ICM

ICM is the docking software developed by MolSoft; its scoring function is based on the all-atom vacuum ff ECEPP/3 with appended terms to account for solvation free energy and entropic

contribution. On the other hand, the search algorithm consists of a biased probability Monte Carlo (BPMC) search algorithm.

ICM was used by Lee and co-workers to screen the 20 k molecules of the AnalytiCon Discovery GmbH library using as a receptor a modified version of the multiple-times-already-cited parallel telomeric G4 (PDB ID: 1KF1). In particular, the G4 motif model was built, starting from the abovementioned parallel G4 which was modified by base replacement or deletion, missing loop nucleotides were added using single-strand B-DNA geometry and potassium ions were placed between the G-tetrad planes to stabilize the overall structure.[204]

The five highest scoring compounds were tested in a polymerase stop assay identifying the natural product fonsecin B as a stabilizing ligand for c-myc G4 DNA. The obtained docking pose revealed that the relatively flat scaffold is stacked on the G-quartet at the 3' terminus, in this model the phenolic and carbonyl oxygen atoms are situated close to the central potassium possibly producing favourable electrostatic interactions.

1.4.7.2 Molecular dynamics simulations

MD simulations can be described as the over-time application of a specific ff to a system of atoms composed of small molecules (*i.e.* ligands, ions and solvents) and also macromolecules as proteins and nucleic acids. The most widely used ffs are called additive as they consider charges as fixed and centered on atoms. On the other hand, polarizable ffs add polarizable dipoles to atoms, so that the charge depends also on the environment. The ff selection is crucial for MD simulations of DNA and in particular of G4. In fact, only few ffs can accurately simulate these particular arrangements. Almost every ff can be easily implemented in different MD software. Among them, the most famous software are GROMACS, AMBER and DESMOND (www.gromacs.org; https://ambermd.org; www, https://www.Schrödinger.com).

CHARMM27 all-atom additive ff for nucleic acids[205] and its evolution CHARMM36 all-atom additive ff for nucleic acids[206] are the specific CHARMM ffs for the simulation of nucleic acids, even if they demonstrated to perform well with the B-DNA arrangement, their efficacy with G4 is not fully established.

Fadrna *et al.* tested the CHARMM27 ff for simulating two telomeric G4 arrangements, namely the antiparallel $d(G_4T4G_4)_2$ dimeric quadruplex with diagonal loops (PDB ID: 1JRN) and the parallel human telomeric monomolecular quadruplex $[AGGG(TTAGGG)_3]$ (PDB ID: 1KF1). In both cases, the ff mostly failed and produced a substantial instability. CHARMM36 was used as itself or in combination with the Drude polarizable ff to simulate different G4 structures, in particular c-kit1, c-kit2, c-kit* and bcl2 promoters. In all cases, the polarizable version of CHARMM36 ff demonstrated

a superior reliability in reproducing the experimental structures. The simulations performed with CHARM36, instead, suffered of inadequate ion interactions and instability, and the expulsion of one of the central ions was also observed.[207–209]

AMBER parmbsc0 ff is a modification of the parm99 ff released in 2007 by the Orozco research group,[210] in particular this ff puts the emphasis on the correct representation of the *α*/*γ* concerted rotation in nucleic acids. As their creators reported, the ff was derived by fitting the models to high-level quantum mechanical data, verified by comparison with high-level quantum mechanical calculations and by the comparison between simulations and experimental data. Moreover, the validation study included long MD simulations and a large variety of nucleic acid structures.[210] In 2015, the same research group reported an upgraded version called parmbsc1 including the modification present in parmbsc0 and additional modifications to the sugar pucker, the χ glycosidic torsion, and the ε and ζ dihedrals.[211]

Another important ff for DNA simulation is OL15 which is based on the parm99/bsco with other following modifications on the χ, ε/ζ and β dihedrals of the sugar-phosphate backbone, it is nowadays alongside bsc1 the most used ff for the simulation of DNA arrangements.[212]

Concerning the parametrization of a possibly present ligand complexed with a DNA structure, this is usually accomplished using generic force fields compatible with the ones used for the receptors, namely CHARMM General Force Field (CGenFF) for CHARMM ff and GAFF/GAFF2 for AMBER ffs.

Using OL15 and GAFF2 ff for G4 and for the ligand respectively, Macchireddy *et al.* proved the binding mode of BRACO-19 to G4 by free-ligand MD simulations.[213] The most stable binding mode was identified as the end stacking and, among the three screened G4 topologies namely parallel (PDB ID: 1KF1), anti-parallel (PDB ID: 143D) and hybrid 1 (PDB ID: 2HY9), the MM/GBSA binding energy analysis suggested that binding to the parallel scaffold was the most energetically favourable for the ligand BRACO-19. Moreover, the binding mode obtained using as a receptor the telomeric intramolecular parallel G4 structure (PDB ID: 1KF1) was consistent with a similar intermolecular parallel G4 structure complexed with BRACO-19 (PDB ID: 3CE5).

A non-trivial topic in the MD simulation of the G4 arrangement is the parametrization of the ions present in the G4 central channel. In particular, the choice of the Lennard Jones parameters can play a central role in the stability of the structures and in the ions retention. Havrila *et al.* tested different DNA and RNA G4s (PDB IDs: 1KF1, 31BK, 1K8P, 143D, 2KF8, 2KM3, 2HY9, 2JPZ, 2MBJ and 3QXR) with OL15 and RNA specific ff.[214] The Joung-Cheatham (JC) SPC/E K^+ parameters showed remarkable performance for the majority of the simulated systems.[215] The choice of the water model is also important and should be done taking into account the specific ffs used for the receptor and for the ions. For G4 simulations, the most employed water models are the three site models SPC/E and TIP3P typically in combination with JC ions parameters.[215] Moreover also the

four site water model TIP4P$_{ew}$ demonstrated to perform extremely well alongside JC ions parameters.[216]

$$V_{\mathrm{AMBER}} = \sum_{\mathrm{bonds}} k(r-r_{eq})^2 + \sum_{\mathrm{angles}} k(\theta-\theta_{eq})^2 + \sum_{\mathrm{dihedrals}} \frac{V_n}{2}[1+cos(n\phi-\gamma)] \; / \\ + \sum_{i<j}\left[\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^{6}}\right] + \sum_{i<j}\left[\frac{q_i q_j}{\varepsilon R_{ij}}\right]$$

$$V_{\mathrm{CHARMM}} = \sum_{\mathrm{bonds}} k_b(b-b_0)^2 + \sum_{\mathrm{angles}} k_\theta(\theta-\theta_0)^2 + \sum_{\mathrm{dihedrals}} k_\phi[1+cos(n\phi-\delta)] \\ + \sum_{\mathrm{Urey-Bradley}} k_u(u-u_0)^2 + \sum_{\mathrm{impropers}} k(\omega-\omega_0)^2 + \sum_{\phi,\psi} V_{CMAP} \\ + \sum_{\mathrm{nonbonded}} \varepsilon\left[\left(\frac{R_{\min_{ij}}}{r_{ij}}\right)^{12} - \left(\frac{R_{\min_{ij}}}{r_{ij}}\right)^{6}\right] + \frac{q_i q_j}{\varepsilon r_{ij}}$$

Figure 1.11. Amber and Charmm ff functional forms (the OPLS ff functional form is almost identical to the Amber's one), for both of them the first summations calculate the potential energies associated with the ligand covalent bonds (bond length, angles and dihedrals) while the last ones regard the potential energies arising from Lennard Jones and charge-charge interactions. The image was obtained from Amber20 manual.

1.5 Ligands stabilizing the G4 structure

As already discussed, the G4 structure highly depends on the different topologies, which it can adopt, and this must be taken in account when designing new molecules able to target this DNA arrangement. The different G4 topologies share similar binding sites, namely by stacking on the external G-quartets or by groove recognition. The stacking mode that may eventually involve interactions with surrounding loops is mainly the feature of planar rigid aromatics. Alternatively, insertion in the grooves of quadruplexes, that possess four grooves, is to be considered, especially for nonplanar, flexible molecules.[217] When thinking to the development of a new G4 ligand, the external-stacking interaction mode is a crucial feature, instead the groove interaction appears as important but not always necessary (an example is telomestatin which will be described later), however most of the ligands use both of them in conjunction to recognize the G4 motif with high selectivity against the dsDNA.

Briefly, the recognition of the G4 arrangement by a ligand molecule highly depends on mostly two types of interaction namely the pi-stacking and the ion pairing electrostatic interactions.

The pi-stacking interaction happens between the ligand aromatic rings and the guanine aromatic units, namely the pyrimidine and imidazole rings.[218] This particular interaction has an electrostatic nature and can be described by a preferential spatial orientation assumed by two similar quadrupoles (in terms of charge distribution) when in proximity, which generally eschew face-centred parallel stacking in favour of perpendicular edge-to-face interactions or off-centred parallel stacking. This

widely accepted model, deriving from the studies of Hunter and Sanders in the early 1990, predicts a preference for face-centred pairing, when an aromatic pair is constituted by an electron-poor and an electron-rich ring. The different models of interacting quadrupoles are represented in the following figure (Figure 1.12).[219–223]

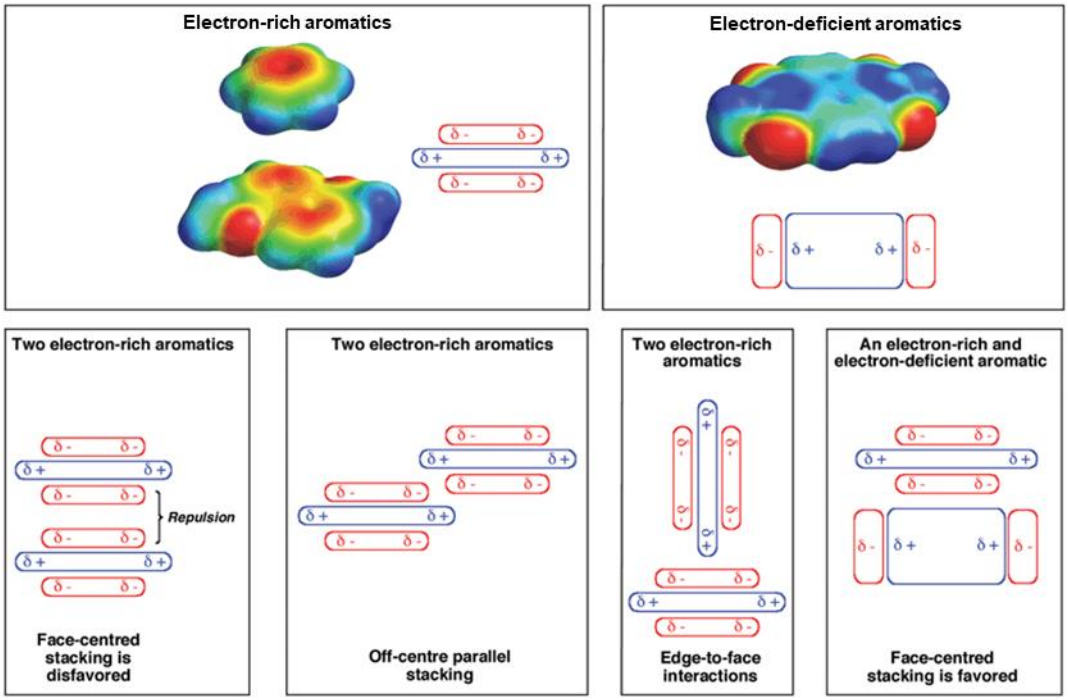


Figure 1.12. Different kinds of stacking interaction between quadrupole moments of aromatic rings.

The ion pairing interaction takes place in the sugar-phosphate backbone of the G4's grooves and is important to increase the overall binding energy of a ligand for the G4. The electrostatic interaction generally takes place between the nucleotides' negative charged phosphate ions and positively charged moieties on the ligands, which are usually amines. Due to the long-range nature (decreasing by r^{-2}) and their non-directionality, the ionic interactions usually favour the binding event to the G4 structure; however they don't possess specificity for the G4 motif and can increase the affinity also for the canonical DNA arrangement which is also negatively charged in its major and minor grooves.[32,224]

The parallel topology offers an upper and a lower planar surfaces formed by guanine tetrads and therefore ideal for pi stacking interactions, the polar, negatively charged, loops instead give rise to electrostatic interactions with positively charged moieties. A particular feature of this arrangement is its intrinsic propension to directly accept the binding of a ligand, usually doesn't requiring a consistent structural adaptation. Instead, other G4 topologies as the hybrid (1, 2) or the antiparallel usually present the guanine tetrads surfaces mostly masked by the loops, therefore needing even large structural modifications upon ligand binding by the stacking mode, also possibly leading to a variation

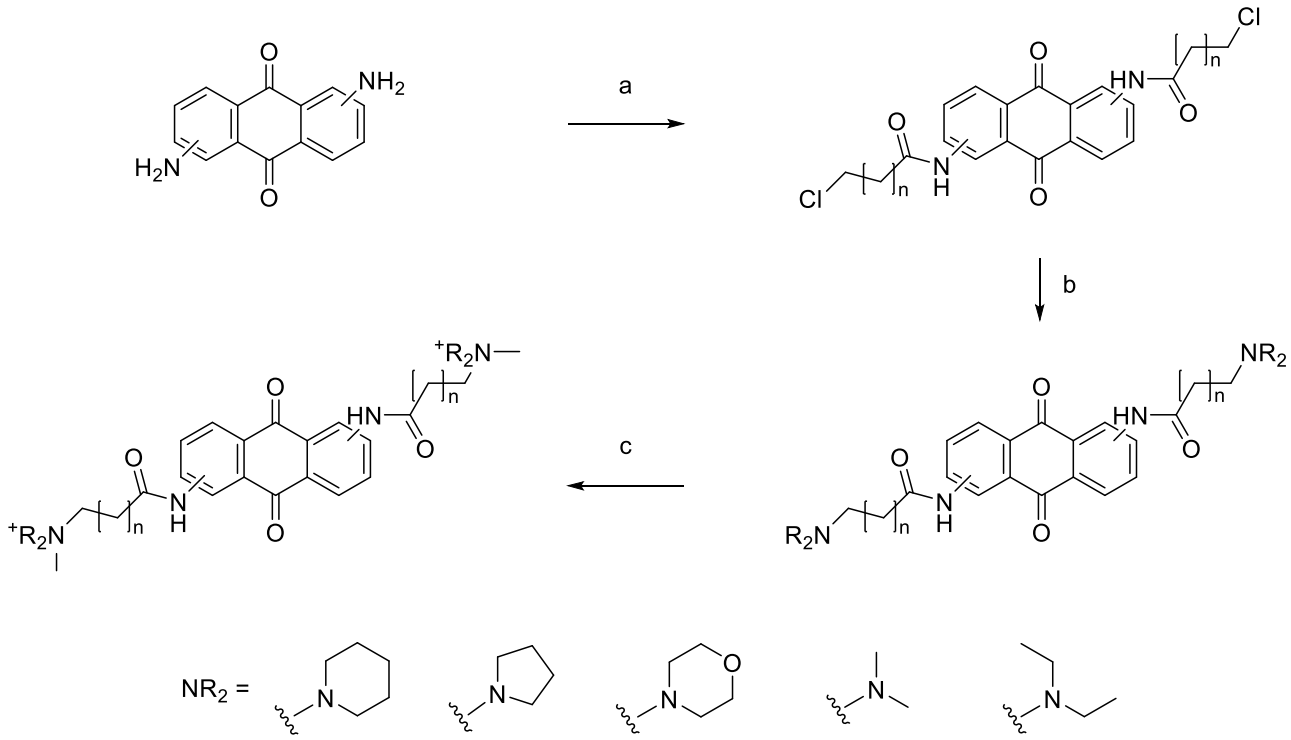
of the starting topology. Indeed, for all these folds, including the parallel fold, it is always possible to think to a ligand driven induction of their formation, however the exact mechanism remains unknown due to the impossibility to directly observe the process or to simulate it by computational methods.[58]

1.5.1 Anthraquinones

Anthraquinones or anthracenediones (AQ) are an interesting planar scaffolds for the development of selective and multifunctional ligands of the G4 DNA, they ability to bind to the nucleic acid was extensively studied in the literature, moreover they possess low redox potential and the ability to photosensitize through a one-electron oxidation reaction.[225] These molecules are strictly related to the natural compounds doxorubicin and daunorubicin which are known dsDNA interacalants.[226]

The first AQ telomeric G4 ligand was reported by Sun and co-workers in 1997, in particular a 2,6-diamidoathraquinone derivative demonstrated to have human telomerase inhibition properties, the study also included a NMR thermal denaturation experiment which indicated a drug-induced increase in the G4-DNA melting temperature (T_m) by about 20 °C.[227] In 1998, Perry and co-workers reported the preparation of a number of 1,4- and 2,6-difunctionalized amidoanthracene-9,10-diones. In particular, among them there were found the most potent non-nucleoside telomerase inhibitors showing IC_{50} levels of 4-11 μM.[228]

In the same year, still Perry and colleagues reported a second set of anthraquinone derivatives able to interact with the human telomeric G4 DNA, in particular di-functionalized amidoanthracene-9,10-diones substituted at the 1,5-, 1,8-, and 2,7-positions were prepared.[228] The synthesis of the compounds provided for three simple steps, namely an initial amidation of the aryl amines of the scaffold with chloro-propionyl chloride which was followed by a nucleophilic substitution reaction on the intermediate's halide with different secondary amines (*i.e.* dimethylamine, morpholine, pyrrolidine) and a final alkyl amine methylation yielding quaternary ammonium salts (Scheme 1.1).



Scheme 1.1. Synthetic scheme for the preparation of di-amido anthraquinones. a: $ClCO(CH_2)_{n=1,2}Cl$; b: HNR_2, EtOH; c: iodomethane, acetone.

The amidoanthraquinones showed only moderate cytotoxicity against a panel of three ovarian cell lines and were generally 2−3 orders of magnitude less potent than the drugs doxorubicin and mitoxantrone, moreover for each isomeric series of compounds, good telomerase inhibition activities were determined, with IC_{50} values in the range of 1.3−17.3 μM. For two compounds of the set, an isothermal titration calorimetry (ITC) assay was performed in order to determine the binding constants related to the binding of the ligands to the human telomeric sequence $[AG_3(T_2AG_3)_3]$, the values obtained were up to 7.9×10^4 M^{-1}. Furthermore, the study investigated the binding event also through computational techniques, the solution NMR structure of the human telomeric repeat $[AG_3(T_2AG_3)_3]$ G4 was modified forming an intercalation site between the diagonal T_2A loop and the G4 segment. The various ligands were then minimized, docked into the intercalation site, and the binding energies were determined.

In 2008, Zagotto and colleagues, developed a series of anthraquinone-amino acid conjugates as novel ligands for the G4 (Figure 1.14), also investigating the ability of the amide bond direction to modulate the G4 recognition and telomerase inhibition. For this purpose, a set of molecules was prepared including the amides deriving from the 2,6/7-diamino anthraquinone derivatives (2,6-NH_2AQ) acylation and from the inverse 2,6/7 dicarboxylic acid anthraquinone derivatives (2,6-COOHAQ) amination.[229,230]

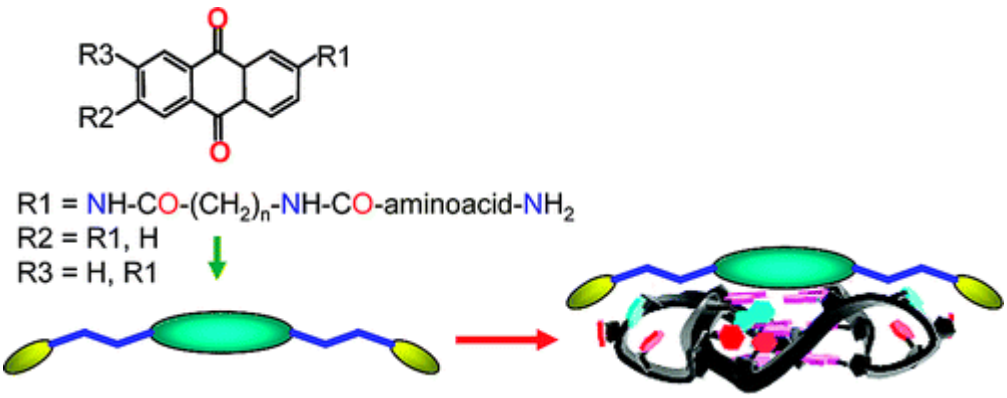
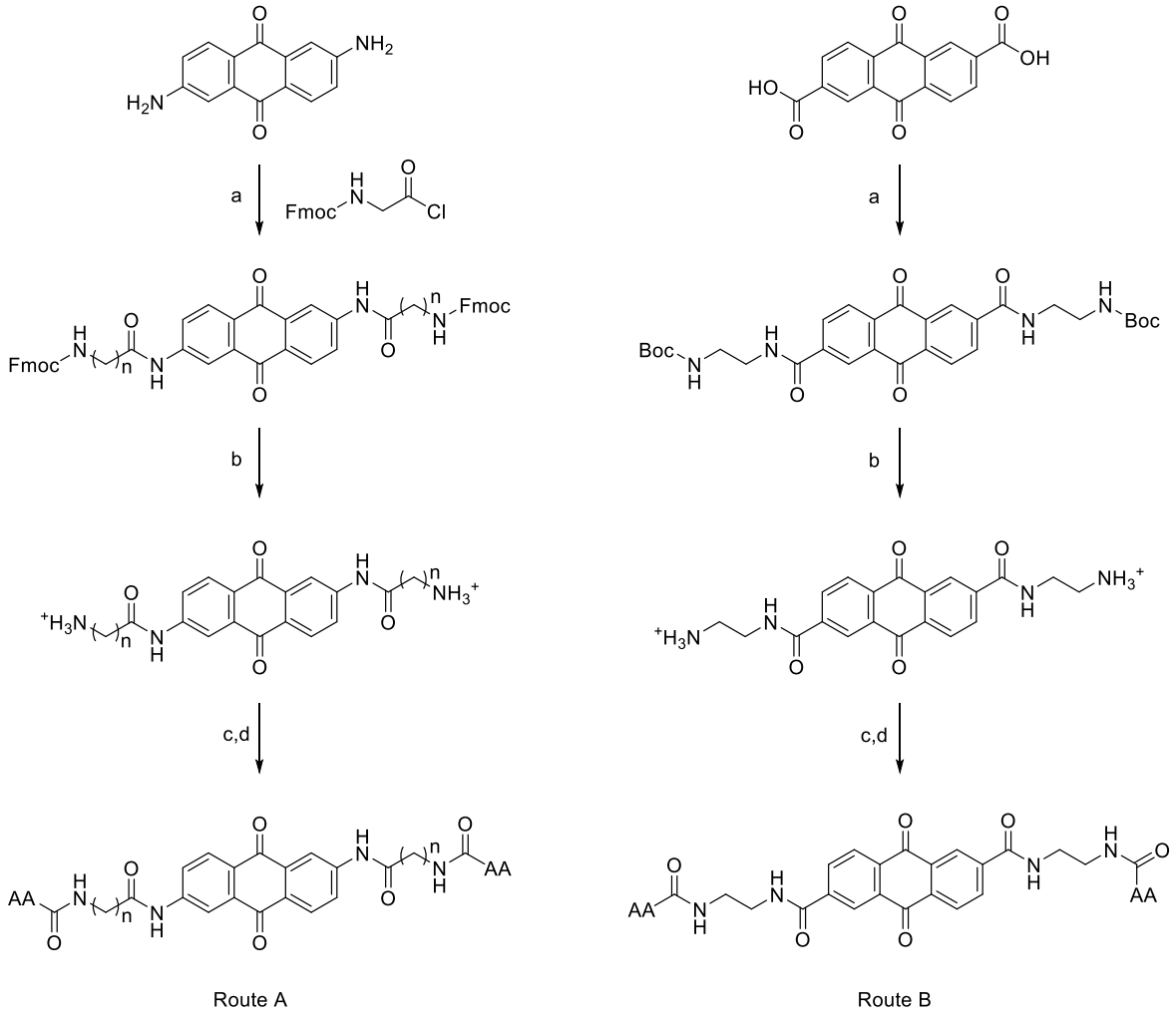


Figure 1.13. General scheme for the development of amino acid - anthraquinone conjugates interacting with G4 arrangement by stacking interaction through the anthraquinone rings and groove recognition by the charged side-chains.

The synthesis of 2,6-NH_2AQ derivatives set provided for an initial acylation of its two aryl amines with Fmoc-protected glycine (Gly) or *β*-alanine (*β*-Ala), then after the Fmoc-deprotection a second amino acid was attached completing the side chains. The preparation of the 2,6-COOHAQ derivatives consisted in the amidation of its two carboxyl groups with mono Boc-protected ethylenediamine followed by Boc-deprotection and attachment of an amino acid of the side-chain. The synthetic schemes adopted are depicted in the following scheme (Scheme 1.2).

Fluorescence melting studies were conducted on the ligands revealing some important aspects for the G4 arrangement stabilization. The introduction of positively charged amino acids like lysine (Lys) or arginine (Arg) generally increased the affinities towards the G4, the effect can be ascribed to the development of relevant electrostatic interactions with the DNA phosphate backbone, on the other hand, apolar amino acids phenylalanine (Phe), leucine (Leu) and isoleucine (Ile) reduced the affinity for the nucleic acid. In terms of ΔT_m, the values obtained were up to 24 °C at no major differences were observed between the 2,6 and 2,7 isomers. The telomerase inhibition activity was then tested for the synthesized compounds obtaining IC_{50} values up to 0.8 µM which was obtained for the 2,6/7-AQ-*β*Ala derivatives having an Arg or a Lys as the side-chain terminal amino acid. Furthermore, all the compounds showed poor short-term cytoxicity against the cell lines HeLa and 293T compared to the reference drug mitoxantrone, in particular the lowest values were obtained for the Lys or Arg containing conjugates, however in a senescence induction assay against HeLa cells, where the treatment was prolonged for 6 days, a marked effect was produced by the highly charged ligands, demonstrating their ability to trigger the envisioned mechanism and to internalize in the cells.



Scheme 1.2. Synthetic schemes for the preparation of amino acid - anthraquinone conjugates, the two schemes represent the synthesis of 2,6-NH_2AQ (Route A) and 2,6-COOH-AQ derivatives (Route B). Route A. a: Fmoc-Gly-Cl or Fmoc-*β*Ala-Cl, THF; b: piperidine, DMF, TFA−water; c/d: various conditions depending on the amino acidic residue. Route B. a: $SOCl_2$, THF then Boc-ethylenediamine, triethylamine; b: TFA, water; c/d: various conditions depending on the amino acidic residue.

As a follow up of the abovementioned work, Zagotto and co-workers reported in 2012 a novel series of aminoacyl-anthracenedione derivatives.[231] These ligands were symmetrically functionalized with N-terminal lysyl residues, incrementing the length of side chains by introducing a Gly, Ala, or Phe spacer, characterized by different flexibility, lipophilicity, and bulkiness. Moreover, 2,6, 2,7, 1,8, and 1,5 regioisomers were examined providing a small library of bis(lysyl-peptidyl) anthracenediones.

The synthesis of all the ligands started from the relative diamino-substituted anthraquinones. While 2,6-diamino and 1,5-diamino AQ are commercially available, the 2,7-diamino and 1,8-diamino AQ were obtained by reduction of the corresponding dinitro derivatives. The first amino acid of the sidechains protected on the amine with the Fmoc group (Fmoc-glycine or Fmoc-*β*-alanine) was introduced on the anthraquinone's aryl amine, using the appropriate acyl chloride derivatives. The

first group of compounds was then prepared by Fmoc deprotection of the amino group and salification with trifluoroacetic acid (TFA). A second and a third group were prepared by Fmoc deprotection followed by standard peptide coupling with one or two amino acids. A total of 17 final compounds were obtained in this way.

A set of assays were conducted on the ligand library to assess their ability to bind and stabilize the G4 DNA. In particular, a FRET fluorescence melting showed a general higher stabilization for the G4 DNA than for the dsDNA with ΔT_m values up to 20.1 °C, the ligands substituted at the 1,5 positions and 1,8 positions demonstrated an increased preference for the dsDNA. A CD study revealed the ability of the ligands to promote a shift of the G4 DNA topology from a mixed type to an antiparallel topology; however, variations on the spectra were also noticed when tested against dsDNA, suggesting an interaction event also with this canonical arrangement. Furthermore, to assess the effects of the ligands on enzymatic processes involving DNA, Taq polymerase and telomerase inhibition assays were conducted, demonstrating that generally the inhibition of telomerase-mediated DNA elongation occurred at lower concentrations in comparison to DNA amplification. Moreover, the observed modulation of telomerase elongation showed a high correlation with G4 binding thermodynamics with an almost linear dependence of the IC_{50} values for telomerase inhibition from $\Delta T_m/T_m$.

1.5.2 PIPER

The telomeric G4 ligand PIPER (N,N′-bis(2-(1-piperidino)ethyl)-3,4,9,10-perylenetetracarboxylic acid diimide) characterized by a broad hydrophobic core with two external amine appendages (Figure 1.14) was discovered by Fedoroff and colleagues,[232] however the molecule was already been tested by Liu and co-workers as a selective ligand for G and C-rich DNA.[233] Through a NMR ligand to G4 titration experiment, using the intermolecular G4 motif arising from the $(TTAGGGTT)_4$ sequence, a 1:1 complex was initially formed, then upon continuation of the titration by the addition of more ligand a 2:1 ligand to G4 was reached, this revealed the presence of two possible binding sites with slightly different affinities. Noteworthy was the formation of a stable tail to tail G4-ligand-G4 complex from the $[dTTAGGG]_4$ sequence, moreover using the probably more biologically relevant $[dTAGGGTTA]_4$ sequence, PIPER showed another binding mode consisting in a threading Intercalation mode at the GT step. The biochemical effects of PIPER were then evaluated, the compound demonstrated good telomerase inhibition activity, furthermore in a DNA polymerase assay the compound did not cause general polymerase inhibition, however the addition of the ligand to the polymerase reaction mixture caused a sequence specific polymerase arrest immediately before the G4 site. A marked polymerase arrest was reached at a 10 μM concentration.

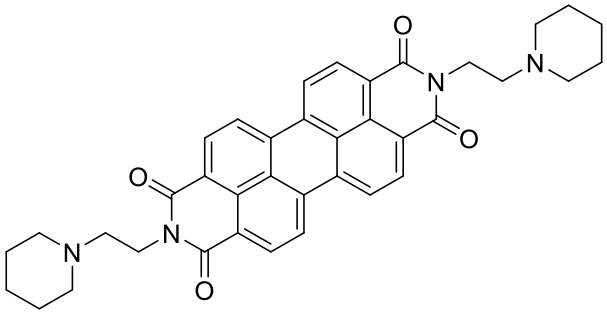


Figure 1.14. Chemical formula of the G4 ligand PIPER (N,N′-bis(2-(1-piperidino)ethyl)-3,4,9,10-perylenetetracarboxylic acid diimide) based on the perylene diimide scaffold.

1.5.3 TMPyP4

The porphyrin derivative TMPyP4 (5,10,-15,20-tetra-(N-methyl-4-pyridyl)porphine (Figure 1.15) developed by Wheelhouse and colleagues is one of the most famous G4 ligands.[234] The molecule was designed to have the appropriate shape and dimensions to stack on top of a G-quartet. A series of binding experiments by means of UV and NMR techniques revealed that TMPyP4 binds to telomeric G4 DNA by stacking externally on the G-tetrads rather than by intercalating between them. Leveraging the photosensitizing ability of the porphyrins, a single-stranded sequence bearing four human telomeric repeats was treated with TMPyP4 and then exposed to light, the assay showed the cleavage of four external guanines, implying the stacking mode of interaction of the compound with the G4 arrangement in consistency with the NMR and UV results. The ability of the molecule to stabilize the telomeric G4 was then measured examining the temperature dependence of the block to Taq DNA polymerase, an increase of the temperature, necessary to unfold the G4 and to allow the enzyme action, of 10 °C was determined between the experiments conducted in presence or not of the ligand. Finally, TMPyP4 was also found to be a potent inhibitor of human telomerase, obtaining an IC_{50} of 6.5 μM in a quantitative cell-free primer-extension assay on an 18-base primer.

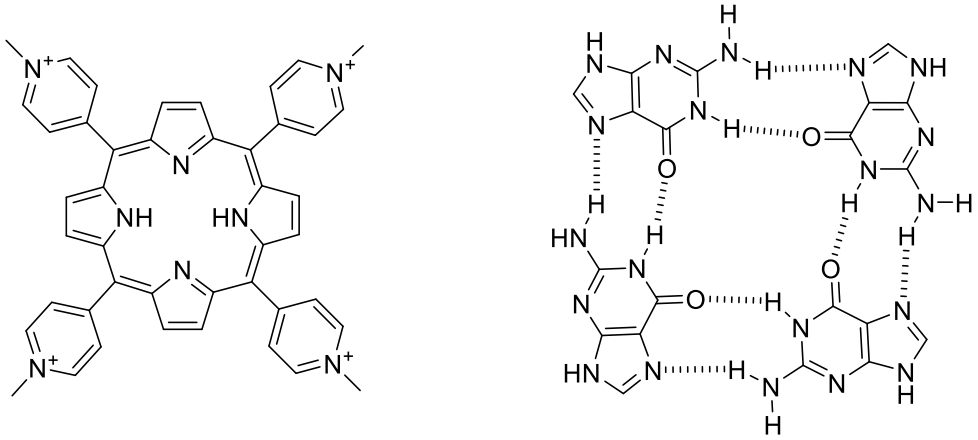
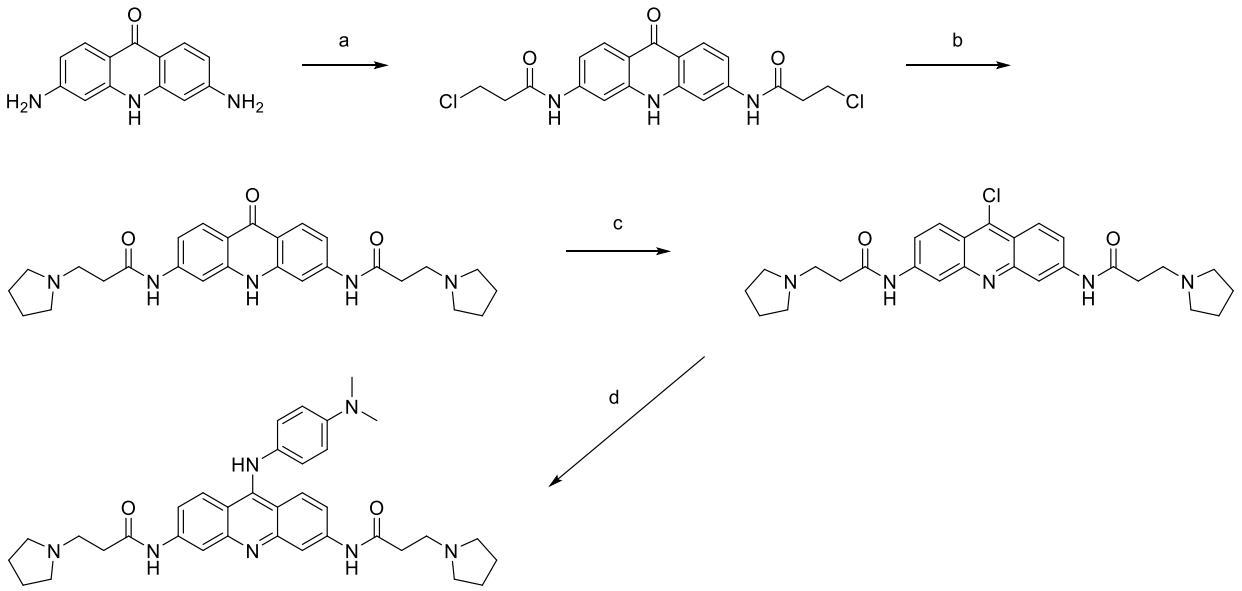


Figure 1.15. Chemical formula of the G4 ligand TMPyP4 and structural affinity with the G-tetrad.

1.5.4 BRACO-19

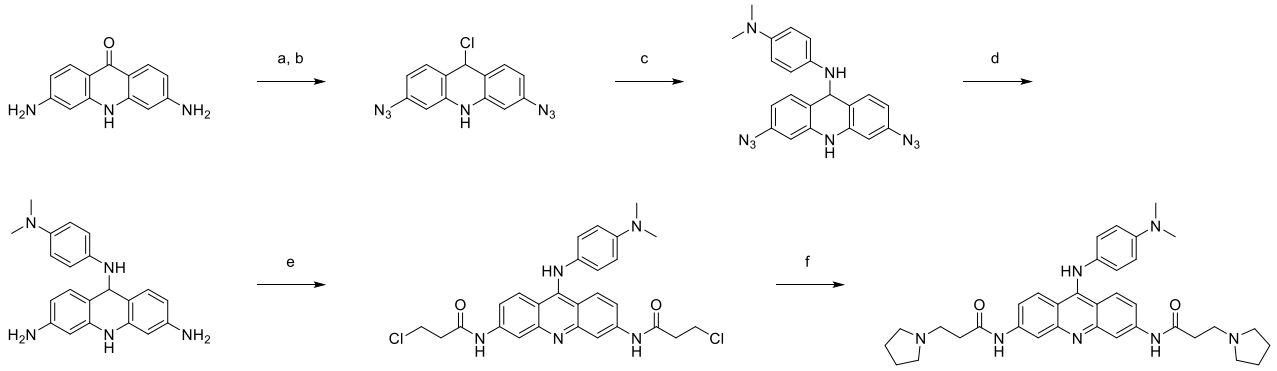
In 2001, Neidle and co-workers developed one of the most well-known G4 ligands, the 3,6-bis(aminoalkyl) substituted acridine BRACO-19. This compound was rationally designed by computer modeling to interact with the G4's grooves in addition to its ability to stack on top of a guanine tetrad. BRACO-19 demonstrated high inhibitory activity in the TRAP assay with a 90 nM IC_{50} and a moderate cytotoxicity towards the uterine carcinoma cell line UXF1138L with an IC_{50} of 2.5 μM, moreover it also showed IC_{50} values of 10.0, 10.1 and 13.0 μM in the telomerase-positive human ovarian carcinoma cell lines A2780, CH1 and SKOV-3, respectively. Its ability to stabilize the G4 arrangement was of 27.5 °C at the concentration of 1 μM.[235–238]

The synthesis of BRACO-19 was accomplished by different research groups here two main pathways reported by the Neidle research group will be showed.[237] The first (9-chloroacridine rute) includes the acylation of diaminoacridone with 3-chloropropionyl chloride followed by amination with pyrrolidine and ketone chlorination with $PCl_5/POCl_3$ which is reacted with N,N-dimethylaminoaniline to obtain the final compound (Scheme 1.3).



Scheme 1.3. Synthetic scheme for the synthesis of BRACO-19 based on the 9-chloroacridine. a: $Cl(CH_2)_2COCl$; b: pyrrolidine; c: PCl_5, $POCl_3$; d: N,N-dimethylaminoaniline, MeOH.

A second higher yielding method is centered on the 3,6-diazido-9-chloroacridine which is easy obtainable through diazotization and treatment with sodium azide of diaminoacridone. This compound can be reacted with the proper aniline to give the 9-substituted diazido acridine intermediate which is subjected to catalytic hydrogenation to reduce the azide moieties at aryl amines that are reacted first with 3-chloropropionyl chloride and then with pyrrolidine to give the final compound (Scheme 1.4).



Scheme 1.4. Synthetic scheme for the synthesis of BRACO-19 based on the diazido chloroacridine. a: $NaNO_2$, dil. H_2SO_4, NaN_3; b: $SOCl_2$; c: aryl amine, cat. HCl, NMP; d: H_2, Pd-C, MeOH, EtOAc; e: $Cl(CH_2)_2COCl$, triethylamine, toluene, MW; f: pyrrolidine.

1.5.5 RHPS4

A novel pentacyclic quinoacridine called RHPS4 (3,11-difluoro-6,8,13-trimethyl-8H-quino[4,3,2-kl]acridinium methosulfate) was reported by the Kelland research group in 2001 and in a second work in 2002 with the synthetic procedures regarding also other prepared derivatives.[239,240] This compound demonstrated to be a potent telomerase inhibitor with an IC_{50} value in the TRAP assay of only 330 nM, whereas in most cell lines, the IC value for cytotoxicity was 20-fold higher at approximately 6 to 8 μM. The binding constants for G4 and duplex DNA were also measured for RHPS4 by fluorescence quenching confirm its selectivity for quadruplex (K_a = 2.25x10^5 M^{-1}) over duplex (K_a = 3x10^4 M^{-1}) DNA, moreover it was able to decrease tumour cell proliferation of breast 21NT cells at concentrations as low as 0.2 μM. In the following figure (Figure 1.16) is represented the chemical structure of RHSP4.

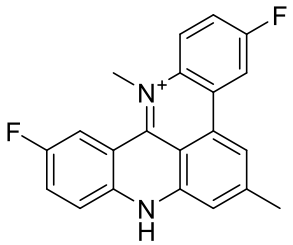
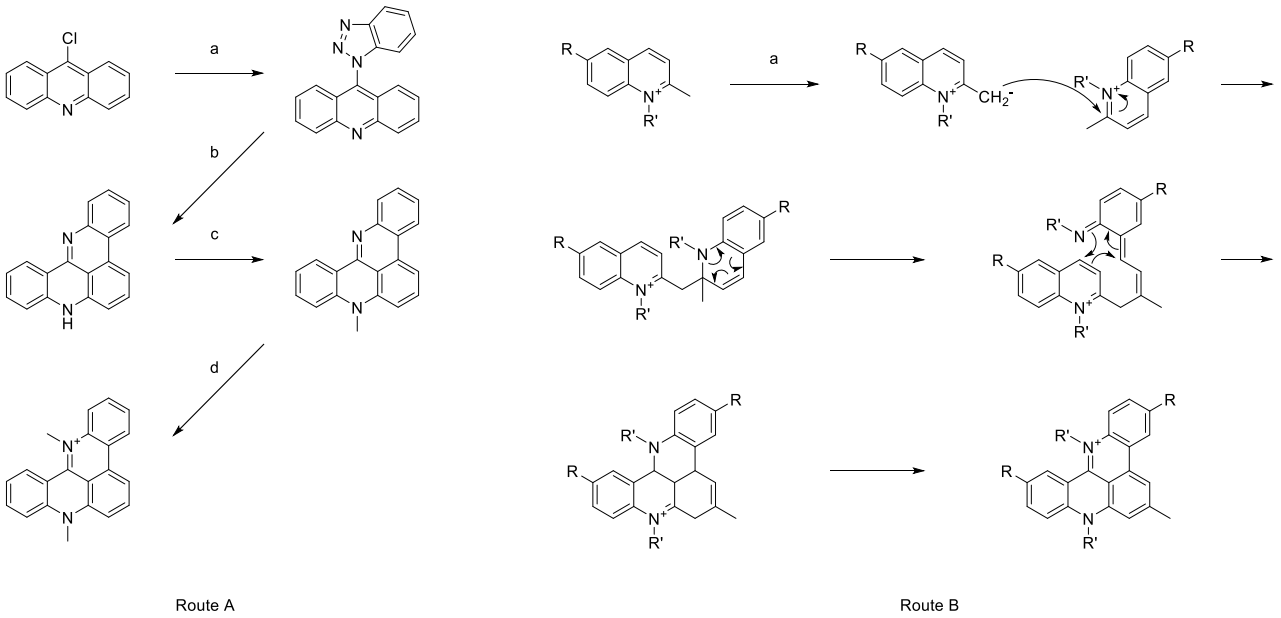


Figure 1.16. Chemical structure of the G4 ligand RHPS4 based on the quinoacridine scaffold.

The synthesis of RHPS4 and other quinoacridine derivatives provided for two different procedures. The first used to synthesize the non-substituted derivatives comprehended the nucleophilic substitution of the 1,2,3-benzotriazole anion on the halide of 9-chloroacridine followed by a thermal *Graebe-Ullmann* reaction yielding a pentacycle intermediate which was two-times methylated on the nitrogen atoms affording the 8,13-dimethyl-quinoacridinium iodide. The second route, even more noteworthy, was first described by Oszczapowicz *et al.*,[241] the proposed mechanism (Scheme 1.5, B) requires the conversion of the 2-methyl group of the initial quinolinium salt to a carbanion by the base piperidine, the carbanion then attacks the electron deficient C-2 of another molecule of quinolinium salt, the opening of the dihydroquinoline ring of the adduct generates an unsaturated system templated for an electrocyclization to the tetrahydro-quinoacridine which upon oxidation, under the reaction conditions, yields the final products.



Scheme 1.5. Synthetic schemes for the preparation of RHPS4, in particular in the depicted schemes are represented the two different strategies adopted to obtain the ligand, one based on the *Graebe-Ullmann* reaction (Route A) and the other on an electrocyclization started by a first intermolecular reaction between two quinoline molecules (Route B). Route A. a: NaH, HOBt, DMF; b: 259 °C in diphenyl ether; c: NaH, dimethyl sulfate, DMF; d: MeI. Route B. a: piperidine, EtOH.

In a following study in 2003, Gavathiotis *et al.* studied more in detail the mode of interaction of RHPS4 with the intermolecular parallel-stranded quadruplex $(TTAGGGT)_4$ formed from the human telomeric repeat.[242] In the titration endpoint of an NMR study, a 2:1 ligand to G4 complex arose, however the formation of the 1:1 complex was also registered at lower amount of ligand. A further investigation by NMR and MD revealed the most plausible structure for the complex which presented one ligand molecules stacked between the upper guanine tetrad and an A-tetrad, and the second molecule stacked on the opposite G-tetrad (PDB ID: 1NZM). The mode of interaction proposed is depicted in the following figure (Figure 1.17).

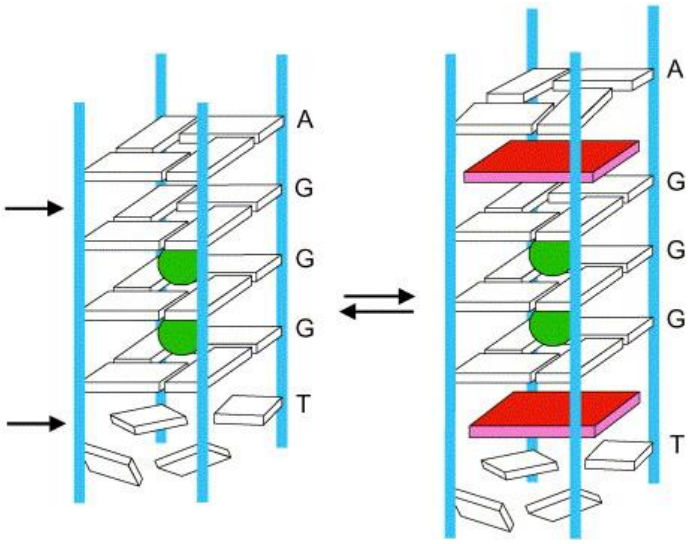


Figure 1.17. Interaction mode proposed for the ligand RHPS4 with the intermolecular parallel G4 arrangement (2:1 stoichiometry) formed by the intermolecular telomeric sequence $(TTAGGGT)_4$. The green balls represent the central cations and the red squares the bound ligands, the latters are positioned either between a G-tetrad and an A-tetrad and between a G-tetrad and an unstructured T-tetrad.

Despite the previously reported resolved structure of the complex between RHPS4 and an intermolecular G4, a high resolution structure of a more biologically relevant, single-stranded telomeric G4, in complex with RHPS4 is still not available in 2021, and for this reason the interaction mode of this important G4 ligand remains elusive. A partial proof of an intercalative binding mode was presented by Mulholland and colleagues in 2017, which investigated the binding mode of RHSP4 with different nucleic acid's arrangement through an exhaustive free-ligand set of MD simulations. In particular, the most stable binding modes were identified for the dsDNA, parallel, antiparallel and hybrid G4 as intercalation, bottom stacking, top intercalation and bottom intercalation, respectively. The intercalation binding mode observed for almost all the DNA motifs, explained the lack of binding selectivity of RHPS4 for G4 over dsDNA and suggests that a ligand modification that destabilizes the duplex intercalation mode but stabilizes the G4 intercalation mode would be crucial to improve the binding selectivity toward G4.[243]

1.5.6 Telomestatin

The Japanese research group leaded by Haruo Seto discovered the natural product Telomestatin, obtained from *Streptomyces anulatus*, as a novel telomeric G4 ligand and telomerase inhibitor (Figure 1.18, left). In a TRAP assay the molecule showed a high ability to inhibit the telomerase activity with an IC_{50} value of 5 nM, whereas it did not show activities against DNA polymerases such as Taq polymerase.[244] Its scaffold is a neutral polycyclic compound, composed of five oxazoles, two

methyloxazoles, and one thiazoline ring, characterized by high hydrophobicity, near complete flatness, and a perfect size shape that provides effective overlap upon stacking on top of the G-quartets, furthermore the structural similarity between telomestatin and a G-tetrad suggests that the telomerase inhibition might be due to its ability either to facilitate the formation of or trap out preformed G4 structures, and thereby sequester single-stranded $[T2AG3]_n$ primer molecules required for telomerase activity. A following structural study was conducted on telomestatin by Kim *et al.*, in the first experiment various concentrations of an oligomer with four repeats of the human telomeric sequence $[T_2AG_3]_4$ were each incubated with increasing concentrations of telomestatin. At high concentrations of telomestatin an intramolecular basket-type G4 was detected, indicating that telomestatin promotes or stabilizes the formation of the intramolecular G4, in particular, at DNA concentrations of 0.005 and 0.2 μM, EC_{50} values of 0.03 and 0.53 μM telomestatin were found. In a parallel experiment with the mutated oligonucleotide [T2AGAG], there was no conversion of the mutated sequence to a G4 structure by telomestatin. Using a simulated annealing molecular docking protocol, four different telomestatin G4 structures were modelled, namely the two possible 1:1 end-stacked complexes, the 2:1 complex (Figure 1.18, right) and a 1:1 complex in which the telomestatin was intercalated between two tetrads.[245] Telomestatin ability to stabilize the telomeric G4 arrangement was assessed towards telomeric G4 through a FRET assay by De Cian and colleagues, in particular using the F21T sequence ΔT_m values of 23.6 and of 22.8 °C were found respectively when in presence of 100 mM Na^+ or 10 mM K^+, a competition experiment with dsDNA in the same conditions demonstrated the complete selectivity of telomestatin for the G4 over duplex DNA.[246] In addition to its high affinity for telomeric overhangs, telomestatin can induce dissociation of telomere-specific binding proteins (TRF2 and POT1) involved in the telomere capping process and, as result, apoptosis is more rapid compared to simple telomerase inhibitor. Thus, telomestatin was classified as a “telomere disrupting agent”. It is also an effective anti-proliferative agent, with IC_{50} values between 0.1 and 5 μM in a panel of cancer cell lines with no significant effect on healthy cells.[247] Unfortunately telomestatin is also afflicted by some drawbacks, the major is the difficulty to obtain it in big quantities either by a total synthesis approach or by natural extraction,[245] moreover its high hydrophobicity and poor water solubility pose a problem for its use as a drug.

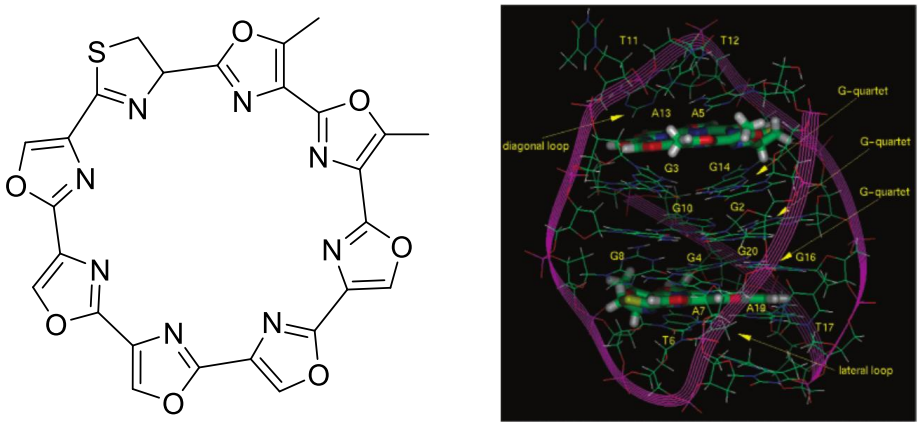


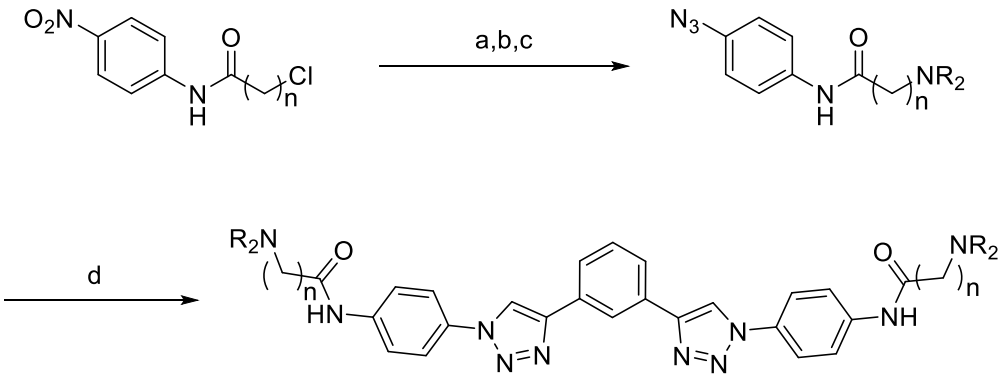
Figure 1.18. Chemical structure of telomestatin (left). Interaction of telomestatin (2:1 stoichiometry) with $d(T_2AG_3)_4$ G4 arrangement (right) by stacking on the external G-tetrads proposed by Kim *et al.* by a simulated annealing docking study.[248]

1.5.6 Development of other relevant G4 ligands through the last 15 years

In 1999, Perry and co-workers reported the development of another series of G4 ligands based on the fluorenone scaffold which showed a range of telomerase inhibition activity towards the 21mer telomeric G4 sequence having a fluorescein/rhodamine fluorophores couple (F21T) with IC_{50} values ranging from 8 to 50 µM measured by the TRAP assay.[249] In the same year, a series of 3,6-disubtituted acridines were designed and synthesized by Harrison and co-workers, in particular their *in vitro* cytotoxicity was evaluated towards human ovarian cell lines (A2780, CH1 and SKOV-3) reporting IC_{50} values between 1.3 and 8.0 µM. Moreover, the inhibitory ability of these ligands was also evaluated in Taq-polymerase and TRAP assays, the 4-methylpiperidine derivative was the most potent with an IC_{50} value of 1.35 µM. Interestingly, none of these acridines showed inhibitory activity toward Taq-polymerase, indicating a marked selectivity for G4 versus duplex DNA.[250]

In 2002, Riou and co-workers reported a series of triazine derivatives with high telomerase inhibition activity in the nM range. The best compounds of the series showed in a FRET melting assay a melting temperature variation (ΔT_m) for the telomeric G4 arrangement of 8 and 20 °C respectively. Using a TRAP assay, both compounds demonstrated potent telomerase inhibition activity in human lung carcinoma A549 cells according with IC_{50} values of 0.041 and 0.13 µM, respectively.[251] Nakatani and co-workers in 2003, reported a series of naphthyridine dimers (the dimeric forms of a 2-amino-1,8-naphthyridine derivative) able to bind to the hTelo sequence. The specific binding affinities were determined through UV thermal denaturation, circular dichroism (CD), and Cold-spray Ionization (CSI) mass spectrometry (MS) measurements. In a thermal denaturation assay the best compounds demonstrated a ΔT_m of 45.6 and 63.6 °C for interaction with the secondary structures of the 9-mer d(TTAGGGTTA) and 15-mer d(TTAGGGTTAGGGTTA), respectively.[252] In 2006, Moorhouse *et al.* developed, through a click-chemistry synthetic approach (CuAAC reaction), a series of bis-triazole

analogues which were tested in a FRET melting experiment to determine their ability to stabilize the G4 telomeric sequence F21T and the dsDNA F10T. For F21T, the best compounds showed, at 0.5 μM ligand concentration, a ΔT_m of 3.9, 8.3 and 12.5 °C, respectively; however they did not effectively stabilize the F10T sequence. The compounds showed interesting telomerase inhibition activity, in particular using a TRAP assay the relative IC_{50}s were determined with values ranging from 13.3 to 23.5 μM.[253] In the scheme below (Scheme 1.6) is represented the synthetic scheme adopted by the authors for the ligands preparation, comprising a first amination reaction followed by nitro group reduction and diazotation of the resulting aryl amine, which upon treatment with NaN_3 formed the aryl azide intermediates. These were used in a final CuAAC reaction with 1,3-diethynyl-benzene yielding the final compounds.

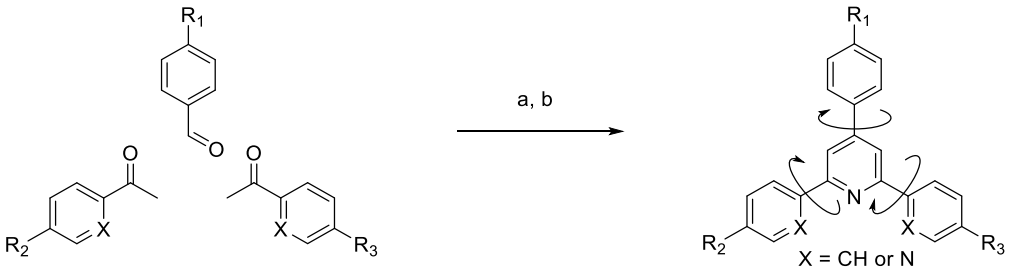


Scheme 1.6. Synthetic scheme for the preparation of the bis-triazole ligands based on the CuAAC reaction. a: HNR_2, MeOH; b: NH_4CO_2, Pd/C; c: tBuONO, HCl, NaN_3; d: Cu(I), 1,3-diethynyl-benzene.

In 2007 Le San *et al.* reported the preparation of a set of novel G4 ligands characterized by the scaffold of the quindolines (2,7-disubstituted 10H-indolo[3,2-b]quinoline derivatives). Two of the prepared molecules were synthesized through a concise and effective synthesis, in particular, the side chains were prepared from acryloyl chloride and N-methylpiperazine, then these were attached to the quindoline core, through the Heck reaction, and a final hydrogenation reaction provided the final compounds. Other compounds of the set were then prepared leveraging the Suzuki coupling to attach aromatic spacers. The best compounds of the series, demonstrated to be active as G4 stabilisers, but still markedly less active than BRACO-19. Two molecules however had a greater selectivity for quadruplex over duplex DNA, representing a real chance to obtain interesting G4 ligands through further optimization, to enhance telomerase and quadruplex potency.[254] Martins and co-workers prepared a series of benzylamino-substituted acridines, which were designed by a molecular modelling study on BRACO-19 ligand. In particular, starting from a molecular modeling study it was found a plausible low-energy binding geometry for the molecule while involved in a stacking interaction with a parallel telomeric quadruplex corresponding to the reported crystal

structure of an intramolecular 22-mer human quadruplex. Envisaging that adding an additional methylene linker at the 9-position would enable the benzyl ring to be more coplanar with the G-platform, thereby improving its affinity to human telomeric quadruplex DNA, a series of compounds with differently substituted benzylamines was prepared. The set of molecules demonstrated to be effective G4 stabilizers with ΔT_m values in the range of 21.0 – 32.5 °C at a ligand concentration of 1 μM. Moreover, the majority of the compounds demonstrated, in a TRAP assay, also telomerase inhibition values lower than 1 μM.[255] Waller and colleagues created a set of triarylpyridine derivatives showing marked G4 binding properties compared to dsDNA, in a FRET melting assay the compounds demonstrated a good ability to stabilize the G4 motif with ΔT_m values up to 21.1 °C at a 1 μM ligand concentration.

Waller and colleagues of the Balasubramanian research group, created a set of triarylpyridine derivatives as new G4 ligands possessing adaptive structural features arising from three rotatable bonds connecting the three aromatic rings to the central one. The set of compounds was prepared with a simple yet effective synthetic strategy, namely through a series of condensation reactions starting with an aldol condensation of an enolisable ketone and a benzaldehyde, followed by Michael addition of the enone with a second enolisable ketone. This was converted to a pyridine *via* a double condensation in the presence of ammonium acetate in acetic acid (Scheme 1.7). The ligands possessed marked G4 binding properties compared to dsDNA, in a FRET melting assay stabilized the G4 DNA with ΔT_m values up to 21.1 °C at 1 μM ligand concentration. The study determined also the dissociation constants (K_d) for the prepared compounds with an overall best result for the hTelo of 3.20 x10^{-7} M related to the best stabilizer in terms of ΔT_m.[256]



Scheme 1.7. Synthetic scheme for the preparation of the triarylpyridine ligands by multiple condensation reaction (right). a: NaOH; b: NH_4OAc, reflux, MeOH, AcOH.

In 2008 the Balasubramanian research group reported a series of bis-indole-carboxamide G4 ligands. Similarly to the molecules of the previously cited study, also part of these compounds were characterized by a conformational flexibility given by the presence of rotable bonds connecting the aryl rings composing the aromatic part of the ligands (Figure 1.19).

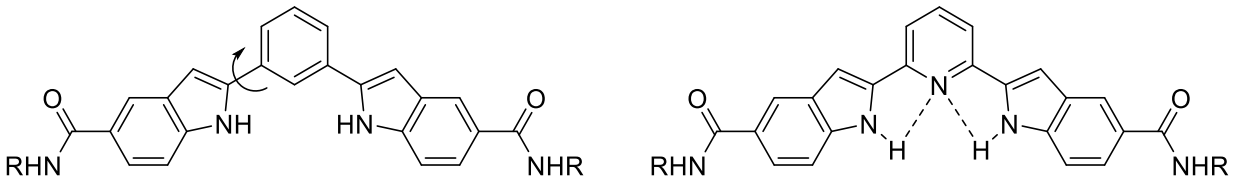
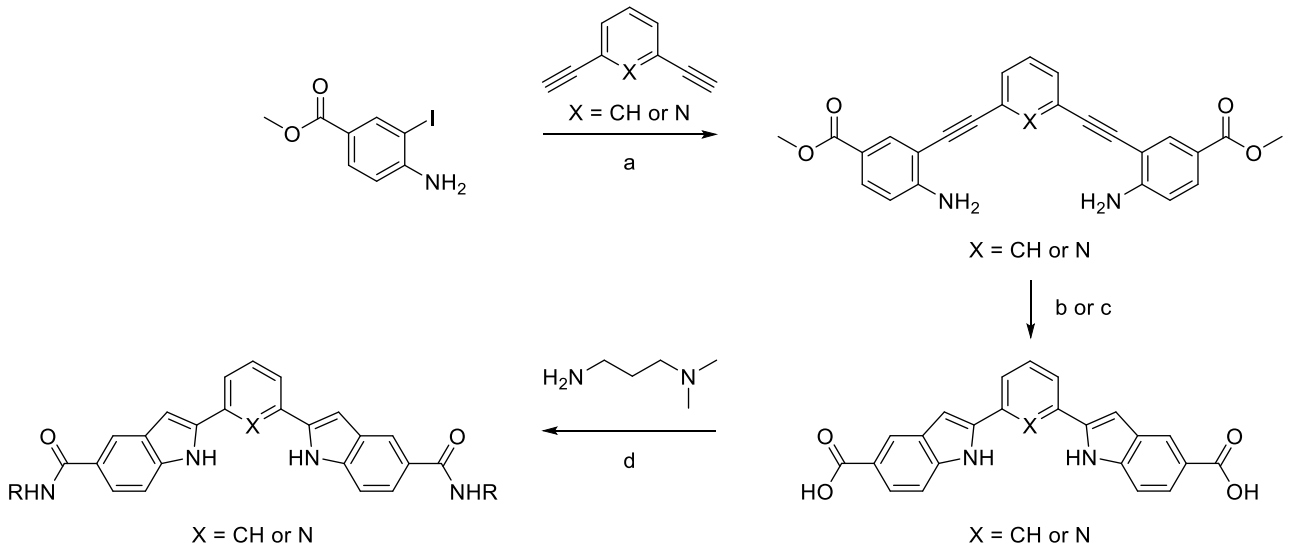


Figure 1.19. Chemical structure of the bis-indole-carboxamide G4 ligands. Free-rotation of the bond connecting the benzene to the indole rings (left) and limited rotation due to H-bonds formation for the pyridine derivatives (right) are also showed.

The synthetic strategy provided for a Sonogashira coupling to connect the three aromatic rings followed by a 5-endo-dig cyclization forming the indoles. The preparation was completed by the connection of the side chains through an amidation reaction (Scheme 1.8) The two final compounds were tested with a FRET melting assay for their ability to stabilize the telomeric G4 motif and one of them showed a ΔT_m value of 21.5 °C. Confirming the importance of the rotables, the lower activity of the second compound can be ascribed to its higher rigidity due to the formation of hydrogen bonds between the NH of the indole ring and the pyridine nitrogen lone pair.[257]

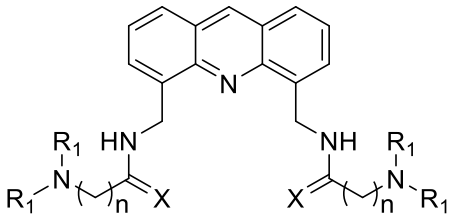


Scheme 1.8. Synthetic scheme for the preparation of the bis-indole-carboxamide derivatives through Sonogashira coupling and amide coupling. a: $PdCl_2(PPh_3)_2$, CuI, Et_3N, DMF; b: $NaAuCl_4 \cdot 2H_2O$; c: CuI, NMP; d: EDC, HOBt, NMM, DMF.

Lu and co-workers developed a series of 5-N-methyl quindoline (cryptolepine) derivatives as telomeric G4 ligands. The compounds possessed a positive charge at the 5-N position on the aromatic quindoline scaffold due to the nitrogen methylation; this modification improved significantly the binding ability and increased the selectivity towards the antiparallel topology whereas the non-

methylated derivative demonstrated a preference for the hybrid G4. All the compounds showed a good ability to stabilize the telomeric G4 with ΔT_m values ranging from 4.0 to 19.0 °C at 1 μM ligand concentration, moreover one compound gave a IC_{50} values of 1.21 and 1.36 μM in a cell viability assay on the HL60 (leukemia) and CA46 (lymphoma) cell lines. A molecular docking study indicated that the ligands stack on the external guanine tetrad of a parallel G4 structure (PDB ID: 1KF1), considered by the authors as the more biologically relevant form. On the same model, a following MD simulation showed that the methylated nitrogen positive charge could contribute to the stability of the complex interacting by an electrostatic interaction with the negatively charged carbonyl channel of the G4, the terminal protonated amino groups tended to interact with the phosphodiester backbone through electrostatic and hydrogen bond interactions. From the MD simulations the MM/GBSA free energy of binding were also determined. Interestingly a high correlation value of 0.75 was found between the MM/GBSA and the ΔT_m values.[258]

In 2009 Laronze-Cochard *et al.* reported the preparation of a series of 4,5-bis(dialkylaminoalkyl)-substituted acridines (Figure 1.20). The compounds, characterized by two side chains, demonstrated a lower ability to stabilize the G4 than BRACO-19 which was used as a reference, the ΔT_m values obtained were in the range of 2.4 to 16.7 μM at 20 μM ligand concentration.[259]

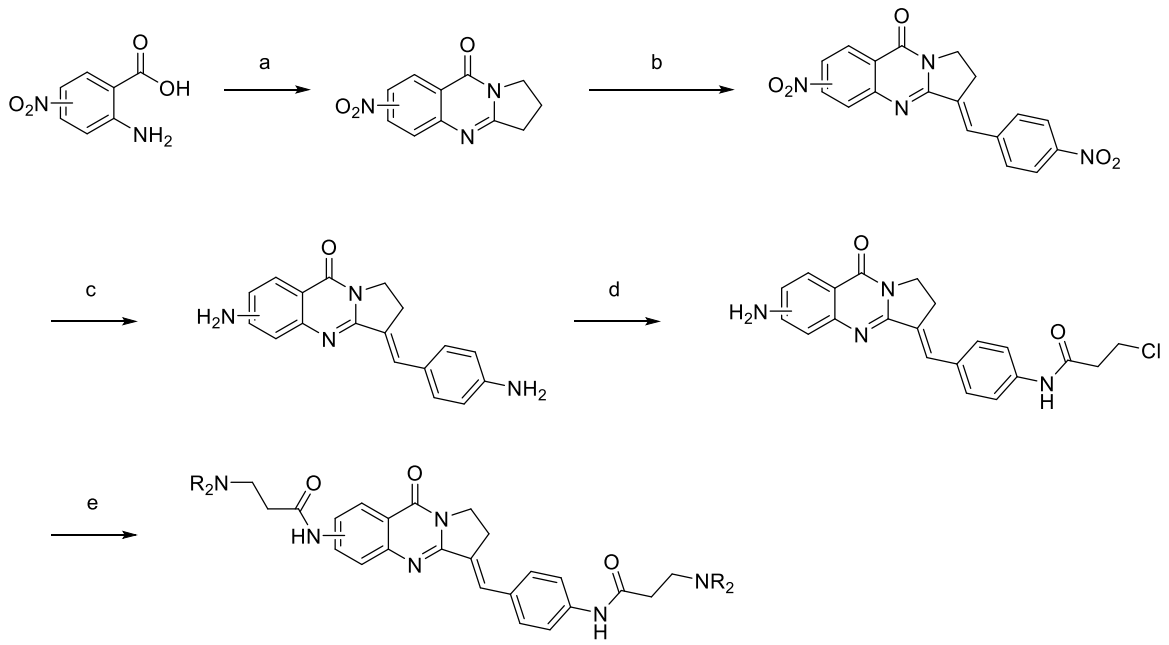


$X = H_2$; $n = 3$; $R_1 = CH_3$

Figure 1.20. Chemical structure of the G4 ligands 4,5-bis(dialkylaminoalkyl)-substituted acridines.

Tan and co-workers reported a set of four isaindigotone derivatives designed as telomeric G4 ligands.[260] Following the flexibility paradigm already pursued by other research groups, the isaindigotone scaffold was selected being neither polycyclic nor macrocyclic and having free rotation around the double bond which enables the adoption of possible twisted and coplanar conformations of the aryl groups. Hence, this adaptive structural feature was thought as ideal to allow the unfused aromatic scaffold not only to occupy the G-quartet but also bind to the G4 grooves. A facile synthetic pathway was adopted by the authors that comprehended the condensation of 2-amino-4-nitrobenzoic acid and pyrrolidin-2-one followed by aldol-like condensation of this intermediate with 4-nitrobenzaldehyde, presumably through an enamine nucleophile. The reduction of the nitro groups followed by their acylation with chloropropionyl chloride and amination reactions create the side

chains and yield the final compounds (Scheme 1.9). In a FRET assay the ligands, at 1 μM concentration produced ΔT_m values in the range of 17.6 to 22.0 °C, moreover showing in a fluorescence assay binding constants up to $10.1x10^6$ M^{-1}. A computational study was also performed to reveal the most plausible binding modes of the ligands towards two G4 structures/topologies, parallel (PDB ID: 1KF1) and hybrid-type (PDB ID: 2HY9) considered by the authors the more biologically relevant forms in the presence of K^+ based on the literature.[261] The combined docking study, MD simulations, and MM/PBSA calculations, indicated that besides the indeed present stacking interactions, the ligands could be readily accommodated in the groove regions of both propeller-type and hybrid-type G4, moreover both of the ligands possessed much more favourable groove binding interactions (lower binding free energy) with hybrid-type G4 than propeller-type G4, and whatever binding of the ligand to the hybrid-type G4 showed much more comparable binding free energy. In conclusion it is noteworthy to mention that for all the ligands the CD binding studies revealed high stoichiometry ratio of up to 4:1 ligand to G4.[260]



Scheme 1.9. Synthetic scheme for the preparation of the isaindigotone derivatives based on the aldol like condensation reaction. a: pyrrolidin-2-one, $POCl_3$; b: 4-nitrobenzaldehyde, Ac_2O; c: $Na_2S{\cdot}9H_2O$, NaOH, EtOH; d: $Cl(CH_2)_2COCl$; e: R_2NH, KI, EtOH.

The derivatization of berberine to obtain G4 ligands was already explored by the Neidle's research group, which prepared a single piperidino-berberine compound which was tested alongside berberine itself and coralyne (Figure 1.21), obtaining good ΔT_m values against the F21T 21-mer human telomeric sequence [GGG(TTAGGG)3].[262] In 2009 Ma and co-workers developed a series of

9-O-substituted berberine derivatives as novel telomeric G4 ligands (Figure 1.21, compound zh-4d). These compounds, characterized by a single side-chain, were obtained by alkylation of the oxygen in position 9 on the berberine scaffold. The CD study indicated the ability of the berberine derivatives to induce the formation of an antiparallel topology on a HTG21 sequence, moreover a FRET melting assay on the human telomeric G4 F21T and F10T sequences determined the ability to stabilize the G4 motif with ΔT_m values up to 25 °C at 2 μM ligand concentration, furthermore negligible stabilization effect was recorded with duplex DNA. Noteworthy also the binding constants obtained by a fluorescence assay showed a good ability for the ligands to interact with the G4 arrangement with K_a values in the 7.3-12.0x10^6 M^{-1} range.[263]

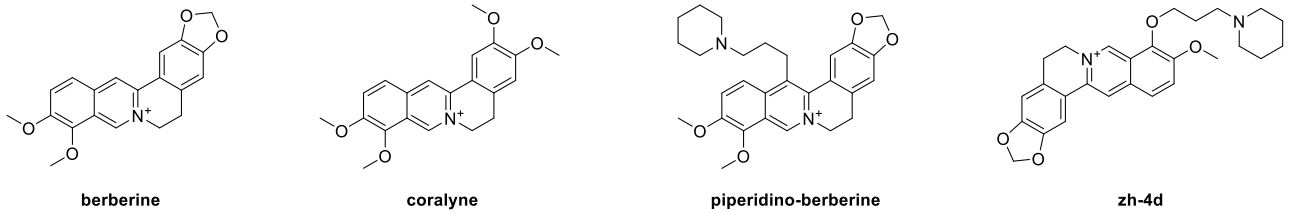


Figure 1.21. Chemical structures of the natural compounds berberine and coralyne and also of the synthetic piperidino-berberine and berberine derivatives.

In 2012 Paul *et al.* reported the preparation of a series of bis-benzimidazole derivatives where the two benzimidazole units were bridged by a polyethylene glycol spacer, the ligands demonstrated to be able to recognize both monomeric $d(TTAGGG)_4$ and dimeric $d(TTAGGG)_8$ G4 arrangements.[264] Some of the compounds demonstrated selective cytotoxicity towards the tumor cell lines HEK293T, HeLa, A549 and NIH3T3, in particular one of them showed an IC_{50} value of 15 μM against HEK293T cells. Furthermore, a TRAP assay revealed IC_{50} values up to 55.4 μM with ligand concentration varying from 2.5 to 60 μM. Larsen and colleagues reported a set of novel G4 ligands having a 4,7-diamino-substituted 1,10-phenanthroline-2,9-dicarboxamide scaffold. In a FRET melting assay using the F21T sequence ΔT_m values up to 27.6 °C at 1 μM concentration of the ligand. Still in 2012 Rahman and co-workers, through a FRET melting assay screening of a library of 2307 molecules, identified 13 new G4 ligands. The sequence used was the human telomeric F21T, for which the compounds showed a high selectivity compared to dsDNA, in particular ΔT_m values ranging between 7.3 and 24.2 °C were found at 1 μM ligand concentration. The same authors discovered in the same year that the natural product Cristatin A is able to stabilize, with high selectivity, the G4 DNA with a ΔTm value of 8.9 °C, moreover a short-term MTS cell viability assay revealed a marked cytotoxic effect against four different tumor cell lines (MDA MB 231, A431, MIAPaCa-2 and HeLa) with IC_{50} values ranging between 2.5-10.4 μM.

In 2014 Maji *et al.* reported the design and the synthesis of a novel series of 4 G4 ligands characterized by a benzimidazole–carbazole conjugate scaffold (Figure 1.22).[265] More in the detail, the compounds were designed to have a central aromatic part for the pi-stacking interaction with the G-quartet, formed by the carbazole and benzimidazole units and two or three side chains having a piperazine/morpholine moieties to induce hydrogen bonding interactions and to get protonated at physiological pH to facilitate the ligand's interaction with the phosphate groups of the G4 DNA.

The molecules demonstrated a high binding affinity for G4 DNA and two of the ligands showed IC_{50} values in the sub-micromolar range in a TRAP assay.

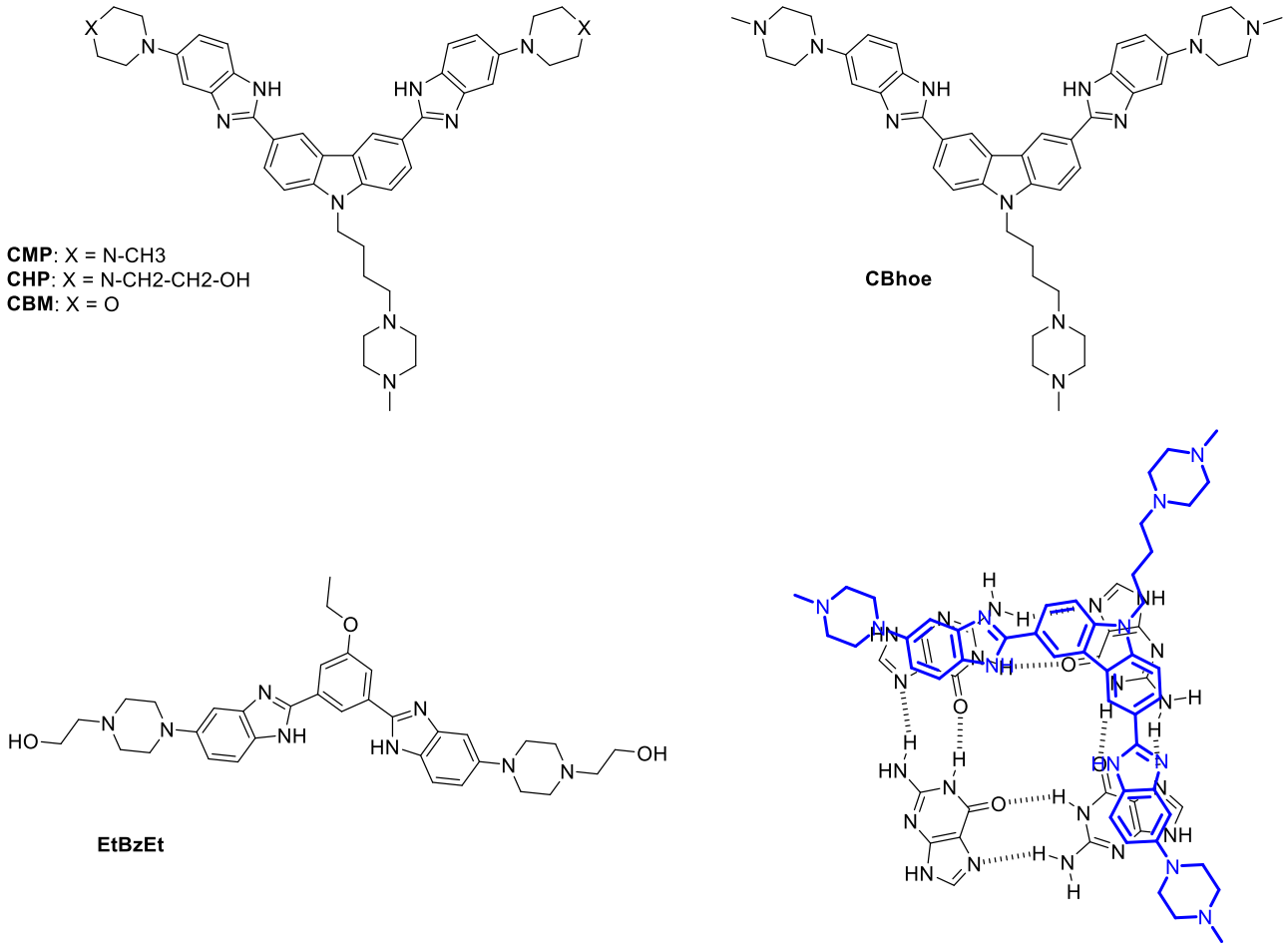
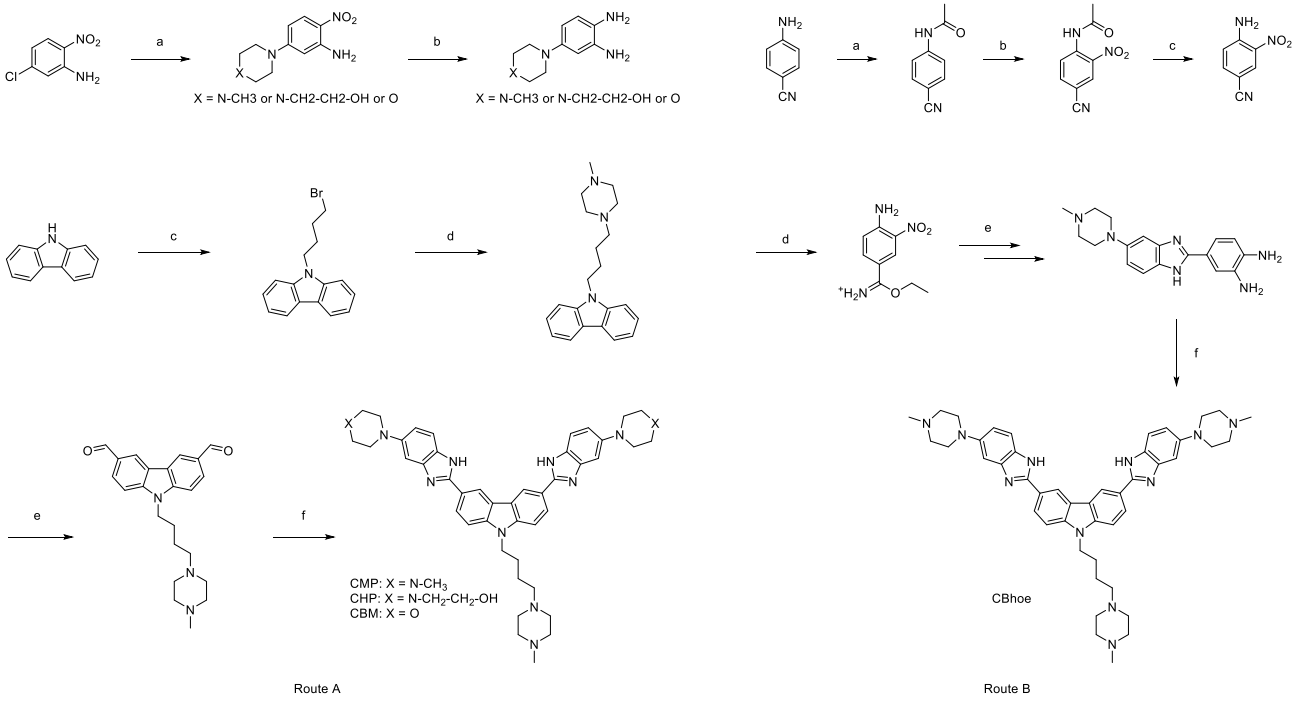


Figure 1.22. Chemical structures of a benzimidazole–carbazole conjugate G4 ligands. In the lower-right part of the figure is represented the ligand CMP on top of a G-quartet.

The synthesis of three of the compounds (CMP, CHP and CBM) provided for carbazole alkylation with 1,4-dibromobutane which was first reacted with N-methyl piperazine and then formylated. The so obtained intermediate was then subjected to oxidative coupling with three different diamines separately prepared obtaining the three ligands. The synthesis of the fourth ligand (CBhoe) started

with 4-aminobenzonitrile acylation followed by nitration and amide hydrolysis, then the so obtained intermediate was converted to its imino-ether hydrochloride salt derivative which upon reaction with the previously cited diamine formed the benzimidazole scaffold. After reduction of the nitro group, the intermediate was used as a diamine itself for the oxidative coupling with the dialdehyde carbazole derivative yielding the final compound. The synthetic schemes are depicted in the following schemes (Scheme 1.10).



Scheme 1.10. Synthetic schemes for the preparation of the benzimidazole–carbazole conjugates. Route A. a: K_2CO_3; b: Pd/C, H_2, EtOH; c: 1,4-dibromobutane, TBAI, NaOH, benzene; d: N-methyl-piperazine, K_2CO_3, acetonitrile; e: $ZnCl_2$, $POCl_3$, DMF; f: $Na_2S_2O_5$, EtOH. Route B. a: Ac_2O, NaOAc; b: KNO_3, H_2SO_4; c: 10% H_2SO_4; d: anhyd. HCl(g), dry EtOH; e: 4-(4-methylpiperazin-1-yl)benzene-1,2-diamine, EtOH, AcOH; f: Pd/C, H_2, EtOH; g: 9-(4-(4-methylpiperazin-1-yl)butyl)-9*H*-carbazole-3,6-dicarbaldehyde, $Na_2S_2O_5$, EtOH.

The compounds synthesized were tested for their ability to interact with the G4 DNA preformed from the 21mer human sequence $G_3(T_2AG_3)_3$. A first UV-Vis assay was used to determine the binding constants of the compounds obtaining values for the G4 DNA from 0.05 to 0.38 μM in the presence of NaCl and 0.03 to 0.33 μM with KCl, for the dsDNA from 25 to 196 μM. A following CD study revealed a hybrid topology for the preformed G4 DNA in the presence of KCl, whereas with NaCl an antiparallel topology was noted. Titration of the preformed G4 DNA with the ligands produced changes in the dichroic spectra suggesting the generation of more than one form of DNA-ligand complexes possibly due to the groove binding interaction. Furthermore, the formation of parallel G4 was noted when forming the G4 structure in the presence of two of the ligands (CMP and CHP). A CD melting assay was also conducted with the G4s preformed in the presence of NaCl and of KCl,

the maximum ΔT_m value registered was of 19 °C. A TRAP assay was then performed on the ligands revealing IC_{50} in the low μM range. In conclusion a molecular modeling study (molecular docking and MD) was conducted using a parallel human telomeric G4 structure (PDB ID: 1KF1), the results suggested that the molecule CMP preferred a stacking binding mode on the 3'-end G-quartet using the three piperazine arms to approach the grooves interacting with the anionic phosphodiester backbone.

As a continuation of the above mentioned work, the same author reported in 2015 a new series of dimeric carbazole–benzimidazole compounds as G4 ligands, which were characterized by the presence of spacers of different polarity linking the two carbazole units.[266] The ligands' monomers were prepared as reported in the previous work of the authors, whereas the ones comprehending the spacers were obtained first by conjugating, by alkylation, two carbazole monomers through either a neutral hexaethylene glycol based unit or one having both oxyethylene and protonable piperazine moieties. The so obtained dimers were then subjected to the same synthetic scheme as the one used for the monomeric ligands affording the final dimeric compounds.

Four of the synthesized compounds demonstrated a good telomerase inhibition activity with IC_{50} values in the low μM region and for one compound up to 0.25 μM. Furthermore, a TRAP assay demonstrated that the best ligand in terms of telomerase inhibition was also fourfold more potent than the corresponding monomer. A computational study was also performed first by docking, with AutoDock 4.0, the compounds on a parallel G4 structure (PDB ID: 1KF1), then the so-obtained DNA-ligand complexes were subjected to a MD simulation which showed a general better efficiency of the dimeric ligand than the monomeric one. More in the detail the combined docking-MD computational approach showed that the monomeric ligand efficiently stacks, through its planar core, onto three guanines of the G-tetrad, furthermore the side chain piperazine approaches the groove interacting with the phosphate backbone. On the other hand, the dimer interacts through one arm with the G-tetrad, whereas the other arm interacts with the groove formed by the T17-T18-T19 nucleotides (Figure 1.23).

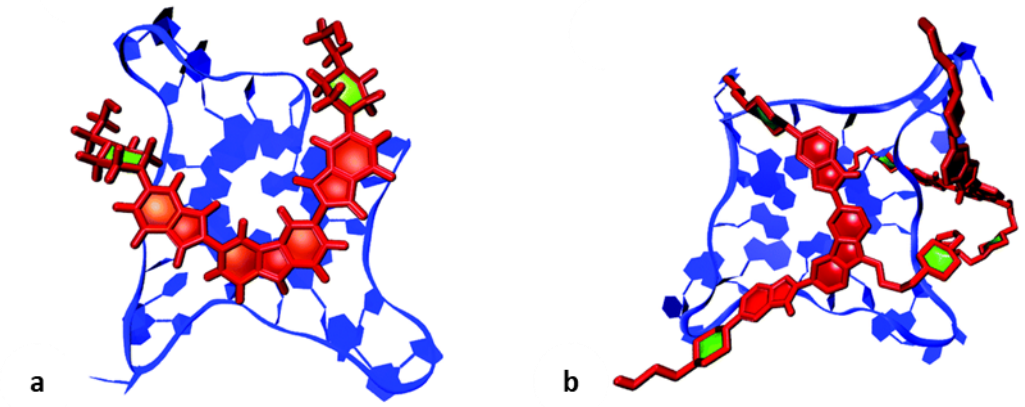
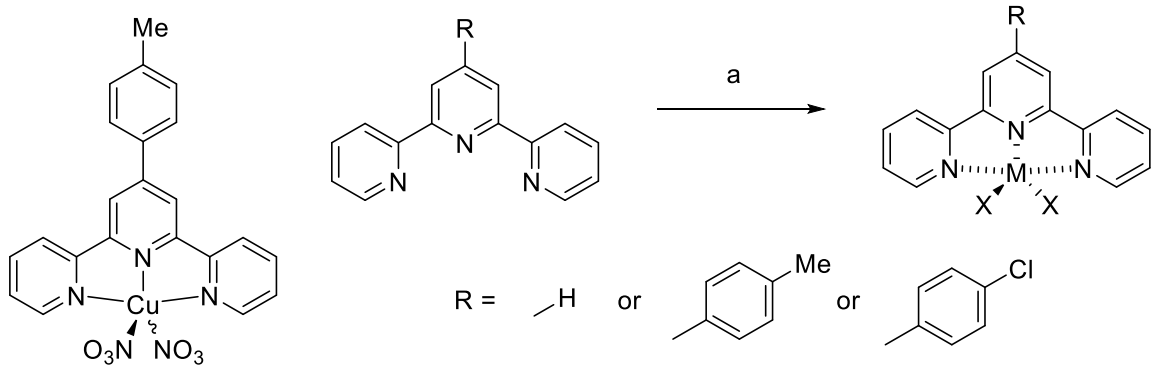


Figure 1.23. Complex of the carbazole–benzimidazole monomeric (a) and dimeric (b) ligands with telomeric parallel G4 arrangements obtained by docking.

In 2016 Marchand and colleagues reported the development of new Cu(II)-terpyridine complexes as selective ligands for human telomeric G4.[267] Metal complexes of planar geometry are known to be able to efficiently target the G4 motif, interacting with the positively charged core with the external G-quartets, which as already discussed exerts a negative charge density on the O6 atoms of the guanine of the tetrads.[268] Therefore, the electropositive metal can in principle occupy the position at the center of the G-quartet, increasing the electrostatic stabilization mimicking the cationic charge of the potassium or sodium that would normally occupy this site. Another advantage of these complex structures is the electron-withdrawing properties of the metal, which reduce the electron density on the coordinated ligand, yielding a system that can display stronger pi-pi interactions.[269–273]

The synthesis of the metal-complex ligands was based on a previously reported procedure.[272,274] In particular a one-step metal chelation (Cu, Zn, Pt and Pd) reaction was performed using as chelators one of the three different-substituted selected terpyridine and either Cl^- o NO_3^- anions (Scheme 1.11).



Scheme 1.11. Synthesis of the G4 ligands Cu(II)-terpyridine complexes and their Zn/Pt and Pd analogues. a: CuX_2 or ZnX_2 or PtX_2 or PdX_2, X = Cl- or NO_3.

The most extensively studied ligand CuII-4′-(p-tolyl)-2,2′:6′,2″-terpyridine (Cu-ttpy) was tested for its ability to bind to different G4 forming DNA sequences through an ESI-MS titration assays. The investigation included: the human telomeric 22mer $G_3(T_2AG3)_3T$, the parallel G4 from the c-myc promoter Pu24, the bimolecular antiparallel G4 $(G_4T_4G_4)_2$, and other telomeric and promotorial G4 forming sequences. Cu-ttpy demonstrated a high specificity for the telomeric sequences tested for which a 2:1 (ligand:DNA) ratio was almost always observed indicating a high cooperativity of the two consecutive binding events. In order to test the effect of the metal center, the ligands with the central ion Zn, Pd, or Pt were then tested. In particular Zn-ttpy demonstrated to be a poor ligand probably due to the trigonal bipyramidal geometry adopted by ZnII, which leads to steric hindrance. Pt and Pd complexes were able to bind to the G4, but without any cooperativity forming only a 1:1 complex. Moreover, in the case of the Pt derivatives, the ligand first bound non-covalently, then after loosing chlorine, bound covalently to the DNA. The topology of the ligand-G4 structures were studied by CD spectroscopy indicating for the telomeric hybrid structures a general tendency to change their folding to antiparallel which is a behaviour already observed for other G4 ligands as telomestatin.[275] To have a deeper insight into the binding mode of the ligand to the G4, compatible with the 2:1 stoichiometry recorded, an ESI-MS assay was performed using as the target sequence the intermolecular-bimolecular G4 forming sequences $(G_4T_4G_4)_2$ and $(G_4T_3G_4)_2$. More in the detail, these arrangements present identical environments on each external G-quartet, which is also a feature not promoting a cooperativity in the two binding events forming the 2:1 Cu-ttpy complexes. Upon collisional activation, the dimers broke into two strands, each containing exclusively either one ligand or one Cu (neutral ttpy ejected from the complex), therefore strongly suggesting that the two ligands are bound at distal positions, which appears in accordance with their position on top of the two external G-quartets.

Gueddouda *et al.* reported in 2016 a novel series of 2,9-bis[(substituted-aminomethyl)phenyl]-1,10-phenanthroline derivatives as G4 Ligands.[27] The 1,10-phenanthroline derivatives, such as PhenDC3 and its structural analogues I, tetraaminophenanthroline II, K35, and phenanthroline-2,9-bistriazole III are known to be good ligands for the G4 DNA thanks to their extended tricyclic phenanthroline aromatic core and to their charged side-chains (Figure 1.24).[276–286] The new phenanthroline derivatives described by the authors were characterized by the substitution with diaminobenzyl moieties at position 2 and at position 9, which were added in order to increase the overall aromatic surface and to enhance the aqueous solubility of the molecules.

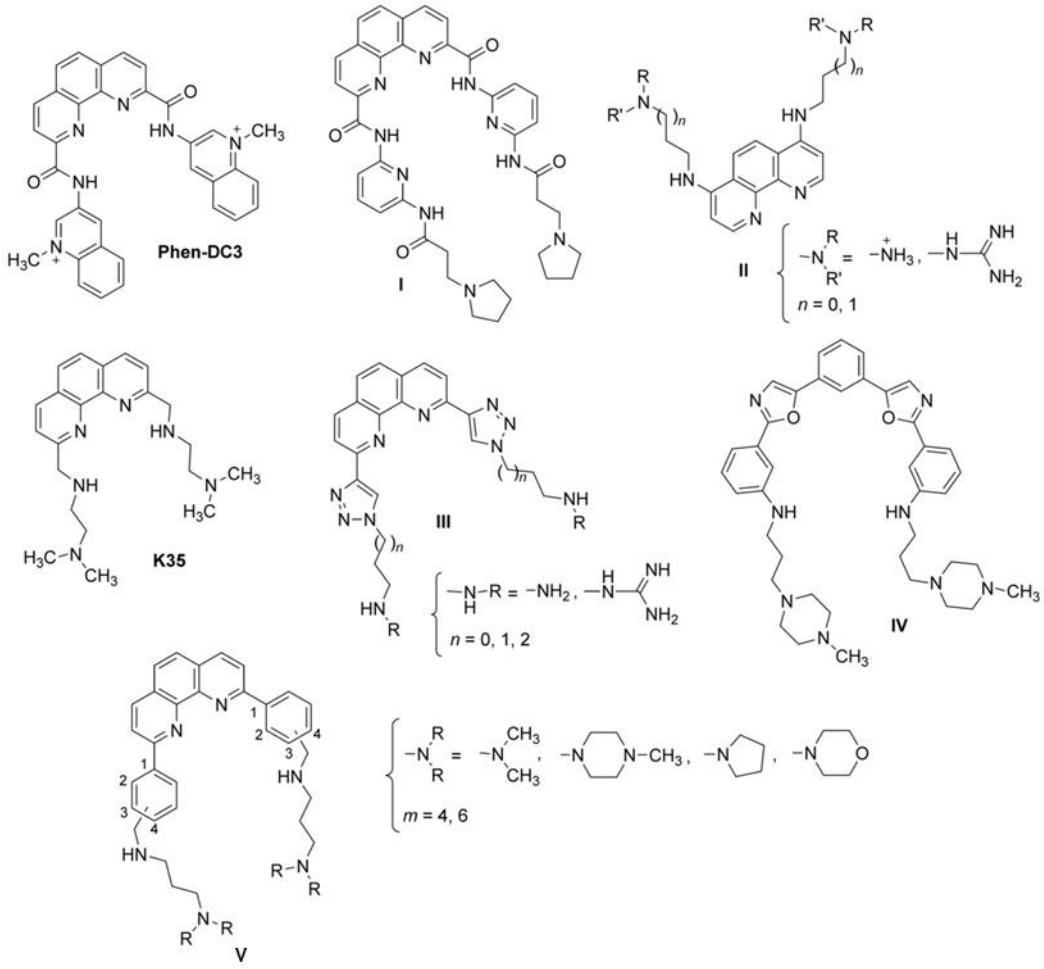
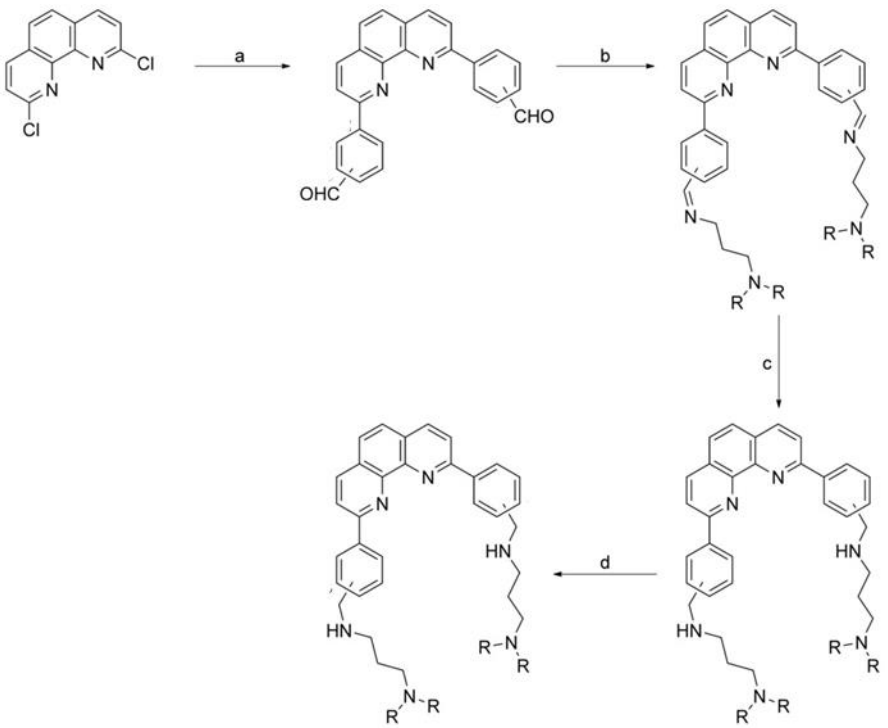
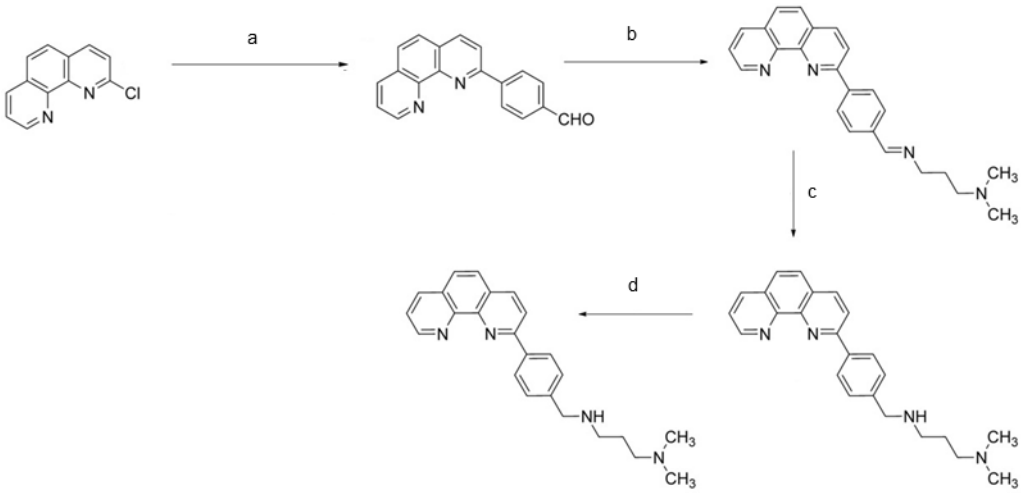


Figure 1.24. Phenanthroline derivatives used as G4 ligands including PhenDC3 and its structural analogues I, tetraaminophenanthroline II, K35, phenanthroline-2,9-bistriazole III and the new ligands developed by the authors IV and V.

The synthesis of the compounds started from 2,9-dichloro-1,10-phenanthroline which was subjected to a direct double *Suzuki–Miyaura* cross-coupling reaction with the preformed 2-, 3- or 4-formylphenylboronic acids. These intermediates were then reacted with primary amines to give the respective di-imines, which were then reduced to the corresponding secondary amines. Furthermore, also a single mono-substituted analogous was prepared substituting the 1,10-phenanthroline core with an aminomethylphenyl moiety substituted only at position 2. The synthetic schemes for the preparation of the mono- and di-subtituted ligands are reported in the figure below (Scheme 1.12).



Route A



Route B

Scheme 1.12. Synthetic schemes for the preparation of the mono- (Route A) and di-substituted (Route B) phenantroline derivatives object of the cited study. Route A. a: Method A: OHC-C_6H_4-$B(OH)_2$, $Pd[P(C_6H_5)_3]_4$, K_2CO_3, toluene/EtOH - Method B: OHC-C_6H_4-$B(OH)_2$, $Pd[P(C_6H_5)_3]_4$, Na_2CO_3/H_2O, DME; b: H_2N-$(CH_2)_3$-NR_2, EtOH; c: $NaBH_4$, MeOH; d: $(COOH)_2$, isopropanol. Route B. a: OHC-C_6H_4-$B(OH)_2$, $Pd[P(C_6H_5)_3]_4$, K_2CO_3, toluene/EtOH; b: 4-formylphenylboronic acid MIDA ester, $Pd(OAc)_2$, S-Phos, K_3PO_4/H_2O, THF; c: OHC-C_6H_4-$B[OC(CH_3)_2]$, $Pd[P(C_6H_5)_3]_4$, Cs_2CO_3, DME; d: H_2N-$(CH_2)_3$-NR_2, EtOH; e: $NaBH_4$, MeOH; f: $(COOH)_2$, isopropanol.

The so-obtained ten final compounds were tested with a FRET melting assay for their ability to interact with the telomeric G4 DNA deriving from the folding of the F21T sequence (Fam-$G_3T_2AG_3T_2AG_3T_2AG_3$ –Tamra), which in the presence of K^+ ions, is known to adopt a mixture of G4 conformations containing an antiparallel and a hybrid type topology, whereas in Na^+ conditions, forms an antiparallel G4 arrangement.

Each of the synthetized compounds stabilized the F21T sequence in both K^+ and Na^+ conditions, and the most stabilizing ligand increased the G4 DNA melting temperature up to 14.9 °C at 1 μM ligand concentration. Moreover, a competition experiment, using 15 or 50 equivalents of dsDNA, demonstrated a general small decrease in the ΔT_m values.

In order to investigate how the best phenanthroline derivative can stabilize other G4-forming sequences, the authors performed a FRET-melting assay with other G-rich, fluorophore tagged, DNA oligonucleotides that are known to adopt various G4 topologies. The study included the sequences F25cebT and FcmycT forming parallel topology G4s, the FtbaT and the Fbom17T folding into antiparallel topology, the F21RT which is a RNA oligonucleotide which folds into a parallel G4, the FctaT, a DNA oligonucleotide which is a mutant variant derived from the DNA human telomeric sequence that adopts an antiparallel G4 topology and the DNA oligonucleotide FdxT, a non-G4-forming control sequence. From the screening, ligand 1a demonstrated to be able to stabilize preferentially the RNA F21RT and the DNA F21T human telomeric sequences. To a lesser extent it stabilized FtbaT and it did not stabilize the other sequences. Furthermore, a CD spectroscopy study suggested the noteworthy ability of 1a to induce a variation of the antiparallel to parallel topology for human telomeric sequence 22CTA, an ESI-MS study revealed a 2:1 stoichiometry of interaction between 1a and the G4 arrangement, and a Pif1 helicase unwinding assay demonstrated the ability of the ligand to unwind the G4 structures. In conclusion, a docking investigation, performed with AutoDock software, was used to determine the binding mode of the ligand towards different G4 topologies including parallel (PDB ID: 1KF1), antiparallel (PDB ID: 2KM3) and hybrid (PDB ID: 2GKU) G4 structures. With the parallel topology the phenanthroline part of the 1a ligand stacked on top of the external G-tetrad pointing the charged amine side chains towards the negatively charged phosphate backbone, similar interaction modes were recorded for the antiparallel and the hybrid conformations even if the interactions observed took place mainly with the external non-guanine bases.

In 2017 Răsădean and colleagues reported the design and synthesis of water soluble, amino acid-functionalised naphthalenediimides (NDIs) as novel G4 potential ligands (Figure 1.25).[287] The symmetrical disubstituted NDI derivatives present a planar core, which has the role to develop a stacking interaction on an external G-tetrad, the amino acid side chains instead were introduced to increase the solubility of the molecules and to provide electrostatic interactions with the loops of the

G4 motif. In particular lysine (also as Boc-protected on the N_α and N_ε) and arginine (as nitro-protected) were selected as the amino acid side chains due to their high charge content at physiological pH and their optimal length to reach the phosphate backbone.

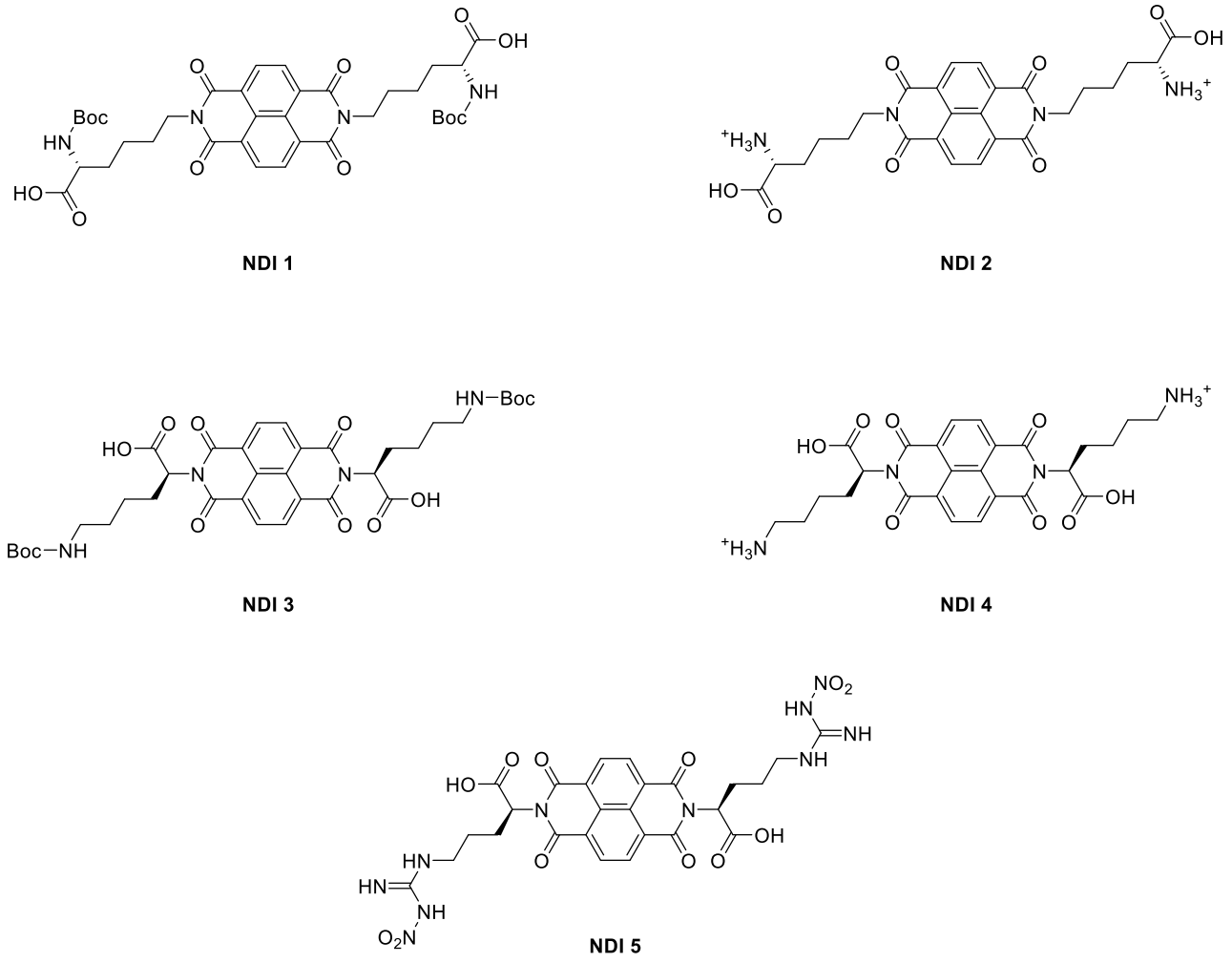


Figure 1.25. Chemical structures of the napthalenediimides (NDI 1-5) ligands object of the cited study.

The synthesis of the 6 new ligands was particularly fast and efficient as it mainly provided for the microwave (MW) thermal reaction between 1,4,5,8-Naphthalenetetracarboxylic di-anhydride and the corresponding protected amino acid, this produced molecules corresponding either to the final compounds or to protected intermediate ready for deprotection in common conditions.

The ligands were then tested for their ability to bind to the G4 DNA by CD melting assay. More in the detail, the compounds were screened against different sequences namely c-kit2, c-kit1, c-myc, k-ras, h-telo and dsDNA, all folding in the parallel topology except the telomeric one which folds into the antiparallel fold. The best results were obtained for the stabilization of c-kit2 with a ΔT_m value of 14.6 °C and a complete selectivity over the dsDNA, a good increase in the melting temperature was

also obtained for c-myc G4 (11.1 °C). Moreover, no remarkable stabilization effects were observed for k-ras G4 and for the antiparallel G4 fold of the telomeric DNA.

Tikhomirov and co-workers reported in 2018 a series of tri-armed ligands of G4 based on a heteroarene-fused anthraquinone scaffold.[288] The study, initially used a combined molecular modelling method to identify the binding mode of 4,11-bis((2-guanidinoethyl)amino)anthra[2,3-b]thiophene-5,10-dione (1a) which was previously reported as a G4 ligand. Then two series of tri-armed ligand based on furan- or thiophene-fused anthraquinone were prepared. The modeling study used as a receptor the NMR resolved structure of a hybrid-1 form G4 arrangement formed by the 24mer human telomeric sequence d($T_2G_3T_2AG_3T_2AG_3T_3AG_3A$) in the presence of K^+ ions and also in the presence of the ligand telomestatin, stacked on top of an external G-quartet (PDB ID: 2MB3). Through this investigation involving the use of docking (ICM-Pro 3.8.6) and MD (Amber16, OL15 and GAFF ffs) techniques, it was possible for the authors, to determine the best conformation of the complex between the G4 motif and 1a in solution. More in the detail, an initial docking was followed by a 6 ns MD simulation which was also followed by a last docking stage to re-validate the ligand position determined by the MD calculation. After the final docking, the location of 1a remained almost the same as the complex obtained from MD, in particular the protonated guanidine groups formed hydrogen bonds with the phosphate backbone. It was then observed that two phosphate groups of the G4 backbone were located close to the thiophene cycle of the ligand, thus suggesting that the attachment of a third side chain with the terminal cationic group, to the aromatic core of ligand 1a would represent a valuable opportunity for 1a ligand optimization. For this reason, the modification of the scaffold of compound 1a was considered and a set of new ligands were prepared. The position 2 of the tetracyclic scaffold was derivatized with a carboxylic group, which was used for the attachment, by amide bond, of a third side chain composed by propylene or ethylene spacers and terminating with amine or guanidine moieties.

The new molecules were synthesized starting from 4,11-dimethoxyanthra[2,3-b]furan- and 4,11-dimethoxyanthra[2,3-b]thiophen-3-carboxilic acids, which were coupled with mono-Boc-protected 1,2-diaminoethane or 1,3-diaminopropane followed by Boc deprotection yielding the corresponding carboxamides. These intermediates were subjected to direct substitution on the methoxy groups with diaminoethane or diaminopropane followed by guanidation of their three free primary amines yielding a first series of final compounds. The second series was obtained through a similar procedure, in particular by the initial formation of the central side chain terminating with the guanidine moiety, followed by the substitution on the methoxy groups with alkyl amines.

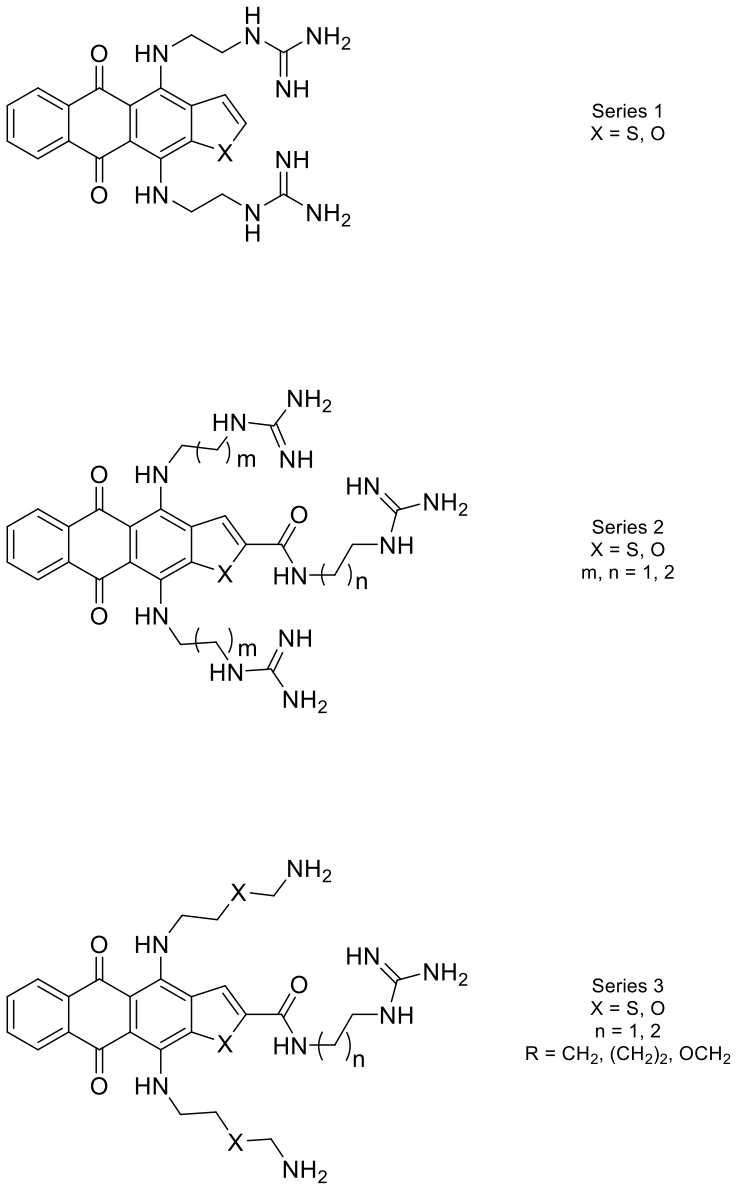


Figure 1.26. Chemical structures of the di- (Series 1) and tri-substituted (Series 2, 3) furan- or thiophene-fused anthraquinone derivatives.

The synthesized compounds were then tested for their ability to stabilize the G4 DNA with a fluorescence melting assay (FRET). In particular, the assay was conducted on the human telomeric 22-mer G-reach sequence G4 ($AG_3T_2AG_3AG_3T_2AG_3$) and on a dsDNA sequence tagged with fluorophores. All the synthesized compounds showed good ΔT_m values up to 25 °C for the G4 arrangement, moreover in a competition experiment, upon addition of 50-fold excess of untagged dsDNA, the G4 stabilization temperatures didn't drop substantially demonstrating a high selectivity for the G4 motif. The ligands were also screened for their antiproliferative activity. The screening revealed that new derivatives possess a modest antiproliferative potency with IC_{50} values up to 9 μM, 11 μM and 9 μM respectively against L1210, CEM and HeLa cell lines, which according to the authors might be related to a low intracellular penetration of the ligands.

Saha *et al.* reported in 2018 three new molecules, based on the berberine scaffold, capable of stabilizing the G4 DNA (BER, BER1 and BER2).[289] These compounds named berberrubine derived from berberine and two of its analogs bearing alkyl phenyl and biphenyl substitutions at 13-position (Figure 1.27).

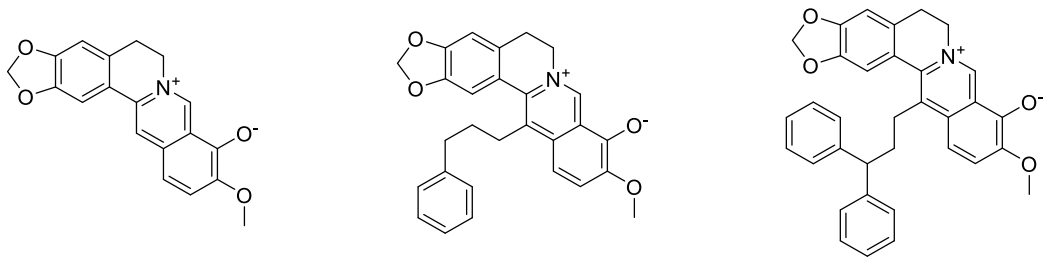
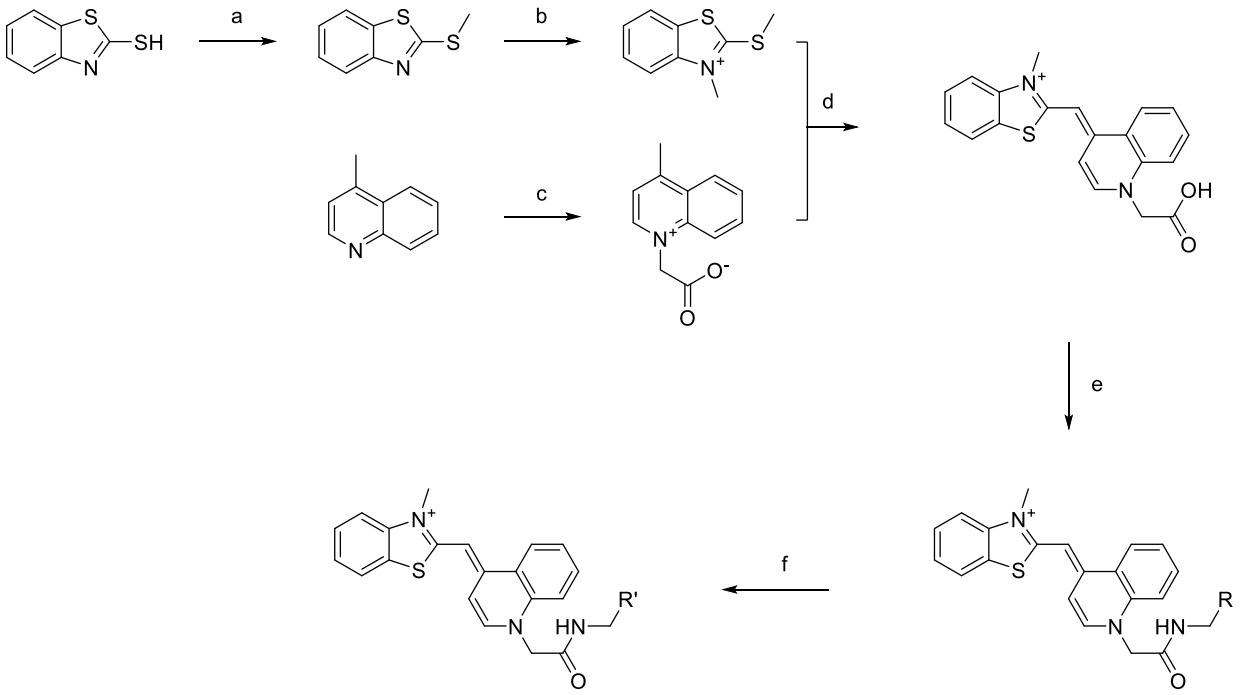


Figure 1.27. Chemical structures of berberrubine (BER) and its derivatives (BER1 and BER2).

The interaction of the series of ligands with human telomeric G4 DNA having the sequence ($G_3T_2AG_3T_2AG_3T_2AG_3$) was investigated by biophysical techniques and computational methods (molecular docking). Both the analogues exhibited higher binding affinity than berberrubine. In particular, the 13-phenylpropyl analogous (BER1) showed the best K_a with a value of 1.45×10^5 M^{-1}. Comparative fluorescence quenching studies evidenced a stronger stacking interaction of the analogs compared to berberrubine, furthermore the thiazole orange displacement assay demonstrated that the analogues were more effective in displacing the end stacked dye in comparison to berberrubine. A molecular docking investigation was also performed with the software AutoDock, using as a receptor the NMR solved structure of an intramolecular two-G-tetrad G4 formed by the human telomeric sequence ($AG_3T_2AG_3T_2AIG_2T_2AG_3T$). The results showed that each ligand binds primarily at the G rich regions of the G4 arrangement.

Wang and colleagues reported in 2019 a novel set of ligands for the interaction with telomeric G4 DNA,[290] in particular the molecules were formed by a thiazole orange-spermine conjugate. The synthesis of these compounds started from a thiazole orange carboxylic acid derivative, which was reacted, by amide coupling, with various alkyl polyamines obtaining the final compounds having the thiazole orange central scaffold and different single side-chains (Scheme 1.13).



Scheme 1.13. Synthetic scheme used for the preparation of the thiazole orange-spermine conjugates object of the cited study. a: K_2CO_3, MeI, DMF; b: MeOTs; c: bromoacetic acid, EtOAc; d: Et_3N, CH_2Cl_2; e: side chain precursor, HOBt, EDC, DMF; f: 10% HCl, EtOAc.

The thermal stabilization of telomeric G4 DNA by thiazole orange derivatives was determined through FRET melting assay. Four of the final compounds demonstrated a small increase of the ΔT_m with values up to 8 °C. A following CD melting study was also performed in order to exclude interferences on the ligand binding given by the presence of the fluorophores, through this assay it was possible to evidence better increases in the ΔT_m values, which for one of the compounds reached 18.4 °C in the presence of K^+ ions. Furthermore, the binding affinities of TO and TO derivatives to G4 DNA were evaluated using fluorometric titration, in particular for three of the synthesized ligands, using a two-site binding mode, a stronger binding site and a weaker binding site were obtained. The K_d values of the stronger binding site for the compounds were up to 0.07 μM. Moreover, pyrrolidine demonstrated to be a better side chain for G4 ligands than piperidine, also the flexible ether linkage was not favorable for the binding of ligands to G4, and the number of amino groups in the side chain, in general, had a positive correlation with the binding of the ligands to the G4 motif.

A docking study was conducted in order to exploit the plausible binding modes of the TO derivatives with telomeric G4 using Dock 6.7. The NMR of the solution structure of the complex of platinum(II) based ligand with human telomeric G4 (PDB ID: 5Z80a) was used as the initial point, in particular, the ligand was removed and the resulting hybrid type G4 arrangement was used as a receptor in a docking experiment. The three ligands demonstrated to be able to bind in the site left free by the

platinum complex, moreover end-stacking and groove interactions were observed in the resulting docked poses. In conclusion, the majority of the TO derivatives which bind to G4 DNA also showed the ability to inhibit the human telomerase activity at low concentrations, moreover, the TO-spermine conjugate exhibited a remarkable IC_{50} of 0.156 μM based on a TRAP assay which is comparable with for example BRACO-19.

Gao and co-workers reported in 2020 a novel small molecule called BMPQ-1 which exhibited a marked ability to selective recognize the hybrid telomeric multimeric G4. BMPQ-1 is characterized by a (5H-pyrazolo [4,3-c] quinoline) scaffold which presents an extended planar system without the classical charged side-chains.[291] The molecule showed ΔT_m values up to 14.6 °C for the stabilization of the dimeric telomeric G4 DNA arrangement formed by the sequence $(G_3T_2A)_7G_3$ (TTA45) at 20 μM ligand concentration. BMPQ-1 was titrated into the TTA45 sequence solution and the conformational changes were registered with CD, as a general result the ligand demonstrated a particular behaviour, namely inducing the transition of basket type motifs present on multimeric G4 to the hybrid topology. Furthermore, the BMPQ-1 molecule showed stronger inhibition against tumor cells than classical G4 ligands like BRACO-19 and RHPS4.

Sasaki and co-workers reported in 2020 a series of macrocyclic hexaoxazole compounds (6OTDs) as G4 ligands having four aminoalkyl, acetamidealkyl, or guanidinylalkyl side chains.[292] In particular, the molecules prepared were characterized by a peculiar scaffold similar in size to the planar structure of the G-quartet, resembling in a certain way the famous G4 ligand telomestatin (Figure 1.28).

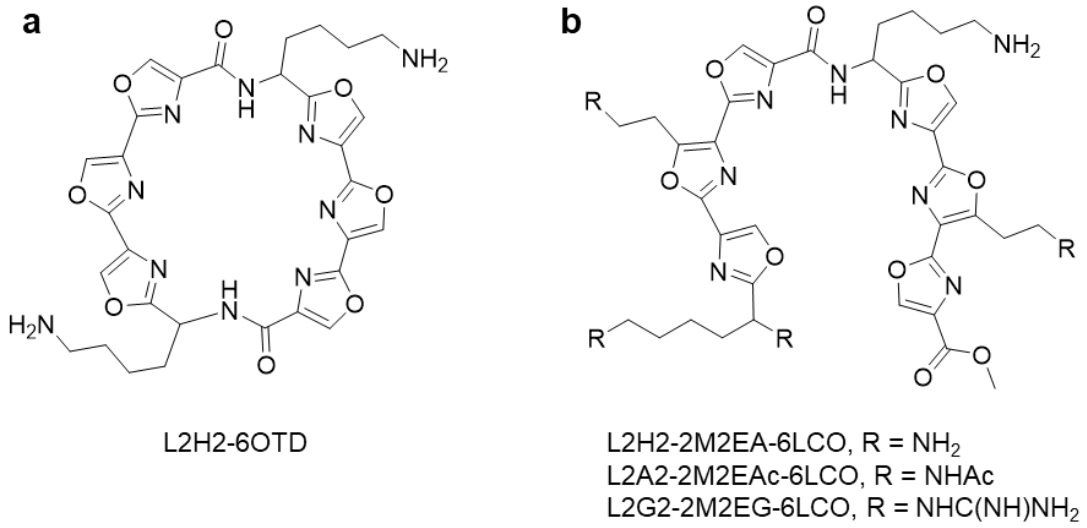
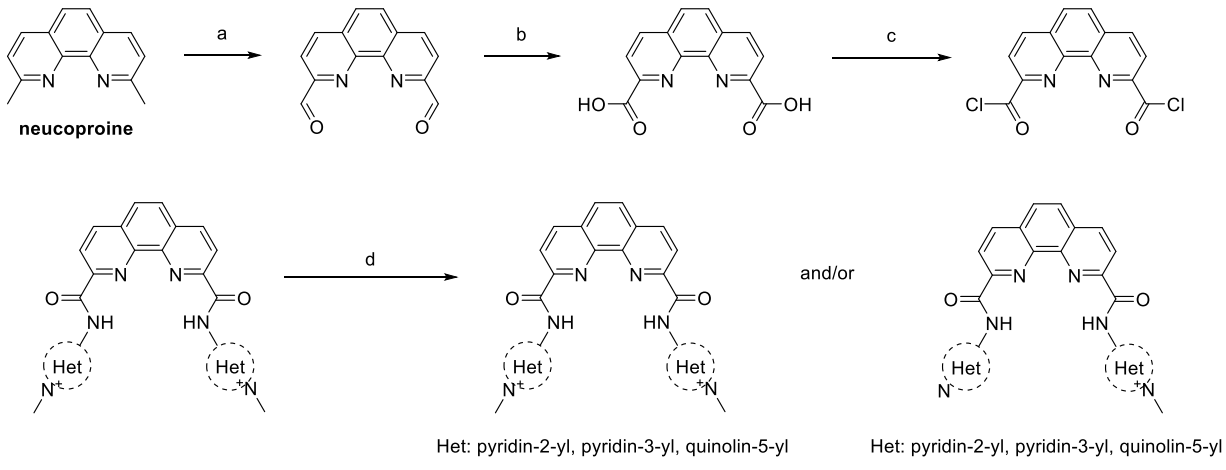


Figure 1.28. Chemical structures of the macrocyclic hexaoxazole compounds studied as G4 ligands.

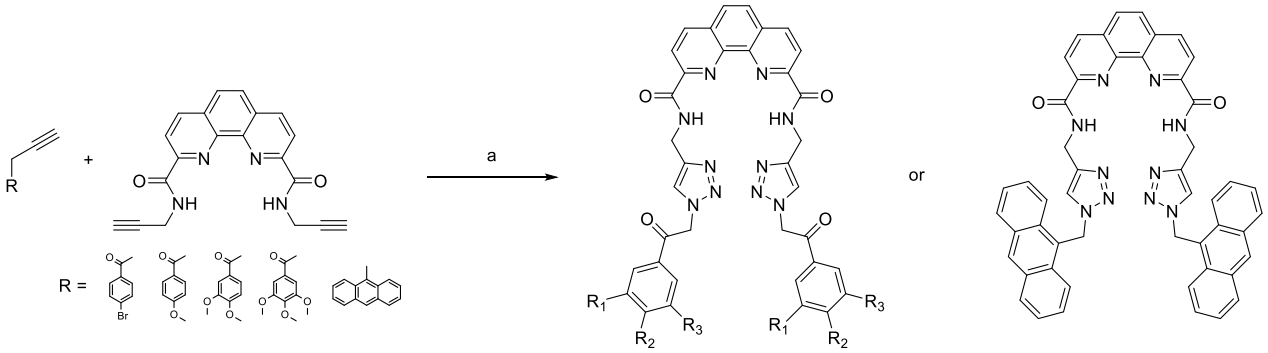
The ability of the ligands to interact and to stabilize the telomeric G4 was assessed by a FRET melting experiment and among the three compounds, only one demonstrated to stabilize the G4 motif with high selectivity over the dsDNA showing, in the presence of K^+ ions, ΔT_m values up to 20 μM at different concentrations of dsDNA which was used as a competitor and at 1 μM ligand concentration. Next, the topologies of the G4 arrangements formed in the presence of the ligands were evaluated by means of CD titration analysis using 24TTA telomeric sequences in presence of K^+ or Na^+ ions, the results indicated that the best ligand (in terms of highest G4 stabilization) of the series was able, in the presence of K^+ ions, to induce a change in the G4 motif topology from hybrid to anti-parallel. Furthermore, in order to investigate more deeply the characteristic topology-inducing ability of the best ligand, different sequences, which topologies were already known, were screened in a CD titration experiment. In all cases the ligand demonstrated to be able to efficiently induce the formation of the antiparallel form, moreover the same results were obtained also varying the type of cations used (K^+ or Na^+). These results were also confirmed by an NMR and a fluorescence spectroscopy studies that also indicated the subtype of the anti-parallel topology, namely the chair subtype.

In 2021, Craciun *et al.* reported a series of new ligands for the G4 DNA. Fifteen new 1,10-phenanthroline derivatives, di-substituted at position 2 and 9 via amide bonds with different heterocycles, were prepared and tested for their ability to interact with G4 motif and for their *in vitro* anticancer activity.[293] The ligands' synthesis started with a two steps oxidation of neocuproine to 1,10-phenanthroline-2,9-dicarboxylic acid which was converted to its acid chloride derivative and reacted with different amines namely 2-aminothiazole, 2-aminopyridine, 3-aminopyridine, 5-aminoquinoline, 8-aminoquinoline and propargylamine. The so-obtained intermediates were then methylated on the heterocycle nitrogen atoms yielding a first series of ligands (Scheme 1.14).



Scheme 1.14. Synthetic scheme for the preparation of the neucuproine derivatives di-substituted with various aryl amine or with propargylamine. a: SeO_2/dioxane; b: HNO_3; c: $COCl_2$, DMF; d: $CF_3SO_3CH_3$, $CHCl_3$.

A second series of ligands was also prepared starting from the di-propargylamine substituted compound, which was reacted, by a click chemistry cycloaddition reaction (CuAAC), with a set of aromatic alkyl azides obtaining in this way 5 final compounds containing the triazole ring moiety (Scheme 1.15).



Scheme 1.15. Synthetic scheme for the preparation of the neucuproine derivatives di-substituted by side chains containing a triazole linker and different phenyl rings or with anthracene. a: CuSO4, sodium ascorbate.

The ligands were then tested against a 60 human tumor cell lines panel at the single dose of 10 μM and 4 of them exhibited a promising growth inhibitory activity against cancer cells. Due to these results, the best two compounds were selected for the same test but at different concentrations. Both the compounds showed valuable IC_{50} values up to 100 nM for some of the tested cell lines including colon cancer HCC-2998, HCT-116, HCT-15, HT29, KM12 and SW-620 cell lines, lung cancer HOP-62 and NCI-H460 cell lines and leukemia K-562 and HL-60(TB) cell lines. Circular dichroism studies were then used to investigate the interaction of the ligands with the G4 motif, in particular with two of the tested compounds a tendency to induce an antiparallel conformation was observed. A CD melting titration was also performed in order to evaluate the ability of the ligands to stabilize the G4 DNA, and the most promising result, was obtained for the methylated pyridin-3-yl substituted compound, which registered a ΔT_m value of 24 °C.

2 Synthesis via A3 Coupling Reaction of Anthracene-Propargylamine as a New Scaffold for the Interaction with DNA

2.1 Introduction

The development of DNA ligands targeting specific sites or arrangements of DNA is an aim that has been pursued for a long time in medicinal chemistry to identify both anticancer agents and chemical probes.[294] The interaction with different sequences of double stranded DNA (dsDNA) can lead to blocking enzymes like topoisomerases that are involved in the replication and transcription of the DNA causing the death of cancer cells.[295] The interaction with the double strand can generally occur following two main patterns: the binding to the major or the minor groove or the intercalation between two base pairs. As already discussed in the introduction section, another intriguing approach for DNA targeting is the stabilization of the nucleic acid in some of its so-called non-canonical structures as the G4, these arrangements arise from guanine-rich sequences such as those that are present in telomeres and promoter genes. Moreover, the stabilization of G4 structures is one of the strategies for the inhibition of telomerase, a reverse transcriptase that maintains the telomere length in cancer cells, thus supporting immortalization.[296] Tricyclic aromatic compounds are known ligands for DNA: drugs like doxorubicin and mitoxantrone share an anthracenedione scaffold and are thought to be able to form a ternary complex with the dsDNA and the topoisomerase II enzyme,[297] BRACO-19 bears an acridine structure and it is considered one of best stabilizers of the G4 arrangement.[298] Anthracene is an aromatic tricyclic compound that exhibits a blue fluorescence.[299] This molecule owns an important characteristic for a DNA ligand as it is completely planar and so able to form pi-stacking, other than hydrophobic, interactions. The role of the anthracene scaffold for the recognition of different kinds of nucleic acids has been for long investigated. The anthracenyl bishydrazone (bisantrene) is an anticancer drug used clinically for years, the molecule is known to bind strongly to dsDNA by intercalation and is an excellent topoisomerase-II poison.[300] 9-aminomethylanthracene demonstrated to bind to DNA with high affinity,[301] then the addition to the latter of a second side chain and the increase to a total of four amines led to an enhancement in its intercalation in the dsDNA.[302] Ostaszewski *et al.* synthesized a tetramethoxy anthracene bearing two central side chains ending with two primary amines which showed a good ability to bind to dsDNA with a binding constant of 4.0×10^4 M^{-1}.[303] Tan and co-workers developed two dsDNA anthracene ligands bearing two central side chains terminating with bulky tertiary amines, which had binding constants in the 10^6 M^{-1} range. In 2010, Folini and colleagues reported a series of anthracene derivatives substituted with one or two 4,5-dihydro-1H-imidazol-2-yl-hydrazonic groups (the bisantrene side chain) at different positions of the aromatic system.[304] The derivatives substituted in position 1,5 showed a good ability to stabilize the G4 DNA with ΔT_m values up to 20 °C at 5 μM concentration of the ligands. As a continuation of the latter study, Ribaudo and colleagues developed a series of constrained

bisantrene derivatives as G4 binders.[305] Using fluorimetric and ESI-MS binding assays, a high tendency of the molecules to interact with the G4 motif was observed.

Gama *et al.* investigated the role as G4 DNA ligands of a series of anthracene-terpyridine metal complexes obtaining a high selectivity for G4 over dsDNA.[306] In 2021 Coban and colleagues prepared a set of 9,10 disubstituted anthracene derivatives and probed their ability to bind to G4 DNA by NMR spectroscopy, moreover the ligands showed good inhibitory activity on the telomerase enzyme with IC_{50} values up to 18 μM against HeLa cells.

In this work, the anthracene scaffold was further explored for the development of new ligands for the canonical and non-canonical DNA arrangements. In particular, its flat-planar core, ideal for the G-quartet recognition by stacking interaction, was modified by the addition in position 1 of a propargylamine side chain. This latter, protonated at physiological pH, was intended to target the negatively charged DNA backbone by ion-pairing electrostatic interactions and by H-bond. Moreover, the rationale behind the propargylamine choice was based both on structural and synthetic reasons. The direct attachment of the alkyne moiety on the aromatic ring in fact served to further increase the overall pi-conjugation of the scaffold and to address the arm towards the G4 backbone. On the other hand, from the synthetic point of view, the alkyne moiety allowed an easy modification with the attachment of various secondary amines leading to tertiary amines.

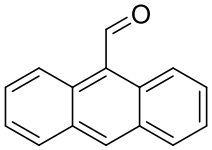
2.2 Materials and methods

General

Commercially available chemicals were purchased from Sigma-Aldrich (Saint Louis, MO, USA) and used as received, unless otherwise stated. ^{1}H and $^{13}C\{^{1}H\}$ NMR spectra were recorded on an Avance III 400 MHz spectrometer (Bruker, Billerica, MA, USA). All spectra were recorded at room temperature; the solvent for each spectrum is given in parentheses. Chemical shifts are reported in ppm and are relative to TMS internally referenced to the residual solvent peak. Datasets were edited with TopSpin (Bruker). The multiplicity of signals is reported as singlet (s), doublet (d), triplet (t), quartet (q), multiplet (m), broad (br), or a combination of any of these. Mass spectra were recorded by direct infusion ESI on a Thermo Fisher Scientific (Waltham, MA, USA) LCQ Fleet ion trap mass spectrometer. A Cary 60 UV-Vis spectrophotometer (Agilent, CA, USA) was used to record UV absorbance spectra and a Cary Eclipse spectrophotometer (Varian-Agilent, CA, USA) for fluorescence emission spectra. UV absorbance spectra were acquired using 50-100 µM compound solutions. For fluorescence experiments, based on UV absorbance spectra, excitation wavelength was set at 405 nm and compound concentration was 1 µM. Emission spectra were recorded in the 410-600 nm range, setting the instrument slits to 5 nm.

Chemistry

Synthesis of anthracene-9-carbaldehyde (**1**)

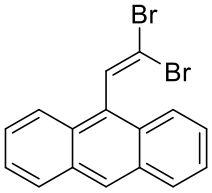


Anthracene (4 g, 22.4 mmol) was dissolved in a mixture of 4.4 mL of DMF and 5.0 mL of 1,2-dichlorobenzene, to the mixture was added under stirring $POCl_3$ (6 mL, 65.3 mmol). The solution was then heated at 100 °C for 2 h, after that it was cooled down to 0 °C and a solution of saturated sodium acetate was added until the pH reached pH 5-6, then more water was added and the mixture was stirred at r.t. for 10 minutes after that it was extracted with DCM. The organic phase was evaporated first under reduced pressure and then with a stream of nitrogen obtaining a yellow solid that was triturated with hexane (1.5 g, 32 %).

^{1}H-NMR (300 MHz, $CDCl_3$) δ: 11.54 (1H, s), 8.97-9.03 (2H, m), 8.72 (1H, s), 8.05-8.10 (2H, m), 7.66-7.73 (2H, ddd, J = 1.4 Hz, 6.6 Hz, 8.9 Hz), 7.53-7.60 (2H, ddd, J = 1.4 Hz, 6.6 Hz, 8.9 Hz).

The compound characterization was in accordance with literature data.[307]

Synthesis of 9-(2,2-dibromovinyl)anthracene (**2**)

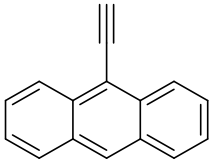


Compound **1** (710 mg, 3.5 mmol) and CBr_4 (2.317 g, 7.0 mmol) were dissolved in 10 mL of DCM, the stirred mixture was cooled down to 0 °C and $P(Ph)_3$ (3.654 g, 14.0 mmol) was added in small portions over a period of 5 minutes. The solution was stirred in ice for 2 h after that TLC (DCM) showed complete conversion of compound **1** to compound **2**. The mixture was brought to r.t. and washed with brine, after that the organic phase was evaporated under reduced pressure giving a viscous dark red liquid that was purified by column chromatography on silica gel (DCM) yielding a yellow solid (616 mg, 49 %).

^{1}H-NMR (300 MHz, $CDCl_3$) δ: 8.51 (1H, s), 8.13 (1H, s), 8.00-8.11 (4H, m), 7.47-7.59 (4H, m).

The compound characterization was in accordance with literature data.[308]

Synthesis of 9-ethynylanthracene (**3**)



Compound **2** (274 mg, 0.76 mmol) was added to 1.5 mL of ACN, after that DBU (0.42 mL, 2.8 mmol) and KOH (500 mg, 8.91 mmol) were added, the mixture was stirred at r.t. for 16 h then the solution was cooled down to 0 °C and it was added first 20 mL of deionized water then 5 M HCl to quench the DBU. The mixture was extracted with DCM and the organic phase was evaporated under reduced pressure, the solid obtained was purified by column chromatography on silica gel (ethyl acetate/hexane 1/4), a light yellow crystalline product was obtained (70 mg, 46 %)

^{1}H-NMR (300 MHz, $CDCl_3$) δ: 8.58-8.64 (2H, m), 8.45 (1H, s), 7.98-8.04 (2H, m), 7.64-7.57 (2H, ddd, J = 1.4 Hz, 6.6 Hz, 8.7 Hz), 7.48-7.55 (2H, ddd, J = 1.4 Hz, 6.6 Hz, 8.7 Hz), 4.00 (1H, s).

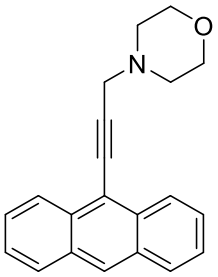
The compound characterization was in accordance with literature data.[309]

General synthesis of compounds **4a-4f**

Compound **3** (0.25 mmol) was dissolved in 0.5 mL of DMSO, then it was added 100 µL of formaldehyde solution (35%), the secondary amine (0.3 mmol) and CuI (0.02 mmol, 3 mg). The reaction flask was deareated, fluxed with nitrogen, tightly stoppered and stirred at r.t. for 6 h. The

reaction progress was checked by TLC (DCM/MeOH/TEA 94.5-97/0,25-0,5/0,5), when the reaction was finished DCM was added and the mixture was washed with 1 M KOH, the organic phase was then evaporated at reduced pressure and the solid obtained was purified by column chromatography on silica gel (DCM/MeOH/TEA 94.5-97/0,25-0,5/0,5).

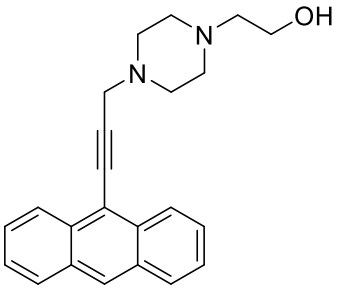
Synthesis of 4-(3-(anthracen-9-yl)prop-2-yn-1-yl)morpholine (**4a**)



The product obtained with the procedure described above, using morpholine as the reactive secondary amine, was a crystalline yellow solid (56 mg, 74%).

^{1}H-NMR (400 MHz, $CDCl_3$) δ: 8.53-8.57 (2H, m), 8.41 (1H, s), 7.98-8.02 (2H, m), 7.60-7.53 (2H, ddd, J = 1.4 Hz, 6.6 Hz, 8.7 Hz), 7.53-7.47 (2H, ddd, J = 1.4 Hz, 6.6 Hz, 8.7 Hz), 3.89 (2H, s), 3.85 (2H, t, 4.7 Hz), 2.83 (2H, t, 4.7 Hz). ^{13}C (100 MHz, $CDCl_3$) δ: 133.9, 131.3, 127.7, 126.8, 126.7, 117.2, 95.5, 82.4, 67.1, 52.7, 48.8. ESI-MS m/z: Calcd for $C_{21}H_{20}NO^+$ ($M+H^+$): 302.15, found 302.04.

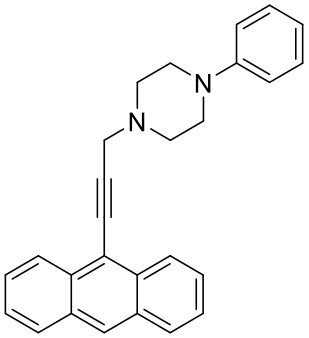
Synthesis of 2-(4-(3-(anthracen-9-yl)prop-2-yn-1-yl)piperazin-1-yl)ethan-1-ol (**4b**)



The product obtained with the procedure described above, using N-hydroxyethyl piperazine as the reactive secondary amine, was a crystalline yellow solid (60 mg, 70%).

^{1}H-NMR (400 MHz, $CDCl_3$) δ: 8.53-8.55 (2H, m), 8.40 (1H, s), 7.96-8.01 (2H, m), 7.60-7.54 (2H, ddd, J = 1.4 Hz, 6.6 Hz, 8.7 Hz), 7.53-7.45 (2H, ddd, J = 1.4 Hz, 6.6 Hz, 8.7 Hz) 3.88 (2H, s), 3.69 (2H, t, J = 5.3 Hz), 2.93 (4H, m), 2.79 (4H, m), 2.67 (2H, t, J = 4.7 Hz). ^{13}C (100 MHz, $CDCl_3$) δ: 132.8, 131.1, 128.7, 127.6, 126.7, 126.6, 125.6, 117.0, 95.4, 82.2, 59.5, 57.5, 52.9, 51.6, 48.2. ESI-MS m/z: Calcd for $C_{23}H_{25}N_2O^+$ ($M+H^+$): 345.20, found 345.25.

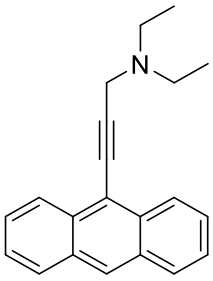
Synthesis of 1-(3-(anthracen-9-yl)prop-2-yn-1-yl)-4-phenylpiperazine (**4c**)



The product obtained with the procedure described above, using phenylpiperazine as the reactive secondary amine, was a crystalline yellow solid (53 mg, 56%).

^{1}H-NMR (400 MHz, $CDCl_3$) δ: 8.54-8.60 (2H, m), 8.42 (1H, s), 7.98-8.03 (2H, m), 7.60-7.54 (2H, ddd, J = 1.4 Hz, 6.6 Hz, 8.7 Hz), 7.54-7.47 (2H, ddd, J = 1.4 Hz, 6.6 Hz, 8.7 Hz), 7.25-7.32 (2H, m), 6.94-7.00 (2H, m), 6.84-6.90 (2H, tt, J = 1.0 Hz, 7.3 Hz), 3.96 (2H, s), 3.35 (4H, t, J = 5.2 Hz), 3.00 (4H, t, J = 5.2 Hz). ^{13}C (100 MHz, $CDCl_3$) δ: 151.3, 132.8, 131.1, 129.1, 128.7, 127.5, 126.7, 126.6, 125.6, 119.8, 117.2, 116.2, 95.6, 82.2, 52.3, 49.2, 48.4. ESI-MS m/z: Calcd for $C_{27}H_{25}N_2^+$ $(M+H)^+$: 377.20, found 377.04.

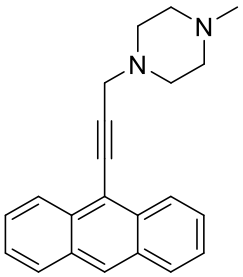
Synthesis of 3-(anthracen-9-yl)-N,N-diethylprop-2-yn-1-amine (**4d**)



The product obtained with the procedure described above, using diethylamine as the reactive secondary amine, was a crystalline yellow solid (47 mg, 65%).

^{1}H-NMR (400 MHz, $CDCl_3$) δ: 8.53-8.59 (2H, m), 8.40 (1H, s), 7.97-8.03 (2H, m), 7.60-7.53 (2H, ddd, J = 1.4 Hz, 6.6 Hz, 8.7 Hz), 7.53-7.46 (2H, ddd, J = 1.4 Hz, 6.6 Hz, 8.7 Hz), 4.03 (2H, s), 2.82 (2H, q, J = 7.2 Hz), 1.25 (3H, t, J = 7.2 Hz) ^{13}C (100 MHz, $CDCl_3$) δ: 132.7, 131.2, 128.7, 127.2,126.8, 126.4, 125.6, 117.6, 95.8, 81.6, 47.7, 41.9, 12.9. ESI-MS m/z: Calcd for $C_{21}H_{22}N^+$ $(M+H)^+$: 288.17, found 288.01.

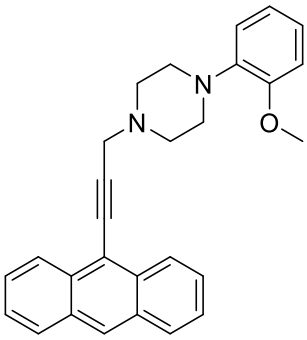
Synthesis of 1-(3-(anthracen-9-yl)prop-2-yn-1-yl)-4-methylpiperazine (**4e**)



The product obtained with the procedure described above, using N-methylpiperazine as the reactive secondary amine, was a crystalline yellow solid (20 mg, 25 %).

^{1}H-NMR (400 MHz, $CDCl_3$) δ: 8.50-8.58 (2H, m), 8.40 (1H, s), 7.95-8.02 (2H, m), 7.61-7.52 (2H, ddd, J = 1.4 Hz, 6.6 Hz, 8.7 Hz), 7.53-7.44 (2H, ddd, J = 1.4 Hz, 6.6 Hz, 8.7 Hz), 3.89 (2H, s), 2.91 (4H, m), 2.66 (4H, m), 2.38 (3H, s) ^{13}C (100 MHz, $CDCl_3$) δ: 132.7, 131.2, 128.7, 127.2,126.8, 126.4, 125.6, 117.6, 95.8, 81.6, 47.7, 41.9, 12.9. ESI-MS m/z: Calcd for $C_{22}H_{23}N_2^+$ $(M+H)^+$: 315.19, found 315.02.

Synthesis of 1-(3-(anthracen-9-yl)prop-2-yn-1-yl)-4-(2-methoxyphenyl)piperazine (**4f**)



The product obtained with the procedure described above, using 1-(2-methoxyphenylpiperazine) as the reactive secondary amine, was a crystalline yellow solid (40 mg, 39%).

^{1}H-NMR (400 MHz, $CDCl_3$) δ: 8.56-8.63 (2H, m), 8.41 (1H, s), 7.96-8.04 (2H, m), 7.61-7.54 (2H, ddd, J = 1.2 Hz, 6.7 Hz, 8.7 Hz), 7.54-7.46 (2H, ddd, J = 1.2 Hz, 6.7 Hz, 8.7 Hz), 6.84-7.05 (4H, m), 3.95 (2H, s), 3.25 (2H, m), 3.05 (2H, m). ^{13}C (100 MHz, $CDCl_3$) δ: 132.8, 131.1, 128.6, 127.5, 126.8, 126.6, 125.6, 117.1, 95.7, 82.1, 55.0, 51.9, 48.3, 45.7. ESI-MS m/z: Calcd for $C_{28}H_{27}N_2O^+$ $(M+H)^+$: 407.21, found 407.11.

Molecular docking

The G4 structure files were retrieved from the RCSB Protein Data Bank (PDB, www.rcsb.org) (PDB IDs: 3UYH, 1KF1) and were prepared with the Protein Preparation Wizard included in the

Schrödinger suite,[310] using default settings, i.e., adding hydrogens, removing waters further than 5 Å from the ligand, adjusting charges, capping termini, optimizing hydrogen bond clusters and performing a final minimization under the OPLS3e force field.[311] The ligands were prepared for docking with Ligprep tool under OPLS3e force field, using Epik ionizer all possible states in the 7±2 pH range were generated, including the possible metal binding states, allowing therefore the interaction with the potassium ions. The docking protocol consisted of rigid-receptor/flexible ligand docking, which was performed with Glide, under the OPLS3e force field, with default settings (*i.e.*, a 0.80 scaling factor for the ligand atom van der Waals radius and 0.15 for the partial charge cut-off, with generation of 1 pose per ligand). For the docking to the intramolecular *holo* G4 (PDB ID: 3UYH) one cubic grid was generated, with Receptor Grid Generation app included in Schrödinger suite, using as the grid site the centroid of the ligand present in the crystal structure and default settings. The precision for the docking was set to SP.

For the docking to the intramolecular *apo* G4 (PDB ID: 1KF1) two cubic grids were generated, with Receptor Grid Generation app included in Schrödinger suite, selecting either the upper or the lower external quartets as the centre, the dimensions of each side of the grids were 15 Å for the internal and 35 Å for the external. The precision for the docking was set to XP.

ESI-MS binding assays

DNA sequences were obtained from Sigma-Aldrich (Milan, Italy) and were purchased with the "desalting" purification grade. Samples were heat-denatured and folded in 150 mM ammonium acetate. A slow annealing procedure was followed: oligonucleotide solutions were heated at 100°C for 3 minutes and slow cooling to room temperature followed. Samples were then stored in fridge (4 °C) for at least one week before incubation with the ligands. Stock solutions of the compounds were prepared in methanol (15 mM) and were diluted in 150 mM ammonium acetate prior to use (1:10). In the final sample, the concentration of the oligonucleotide was 5-10 μM in 150 mM ammonium acetate, with a 10:1 compound/oligo ratio (methanol percentage present in the incubated sample: < 0.5 %). Samples were acquired after an equilibration time of 30 minutes. Mass spectra were recorded by direct infusion ESI on a Thermo Fisher Scientific (Waltham, MA) LCQ Fleet ion trap mass spectrometer. The instrument was set in negative ionization mode with a 3.4 kV capillary voltage, 120 °C capillary temperature and a flow rate of 5 μL/min. Binding affinity (BA) was calculated according to the formula: $BA = (\Sigma DNA_{bound}/(\Sigma DNA_{free} + \Sigma DNA_{bound})) \times 100$. Since the resulting value is an expression of the bound fraction of the DNA with respect to the total DNA (free and bound), internal standard was not adopted in this experiment. For the calculation of BA, signals of both 1:1 and 2:1 complexes were considered. Collision-induced dissociation (CID) experiments were performed on the complexes by isolating the precursor ion in the trap. $\text{Relative I(\%)} = (I_{COMPLEX}/(I_{COMPLEX}+I_{DISSOCIATION\ PRODUCTS})) \times 100$.

Fluorimetric titration assay

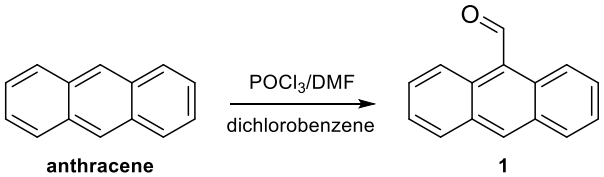
Titration experiments were carried out on a Varian Cary Eclipse spectrophotometer at 25 °C. Ligand-DNA binding was evaluated by addition of increasing amounts of folded DNA (G4 or dsDNA) to a freshly prepared 1 μM compound solution in 150 mM ammonium acetate in a 3 mL quartz cell. Titrant concentration was 100 μM in 150 mM ammonium acetate. Excitation wavelength was set to 405 nm and emission spectra were recorded in the 410-600 nm range, setting the instrument slits to 5 nm. The intensity of the emission peak at 435 nm was considered for the calculations.

2.3 Results and discussion

For the preparation of the mono substituted compounds, the alkyne in position 9 was introduced by *Vilsmeier-Haack* formylation followed by a *Corey-Fuchs* reaction. This intermediate was modified with different secondary amines with a Cu(I)-catalysed A3 coupling reaction, which is a transition-metal catalyzed three-component coupling reaction of amines, aldehydes, and terminal alkynes that produces propargylamines. Moreover, it is characterized by high atom economy (with water as the only by-product), versatility, applicability to a wide range of substrates, chemoselectivity and high product yields. All the compounds of the present study were designed performing molecular docking calculations exploring the binding to dsDNA and G4 motifs. Their predicted physicochemical properties were also considered and calculated. Then, the interaction of the obtained compounds with dsDNA and G4 DNA sequences was studied by ESI-MS, obtaining their selectivity ratio for the two forms. Moreover, collision-induced dissociation (CID) experiments were used to investigate the relative gas-phase kinetic stability of small molecule-DNA complexes. In conclusion of the study, a fluorescence titration experiment was also used as a second independent experimental technique to confirm that the compounds can effectively interact with DNA in solution.

Chemistry

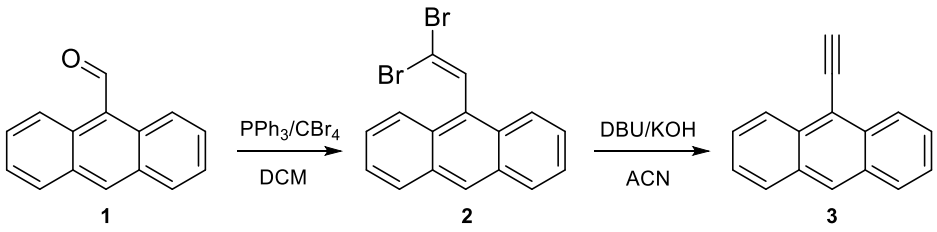
The designed compounds (reported in Figure 2.1), were synthesized as illustrated in the following schemes (Scheme 2.1, 2.2 and 2.5). Anthracene aldehyde derivative **1** was obtained by *Vilsmeier-Haack* formylation of anthracene. This reaction is highly regioselective towards the 9 position of the tricyclic system, a nucleophilic site that reacts readily with the electrophile *Vilsmeier* reagent formed *in situ* with $POCl_3$ and DMF.



Scheme 2.1. Synthesis of compound **1** by *Vilsmeier-Haack* formylation reaction of anthracene.

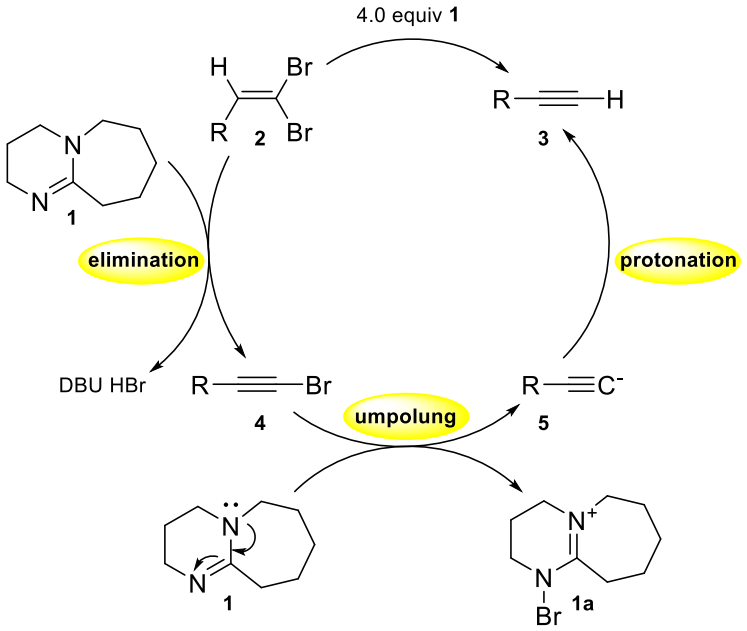
The synthetic scheme proceeds with the alkyne formation accomplished with a variant of the *Ramirez-Corey-Fuchs* two-step reaction. The aldehyde was treated with triphenylphospine and CBr_4 to produce the dibrominated alkene 2 via *Ramirez* olefination. The reaction proceeded smoothly and

quantitatively yielding the desired intermediate **2**. In the classical *Ramirez-Corey-Fuchs* reaction the alkyne formation is usually accomplished using n-butyllithium in apolar solvents at -78 °C: this induces the production, by lithium-halogen exchange, of a reactive intermediate that is subjected to alpha-elimination of the bromide producing a carbene that readily rearranges to a terminal alkyne.[312] In this work, the elimination step was performed with an already reported alternative procedure.[313] Compound **2** was treated with 4 equivalents of DBU together with solid KOH in acetonitrile at r.t. for 48 h, giving intermediate **3** with good yield (Figure 2.2).



Scheme 2.2. Synthesis of **3** from **1** by *Ramirez-Corey-Fuchs* reactions.

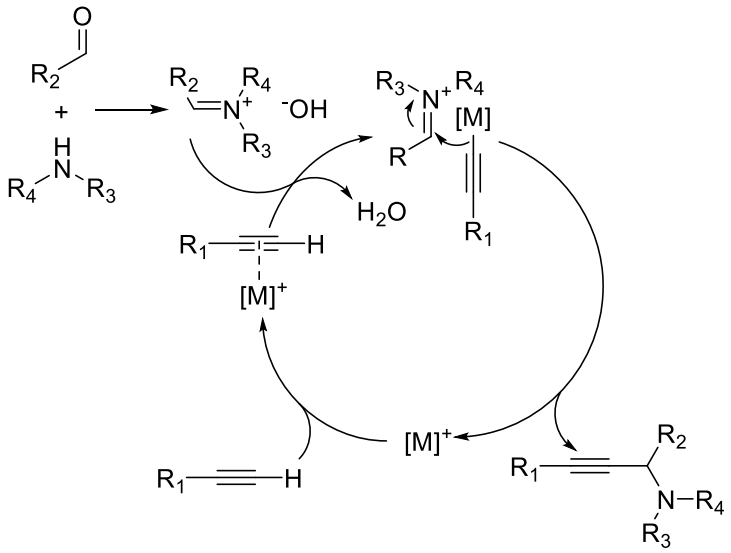
A possible reaction mechanism (Scheme 2.3) comprehends a first elimination reaction forming a bromoalkyne intermediate which is subjected to a nucleophilic attack by a second DBU molecule forming an acetylide ion that is protonated upon aqueous work-up.[314]



Scheme 2.3. Plausible mechanism for the *Corey-Fuchs* reaction with DBU acting first as a base and then as a nucleophile.

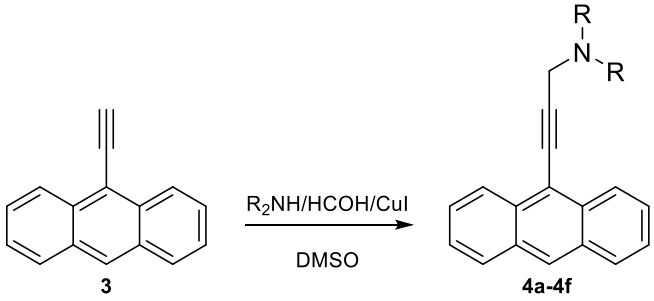
The final step for the synthesis of the compounds (reported in Scheme 2.5) consists in the Cu(I) - catalyzed A3 coupling reaction between compound **3**, formaldehyde and various secondary amines (*i.e.* morpholine, N-hydroxyethylpiperazine, phenylpiperazine, diethylamine, N-methyl piperazine and 2-methoxy phenylpiperazine).

The multi-component reaction is thought to proceed similarly to a Mannich reaction. At first the alkyne activation happens forming a copper acetylide, which performs a nucleophilic addition on the intermediate formed by the reaction of formaldehyde and a secondary amine, yielding the propargylamine moiety (Scheme 2.4).[315,316]



Scheme 2.4. Proposed mechanism for the A3 coupling reaction with secondary amines.

DMSO was used as a solvent to dissolve all the species and the reaction went to completion usually in less than 3 hours with high selectivity and good to high yields for the desired products **4a-f** (Scheme 2.5).



3 **4a-4f**

Scheme 2.5. Synthesis of compounds **4a-4f** by A3 coupling reaction from compound **3**.

The six final compounds, bearing different tertiary amines (**4a**: morpholine, **4b**: N-hydroxyethyl piperazine, **4c**: phenylpiperazine, **4d**: diethylamine, **4e**: N-methyl piperazine and **4f**: 2-methoxy phenylpiperazine) are reported in the figure below (Figure 2.1).

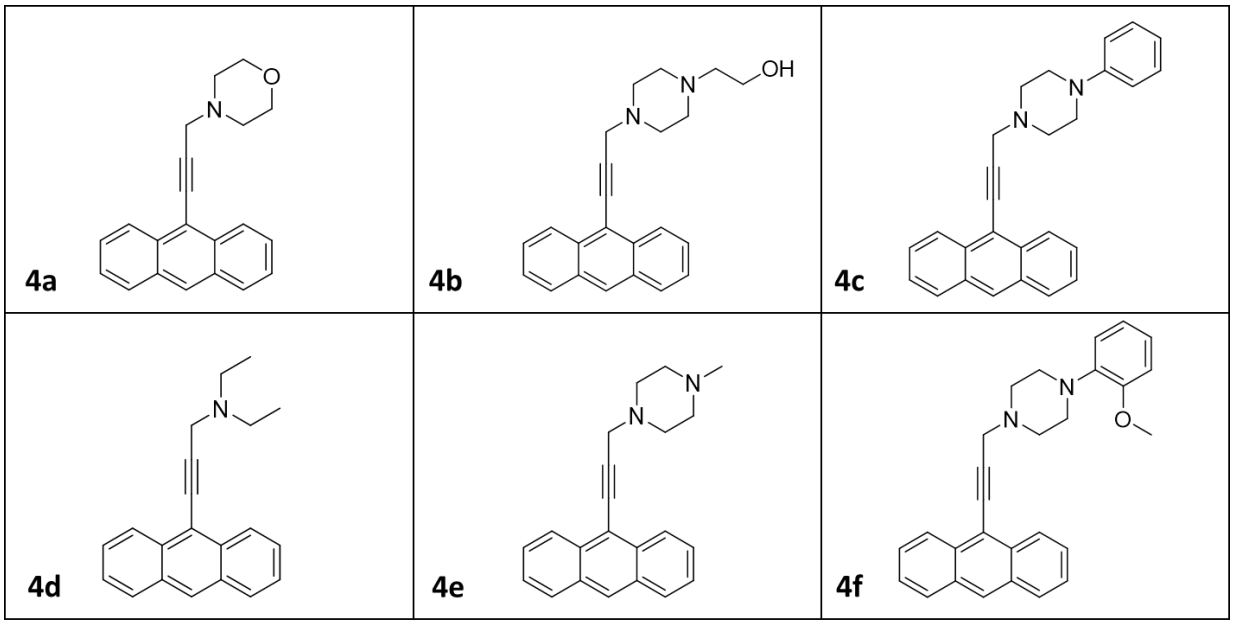


Figure 2.1. Chemical structures of compounds **4a-4f**.

Physicochemical properties

The designed compounds were screened for their physicochemical properties, that were predicted using Molinspiration Cheminformatics software[317] which takes in account parameters that can directly influence properties such as the solubility and cell permeability, like molecular weight, partition coefficient and the number of hydrogen bond donor or acceptor atoms.[318] The calculated physicochemical properties of the molecules **4a-4f,** reported in the table below (Table 1.1) demonstrated to fall within the guidelines from the most recent reports on drug-likeness.[318]

Table 1.1. Calculated physicochemical properties. miLogP: octanol-water partition coefficient, TPSA: total molecular polar surface area, natoms: atoms number in the molecule, MW: molecular weight of the molecule, nON: number of nitrogen and oxygen atoms of the molecule, nOHNH: number of hydrogen atoms bound to oxygen and nitrogen atoms, nviolations: number of violations of the parameters with respect to Lipinski's rules, nrotb: number of rotable bonds, volume: molecular volume.

Compound	miLogP	TPSA (Å^2)	natoms	MW (g/mol)	nON	nOHNH	nviolations	nrotb	Volume (Å^3)
4a	4.31	12.47	23	301.39	2	0	0	1	289.11
4b	3.72	26.70	26	344.46	3	1	0	3	334.53
4c	6.05	6.48	29	376.50	2	0	1	2	364.31
4d	5.21	3.24	22	287.41	1	0	1	3	290.48
4e	4.35	6.48	24	314.43	2	0	0	1	309.47
4f	6.06	15.71	31	406.53	3	0	1	3	389.86

Molecular docking

The molecular docking calculations were performed with Schrödinger Maestro 2019 through its embedded Glide docking algorithm.[319] The d(CGTACG) sequence, crystallized with a morpholino doxorubicin ligand, was used as a model for the dsDNA (PDB ID: 2DES). On the other hand, for the human telomeric DNA, the crystallized structure of a human telomeric G4 with a naphthalene diimide (MM41) ligand was used (PDB ID: 3UYH). First of all, to validate the docking protocol, for both the structures their native ligands were re-docked obtaining poses that fully resembles the native pose.

The binding energies of the compounds **4a-4f** for the formation of the complexes with the nucleic acids structures were calculated in terms of Glide Score and are reported in the table below (Table 2.2). For both the DNA arrangements **4b** was predicted as the best ligand. As depicted in Figure 2.2, the molecule intercalates in the dsDNA and stacks on the external G-quartet of the G4 arrangement. Concerning dsDNA, **4b** is followed by the N-methyl piperazine derivative **4e**. For G4 complexes, the second-best ligand is the diethylamino derivative **4d**.

In a second moment, in order to harmonize the computational data with those obtained in the later studies, which will be presented in the next chapters in this thesis, and also to increase the significance of the investigation, an additional docking experiment was performed. Specifically, an intramolecular parallel G4 structure deriving from the telomeric DNA sequence (5'-$AG_3T_2AG_3T_2AG_3T_2AG_3$-3) almost identical to the one used in the ESI-MS assay (5'-$AG_3T_2AG_3T_2AG_3T_2AG_3T$-3), was used as G4 receptor. Interestingly, also for this receptor, the best ligand in terms of Glide Score resulted **4b** with the value of -8.8 kcal/mol. Also for this other parallel G4 structure, **4b** demonstrated a similar behaviour of interaction. In particular, the anthracene core positioned itself on top of an external G-quartet with which it developed pi-stacking interactions and its polar sidechain was directed towards the G4's grooves (Figure 2.3).

Table 2.2. Glide Score values calculated for the binding interaction of compounds **4a-4f** with G4 and dsDNA (kcal/mol).

Compound	Glide Score G4 (PDB ID: 3UYH)	Glide Score G4 (PDB ID: 1KF1)	Glide Score dsDNA
4a	-5.8	-5.3	-8.2
4b	-7.7	-8.8	-9.2
4c	-4.7	-6.2	-7.8
4d	-6.6	-5.1	-8.0
4e	-6.2	-8.0	-8.6
4f	-6.4	-6.3	-7.2

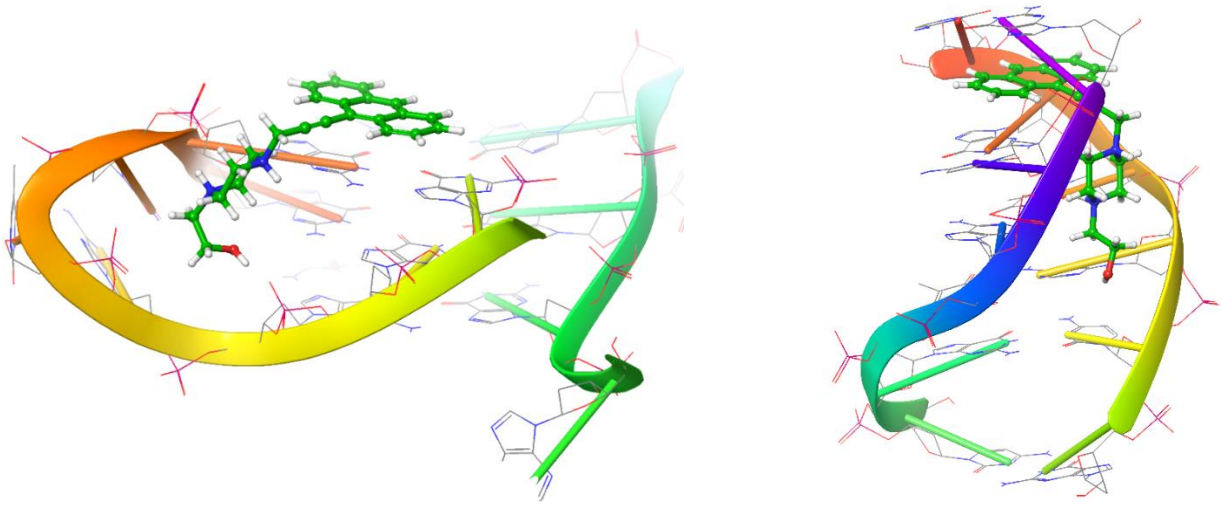


Figure 2.2. Binding poses for **4b** in complex with G4 (PDB ID: 3UYH) (left) and dsDNA (right).

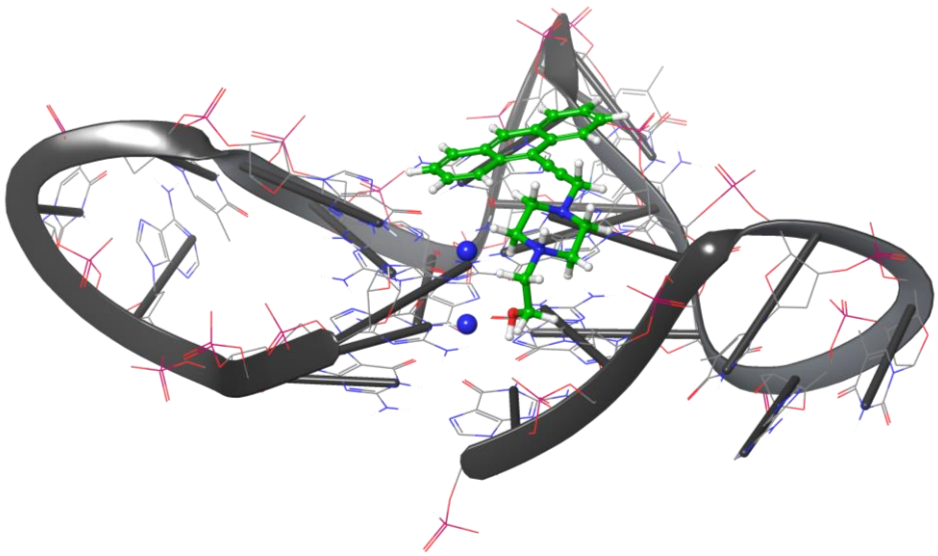


Figure 2.3. Binding pose for **4b** in complex with G4 (PDB ID: 1KF1).

ESI-MS binding study

Using an ESI-MS experiment, the 6 anthracene derivatives were tested against a G4 and a dsDNA forming sequences and their ability to form a small molecule-nucleic acid complex was evaluated.[320,321] In particular, the G-rich DNA oligonucleotide containing the telomeric TTAGGG repeat (5'-$AG_3T_2AG_3T_2AG_3T_2AG_3T$-3) and the dsDNA (5'-ACTATTTACGTATAATGA-3', 5'-TCATTATACGTAAATAGT-3') sequences were used in the experiments. For all the ligands, the BA affinity was then calculated. All the compounds showed an overall preference for dsDNA, with **4b** showing similar and relatively high BA values towards both the oligonucleotides (Table 2.3). In all the binding experiments, mitoxantrone, a dsDNA intercalant was used as a positive control. An illustrative example of a ligand-G4 interaction experiment is reported in the figure below (Figure 2.4).

Table 2.3. Binding affinity (BA) values for compounds **4a-4f** and for mitoxantrone towards G4 and dsDNA, as calculated from ESI-MS binding experiment. Selectivity ratio was calculated dividing G4 BA value by dsDNA BA value. For the BA calculations, the peaks corresponding to the different states of charge and 1:1 or 2:1 stoichiometries were considered.

Compound	BA G4	BA dsDNA	Selectivity ratio (G4/dsDNA)
4a	41.7	65.5	0.64
4b	68.4	79.8	0.86
4c	33.6	49.3	0.68
4d	47.3	61.5	0.77
4e	48.4	85.9	0.56
4f	28.3	45.8	0.62
mitoxantrone	27.2	45.5	0.60

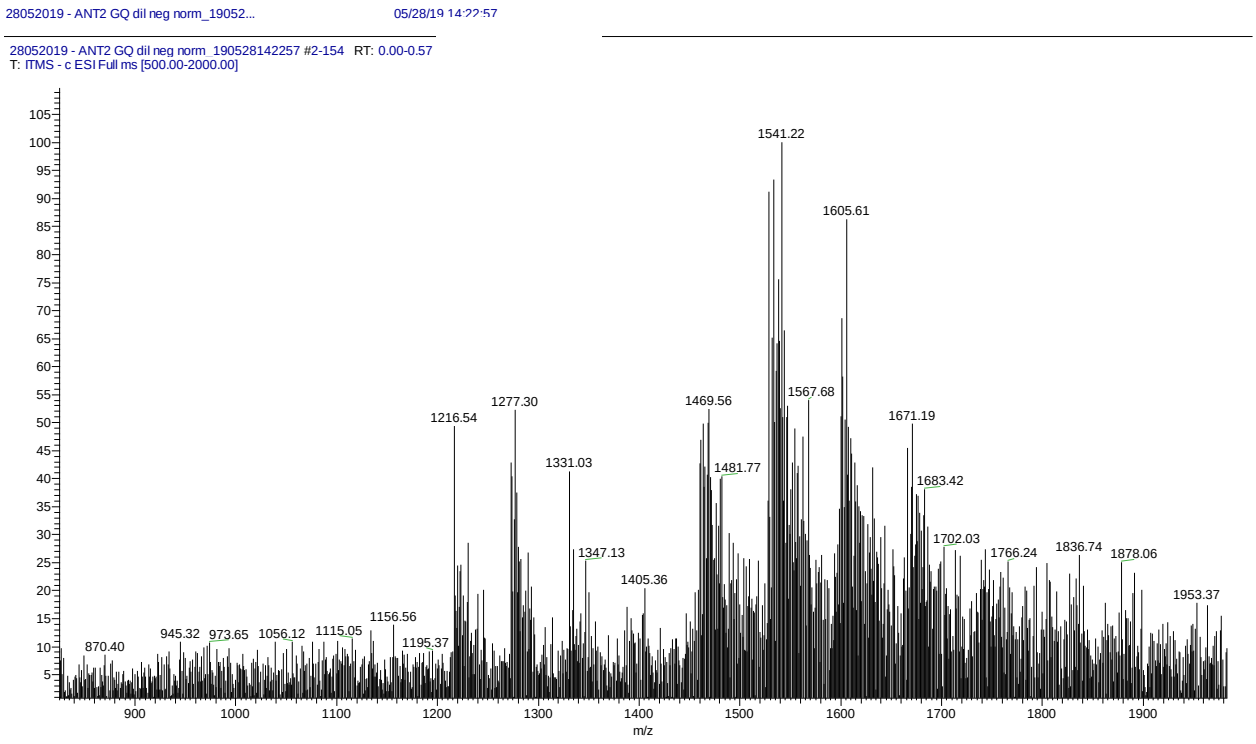


Figure 2.4. ESI-MS experiment showing the interaction of compound **4b** with G4 DNA (z = -5 and z= -6).

Two additional competition experiments were performed to directly compare the DNA-binding properties of the compounds. In the first experiment, the 3 compounds with higher BA towards G4 (**4b, 4d, 4a**) were mixed and incubated with the oligonucleotide. The sample was analyzed under the same conditions and for **4b** the observations of the previous experiments were confirmed, as even if in the presence of the other ligands **4a** and **4d**, it demonstrated to be the best binder of the

group. Then **4b** was incubated with an equimolar solution of dsDNA and G4 and, as already observed in the first experiment, a limited preference for dsDNA was highlighted.

Collision-induced dissociation (CID) experiments were used to investigate the relative gas-phase kinetic stability of complexes, in terms of $E_{COM}^{50\%}$ which is calculated from the dissociation curve obtained from the relative intensities of the adduct and of the dissociation products in the mass spectrum at different collision energy values. Since the formation of 1:1 and 2:1 ligand/DNA complexes were observed in most of the cases, the stability of the complexes with different stoichiometry was investigated. Compound **4a** and **4d** demonstrated to efficiently stabilize G4 in the 2:1 stoichiometry. **4b**, again, showed similar results in the G4 and dsDNA stabilization, but with limited efficiency in the 2:1 complex. No 2:1 ligand/G4 complex was detected for bulkier compounds **4c** and **4f**. The results are reported in the table below (Table 2.4).

Table 2.4. Relative gas-phase kinetic stability ($E_{COM}^{50\%}$) investigated by collision-induced dissociation (CID).

Compound	$E_{COM}^{50\%}$ 2:1 ligand/G4 (eV)	$E_{COM}^{50\%}$ 1:1 ligand/G4 (eV)	$E_{COM}^{50\%}$ 2:1 ligand/dsDNA (eV)
4a	60.67	36.29	47.39
4b	34.57	40.07	39.43
4c	-	39.98	76.01
4d	67.02	25.79	59.08
4e	19.59	22.37	32.46
4f	-	29.54	58.40
mitoxantrone	-	44.40	44.64

The overall experimental approach was validated using mitoxantrone, which is known to directly interact by intercalation with DNA, as a positive control. In particular, BA was calculated using the same ESI-MS binding experiment described above. As expected, mitoxantrone interacts preferentially with dsDNA, giving a strong 2:1 ligand/DNA complex signal (Table 2.3). Although, interaction with G4 was also observed with a 1:1 stoichiometry (Table 2.4 and Figures S48 and S50). The gas-phase kinetic stability of these complexes was evaluated, showing similar values for 1:1 ligand/G4 and 2:1 ligand/dsDNA (Table 2.4). Interestingly the best ligand **4b** outperformed mitoxantrone in terms of BA for either the G4 and the dsDNA arrangements.

Fluorescence binding assay

In parallel, the interaction of **4b**, the best DNA binder of the set, was investigated using fluorescence titration as a second, independent experimental technique to confirm that the compounds can effectively interact with DNA in solution (see Figure S53 for spectra). The experiments confirmed the observations from docking and ESI-MS, highlighting that **4b** interacts preferentially with dsDNA. The binding experiments with duplex and G4 DNA showed a decrease in fluorescence resulting from the interaction. The intrinsic binding constants of the compound/DNA complexes were obtained by fitting the experimental fluorescence values as a function of the molar concentration of duplex and G4 DNA using a previously reported binding model.[31,32] The obtained values for **4b**-DNA complexes are reported in Table 2.5.

DNA sequence	**K_b (M^{-1})**
G4	$(3.8 \pm 0.5) \times 10^5$
dsDNA	$(1.3 \pm 0.3) \times 10^6$

Table 2.5. DNA-binding constants (K_b) for compound **4b** towards G4 and dsDNA.

2.4 Conclusion

A set of new DNA ligands was prepared through an efficient synthetic approach based on the A3 coupling reaction. This versatile synthetic pathway allows the preparation of planar, fluorescent and easily derivatizable DNA binders and allows the future expansion of the series through scaffold optimization. The synthesized compounds were initially studied *in silico* and their predicted physicochemical properties suggested that they satisfy the main drug-likeness criteria. Concerning their interaction with DNA, molecular docking calculations showed a general preference of the molecules for the dsDNA with respect to the G4 arrangement. Mass spectrometry binding experiments proved that the compounds bind both DNA structures with a general preference for dsDNA that confirms the docking results. This preliminary observation was also supported by fluorescence titrations for compound **4b**. Nevertheless, the relative gas-phase kinetic stability study highlighted that the complexes formed by compounds **4a**, **4d** and **4b** with G4 are more stable, which is crucial for triggering biological effects. These intriguing findings pave the way for a future development of this novel DNA targeting scaffold. Given the high ability to interact with the dsDNA and the good calculated physicochemical properties, these promising ligands, which in some assays outperformed the famous intercalant drug mitoxantrone, will be tested against various cancer cell lines. On the other hand, for what concerns the interaction with G4 DNA, an optimization of the scaffold appears necessary in order to increase the selectivity of the ligands for this non-canonical fold.

2.5 Supporting material

*Characterization of compounds **4a-4f***

Docking study

ESI-MS binding study

Fluorimetric titrations

*Characterization of compound **4a***

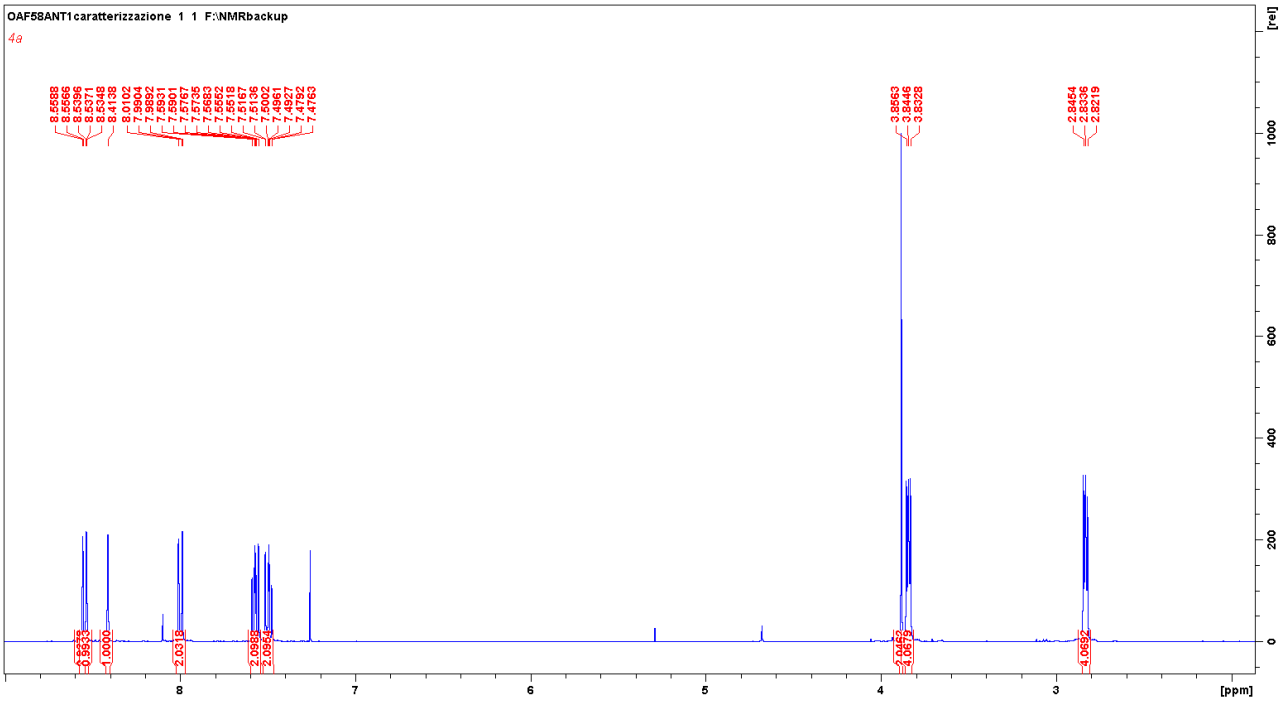


Figure S2.1. ^{1}H-NMR spectrum of compound **4a**.

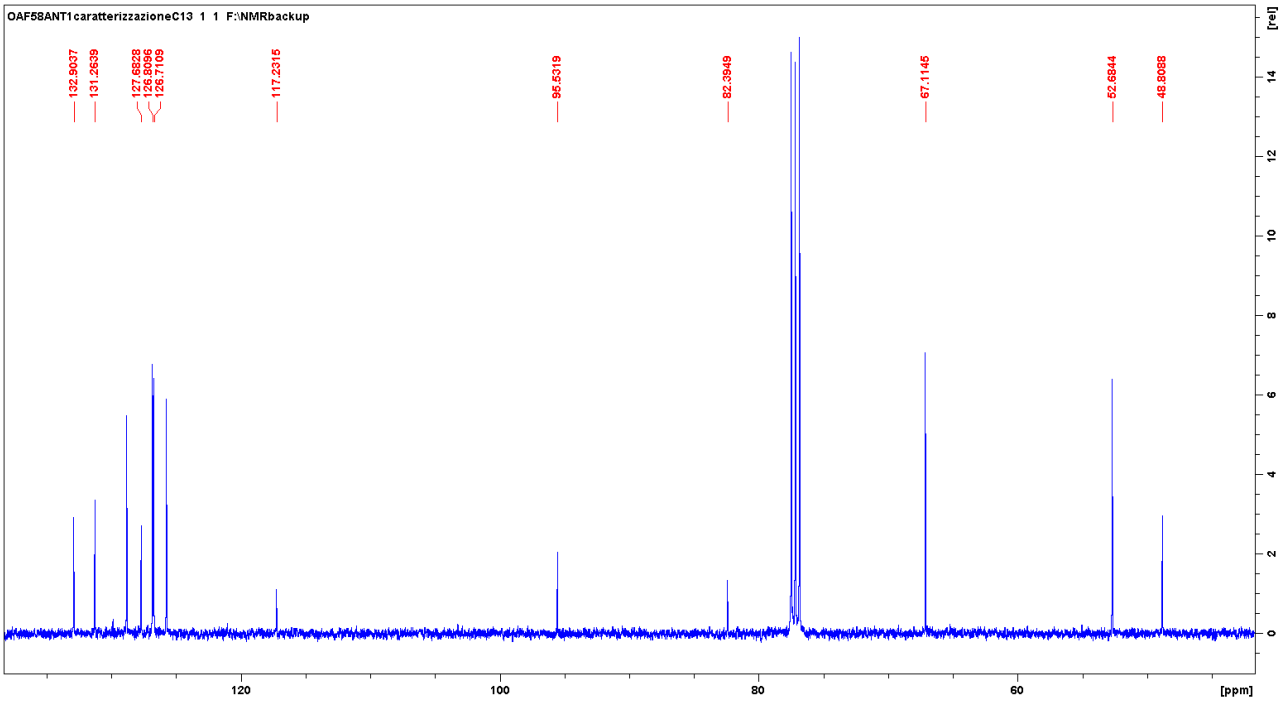


Figure S2.2. ^{13}C-NMR spectrum of compound **4a**.

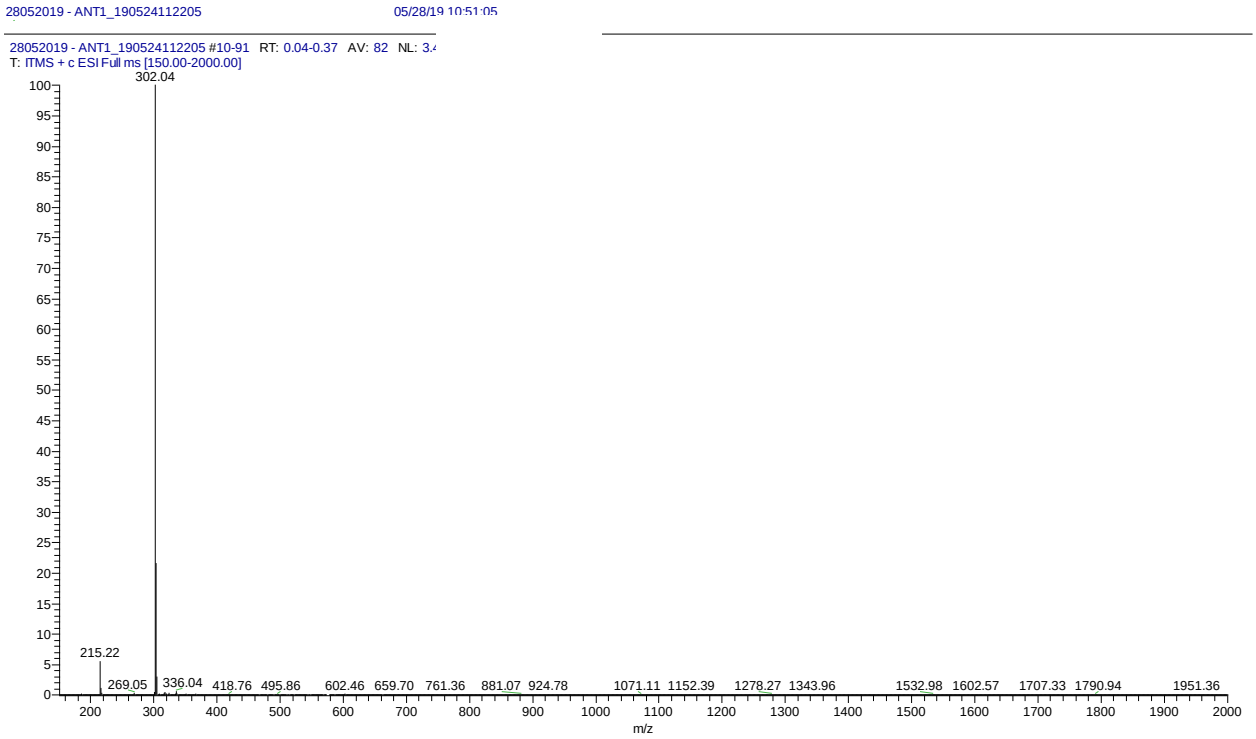


Figure S2.3. ESI-MS spectrum of compound **4a**.

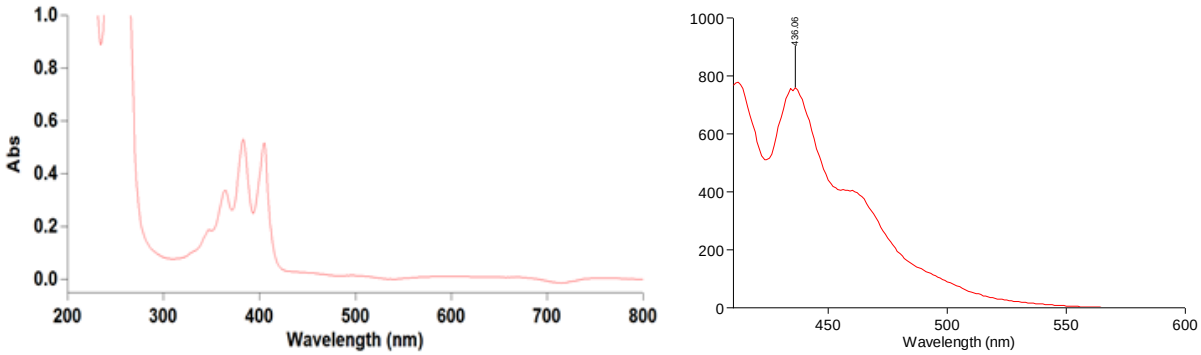


Figure S2.4. UV-Vis absorbance spectrum (left) and fluorescence emission spectrum (right).

*Characterization of compound **4b***

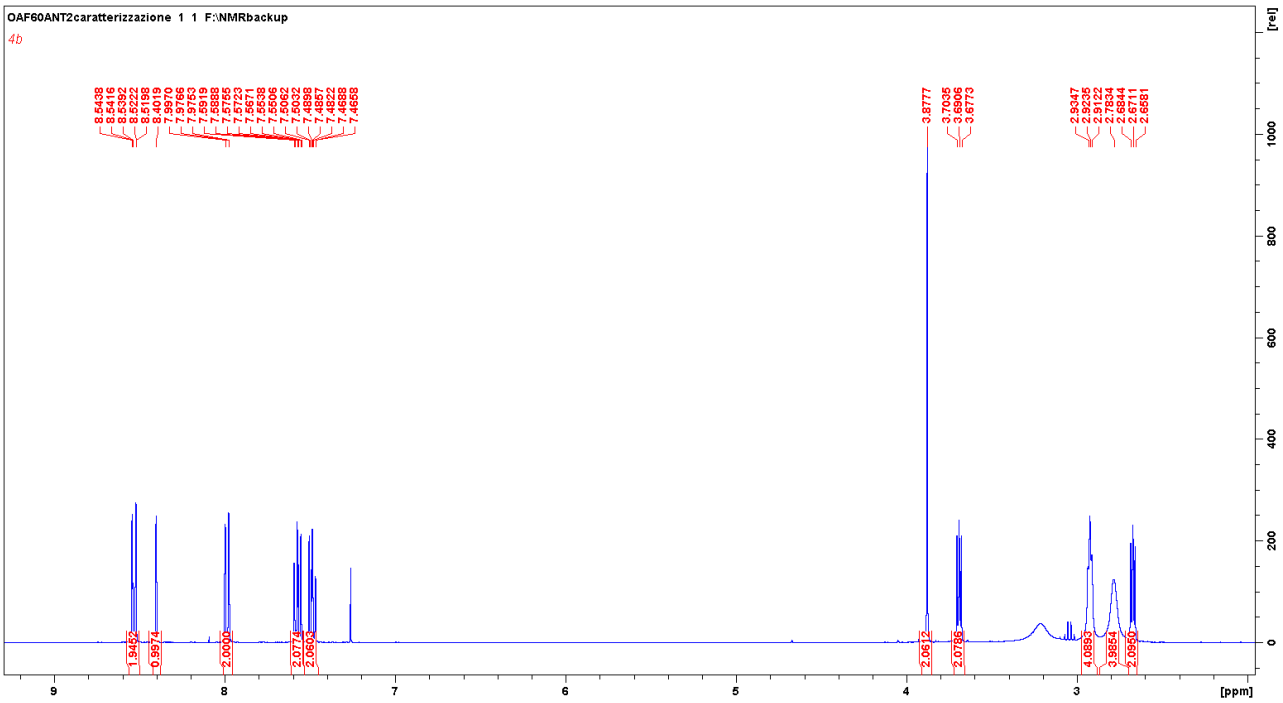


Figure S2.5. ^{1}H-NMR spectrum of compound **4b**.

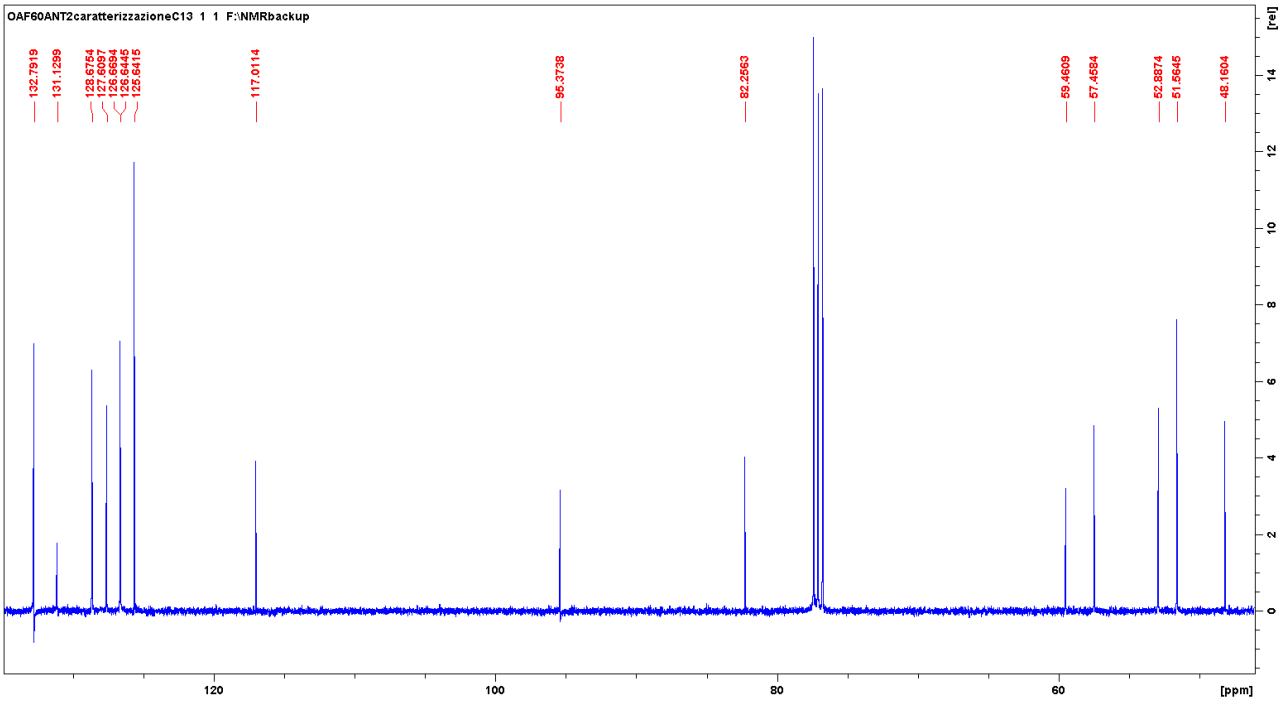


Figure S2.6. ^{13}C-NMR spectrum of compound **4b**.

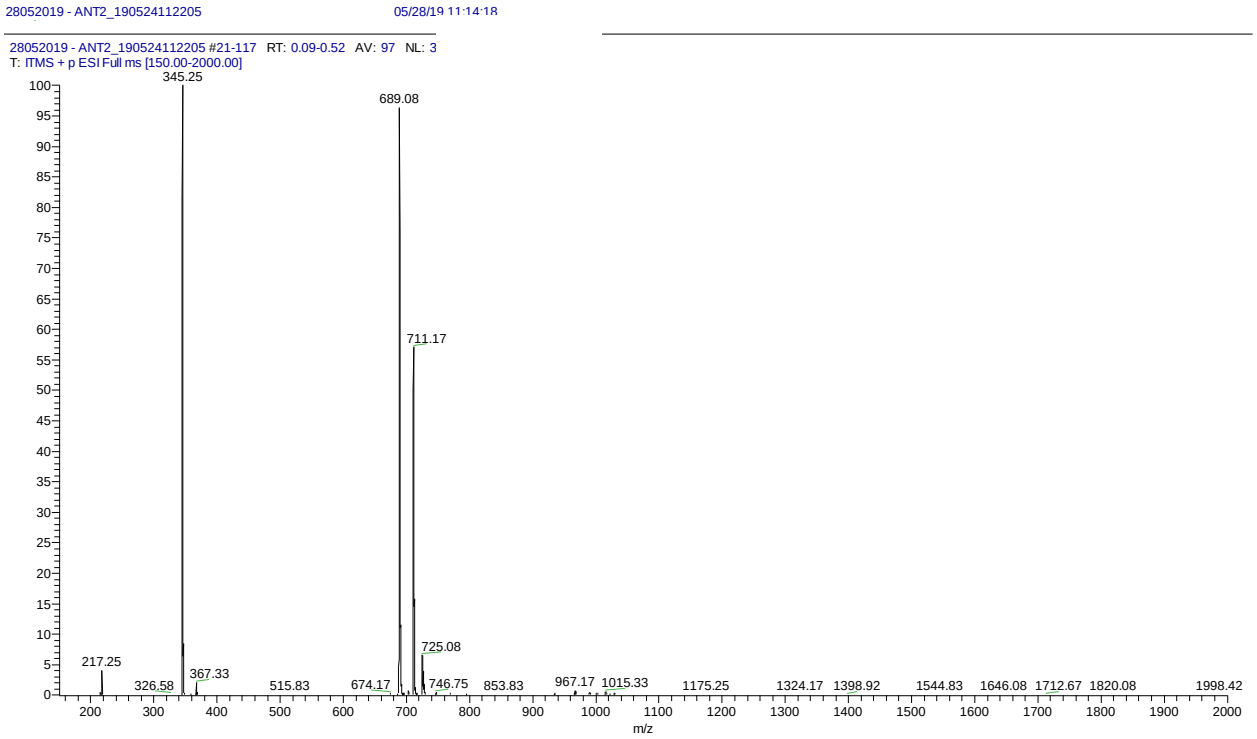


Figure S2.7. ESI-MS spectrum of compound **4b**. Molecular ions corresponding to a dimer and its sodium adduct can be observed.

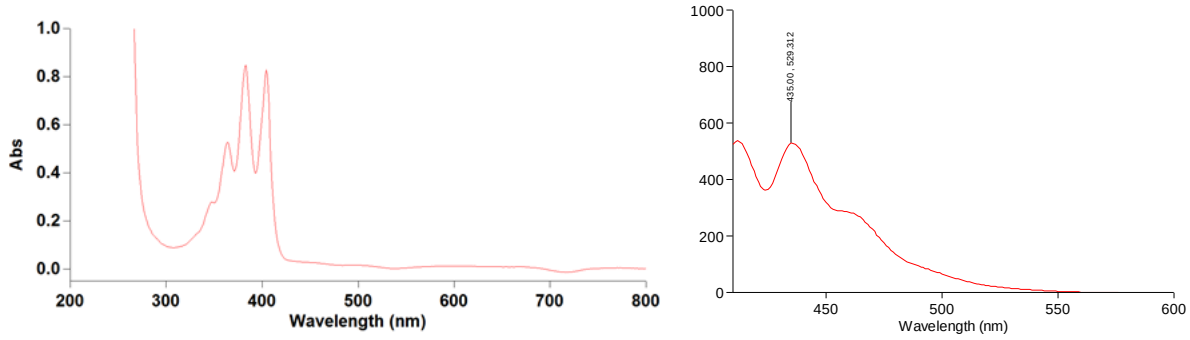


Figure S2.8. UV-Vis absorbance spectrum (left) and fluorescence emission spectrum (right).

Compound **4b** was investigated more in detail and its spectral behavior was studied at different concentrations acquiring multiple UV-Vis and fluorescence emission spectra. As can be noticed from the spectra reported below, no major spectral changes can be observed increasing compound concentration. This suggests that aggregation does not occur at the tested concentrations.

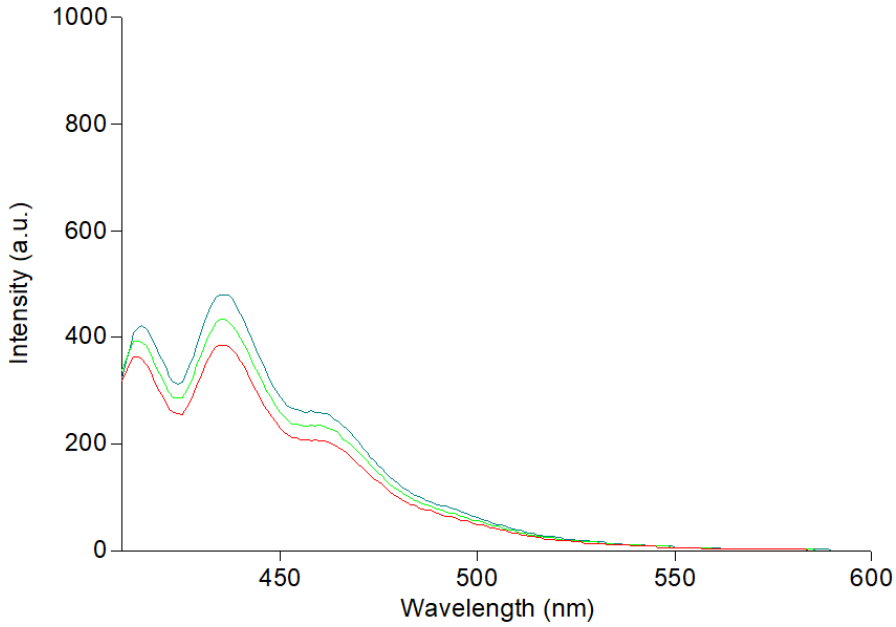


Figure S2.9. Fluorescence emission spectra at different concentrations (50-100 µM), slit: 2.5 nm.

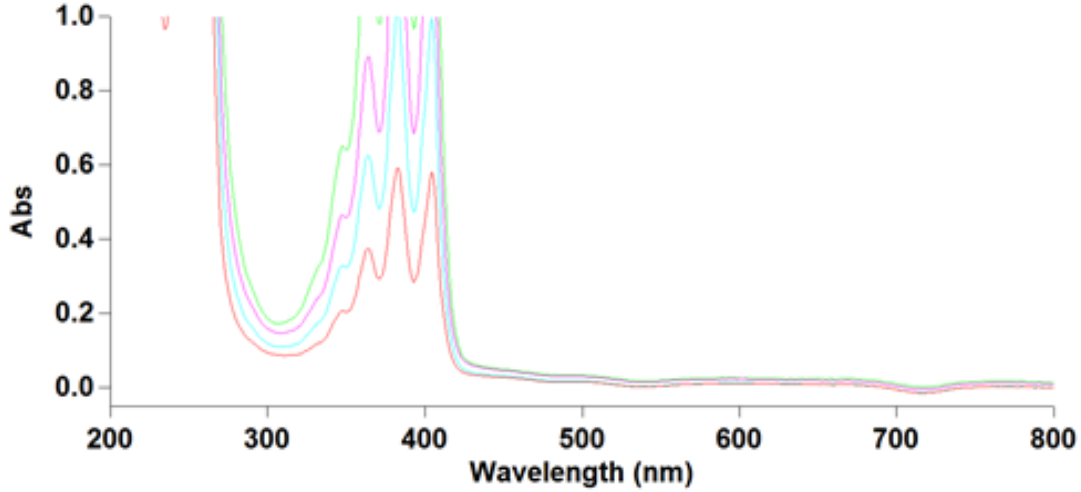


Figure S2.10. UV absorbance spectra at different concentrations (50-200 µM).

*Characterization of compound **4c***

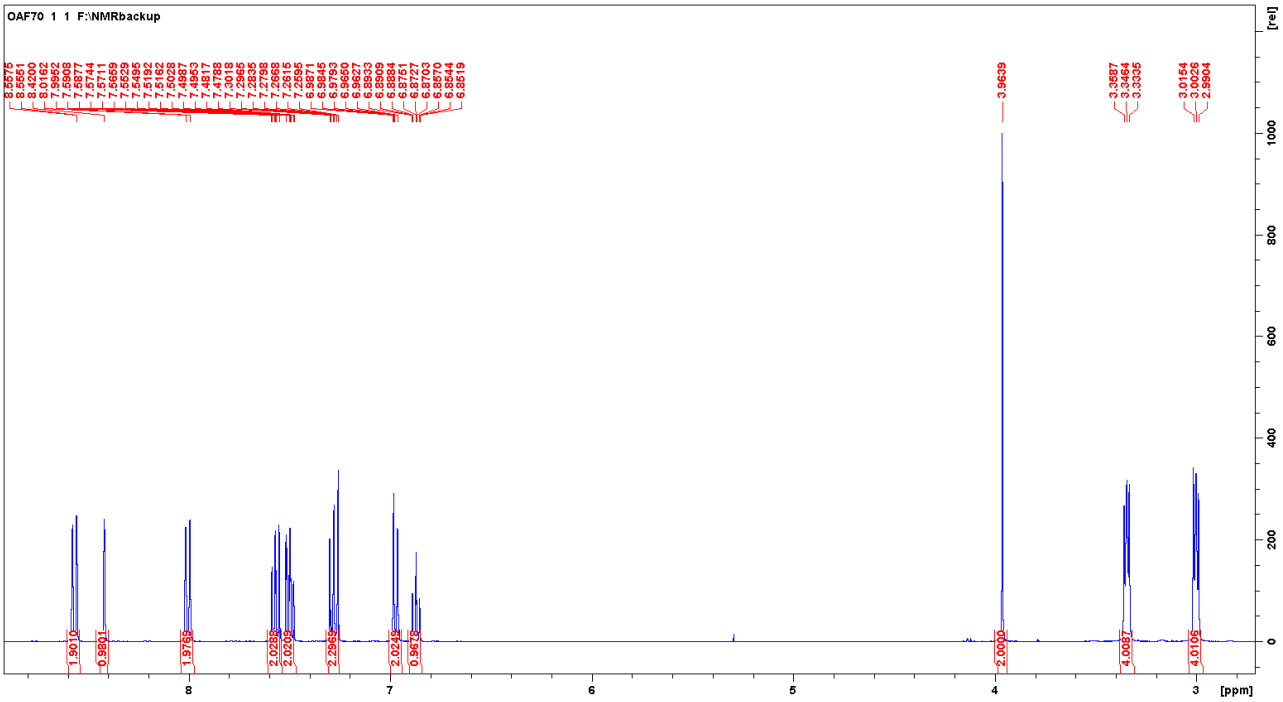


Figure S2.11. ^{1}H-NMR spectrum of compound **4c**.

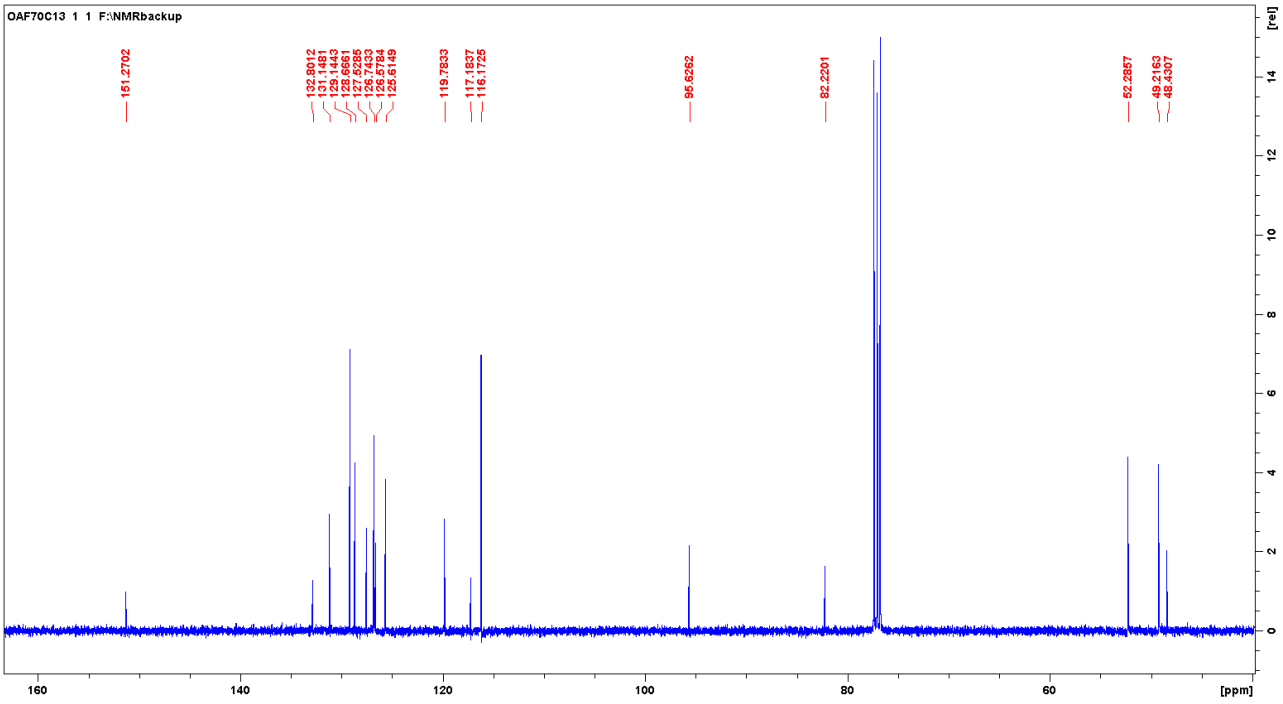


Figure S2.12. ^{13}C-NMR spectrum of compound **4c**.

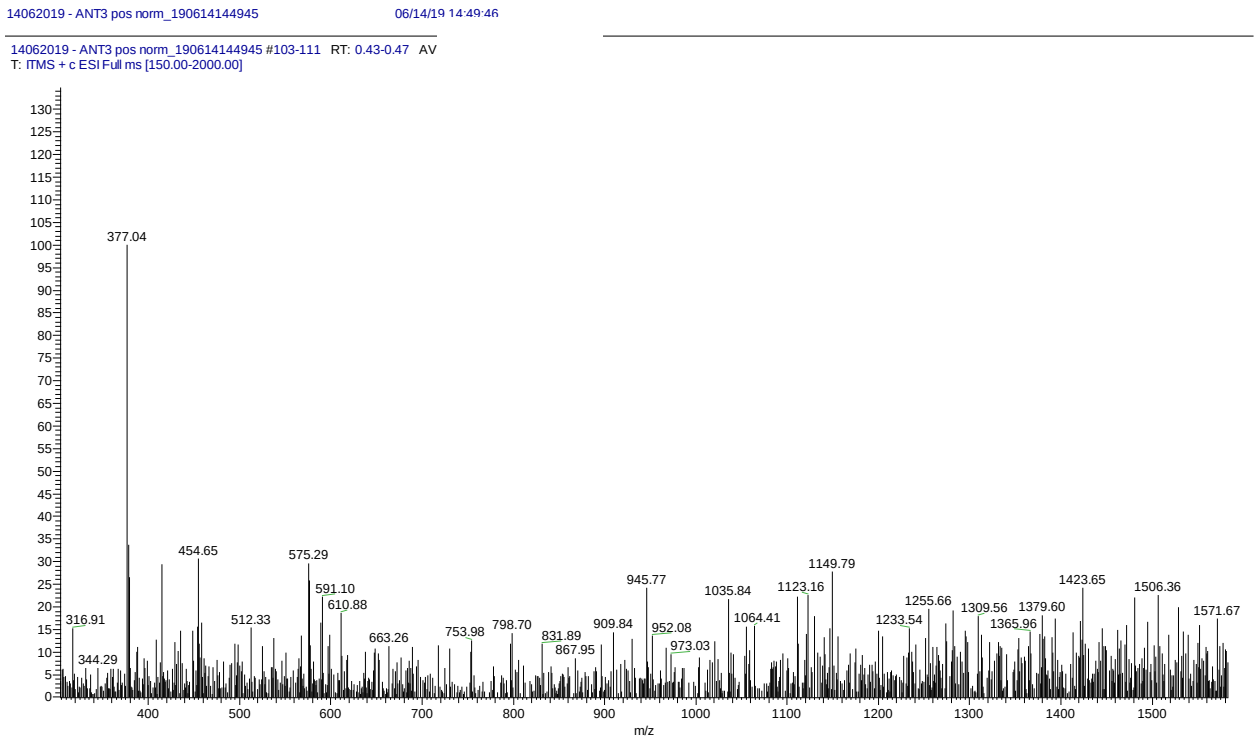


Figure S2.13. ESI-MS spectrum of compound **4c**.

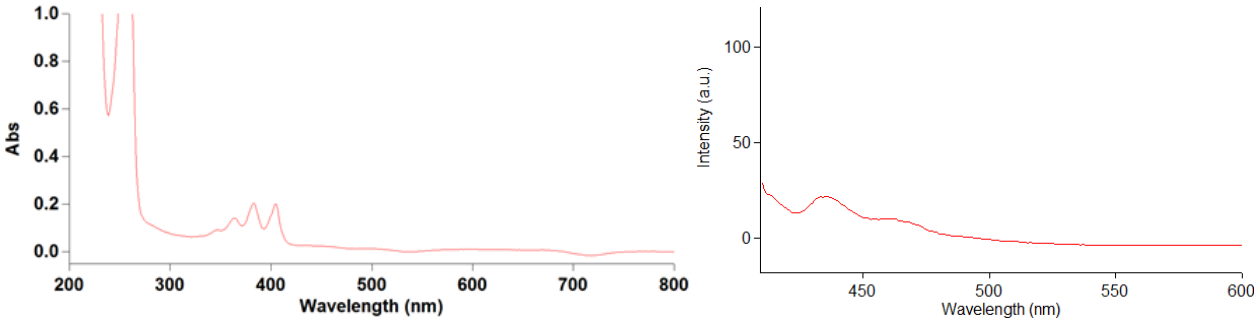


Figure S2.14. UV-Vis absorbance spectrum (left) and fluorescence emission spectrum (right).

*Characterization of compound **4d***

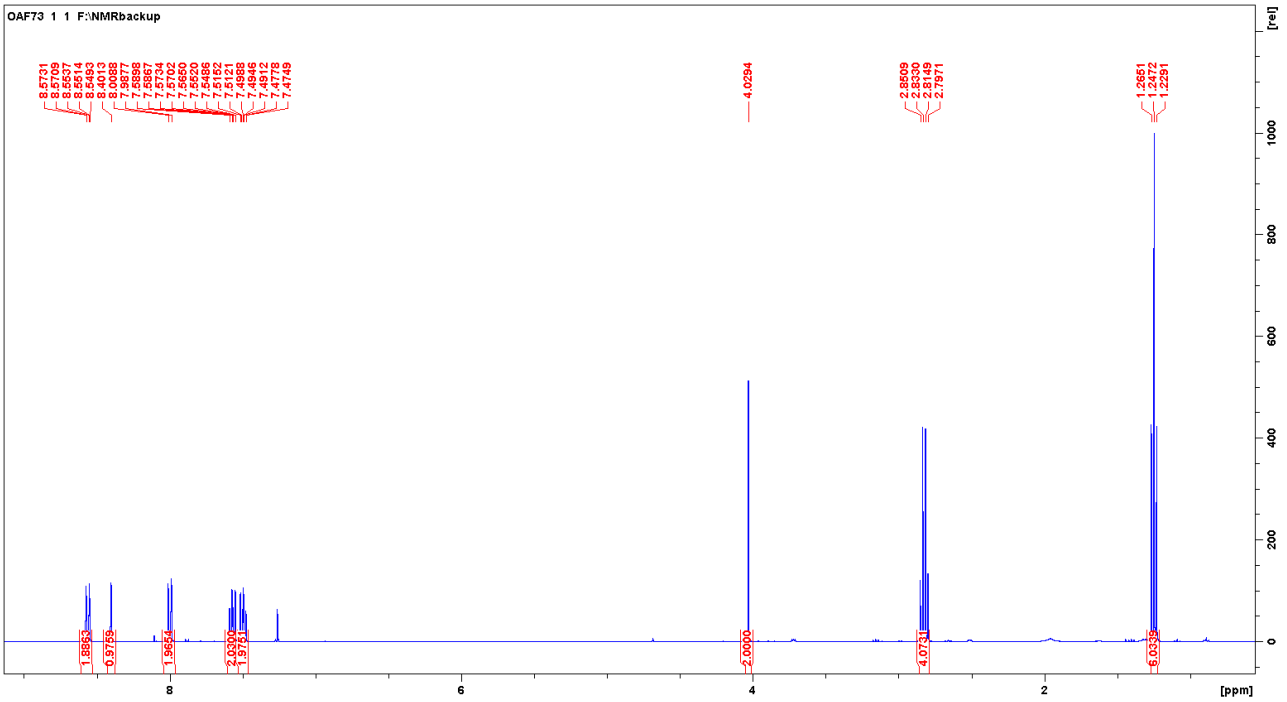


Figure S2.15. ^{1}H-NMR spectrum of compound **4d**.

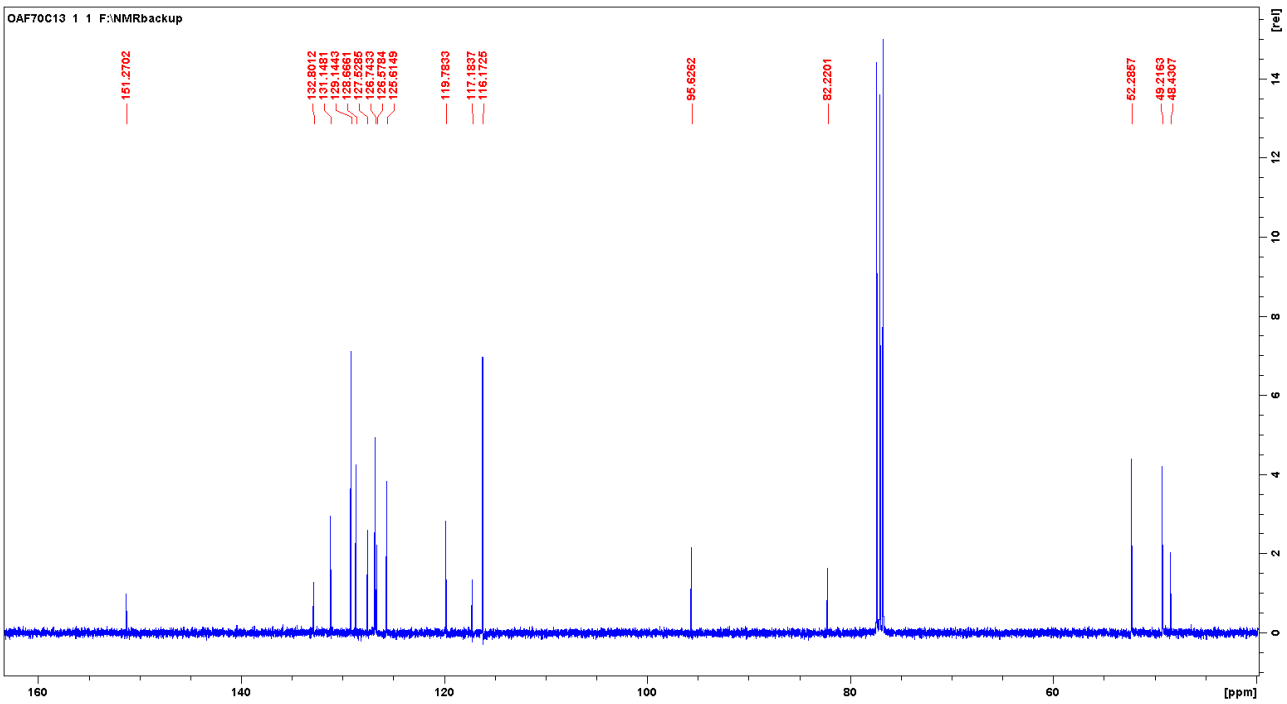


Figure S2.16. ^{13}C-NMR spectrum of compound **4d**.

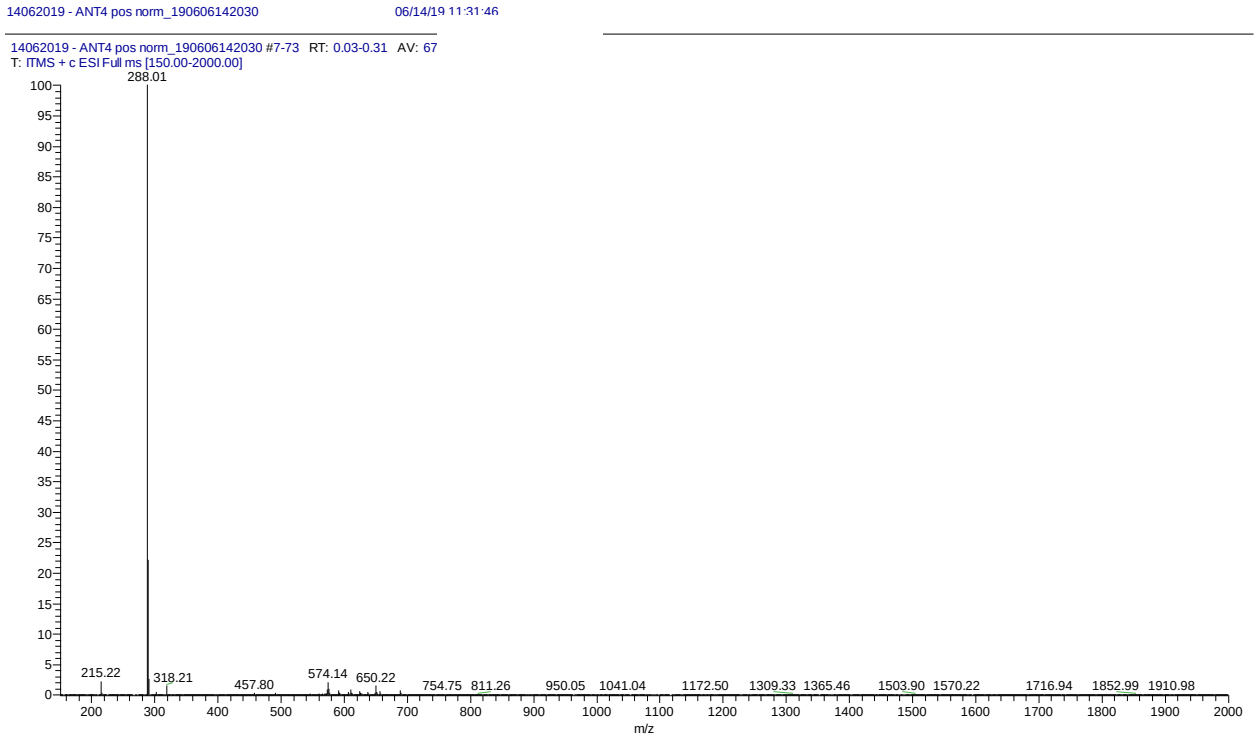


Figure S2.17. ESI-MS spectrum of compound **4d**.

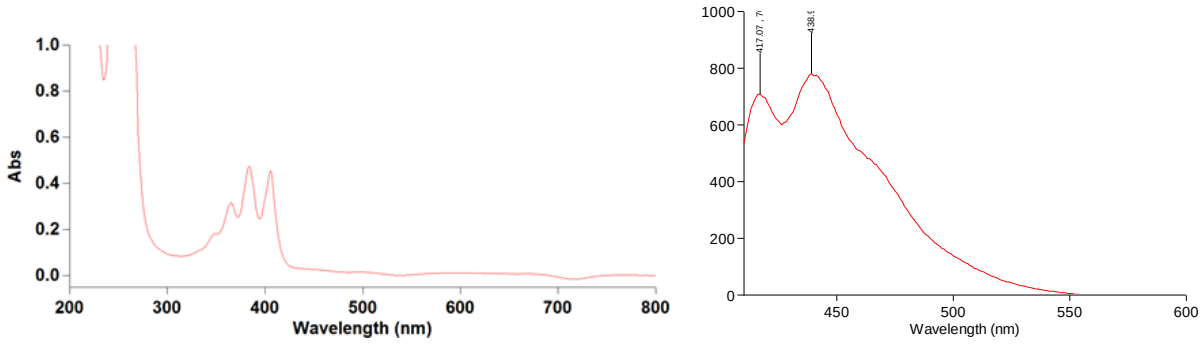


Figure S2.18. UV-Vis absorbance spectrum (left) and fluorescence emission spectrum (right).

*Characterization of compound **4e***

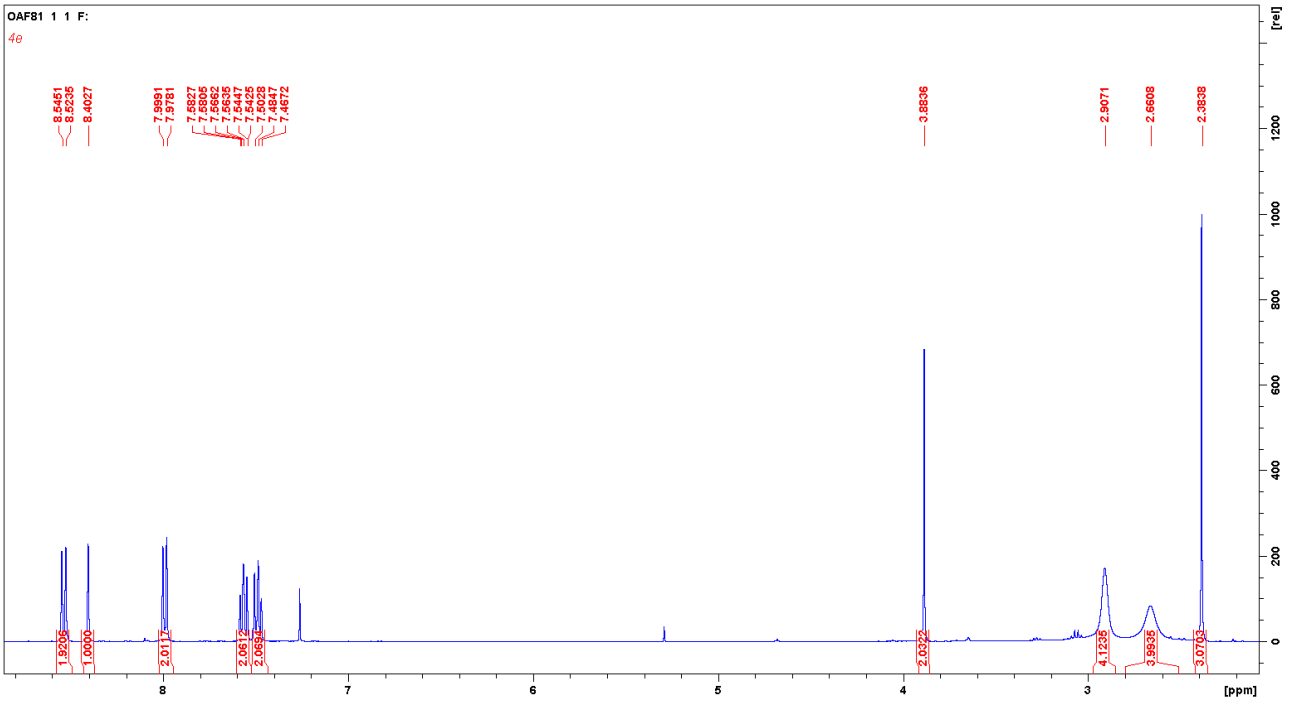


Figure S2.19. ^{1}H-NMR spectrum of compound **4e**.

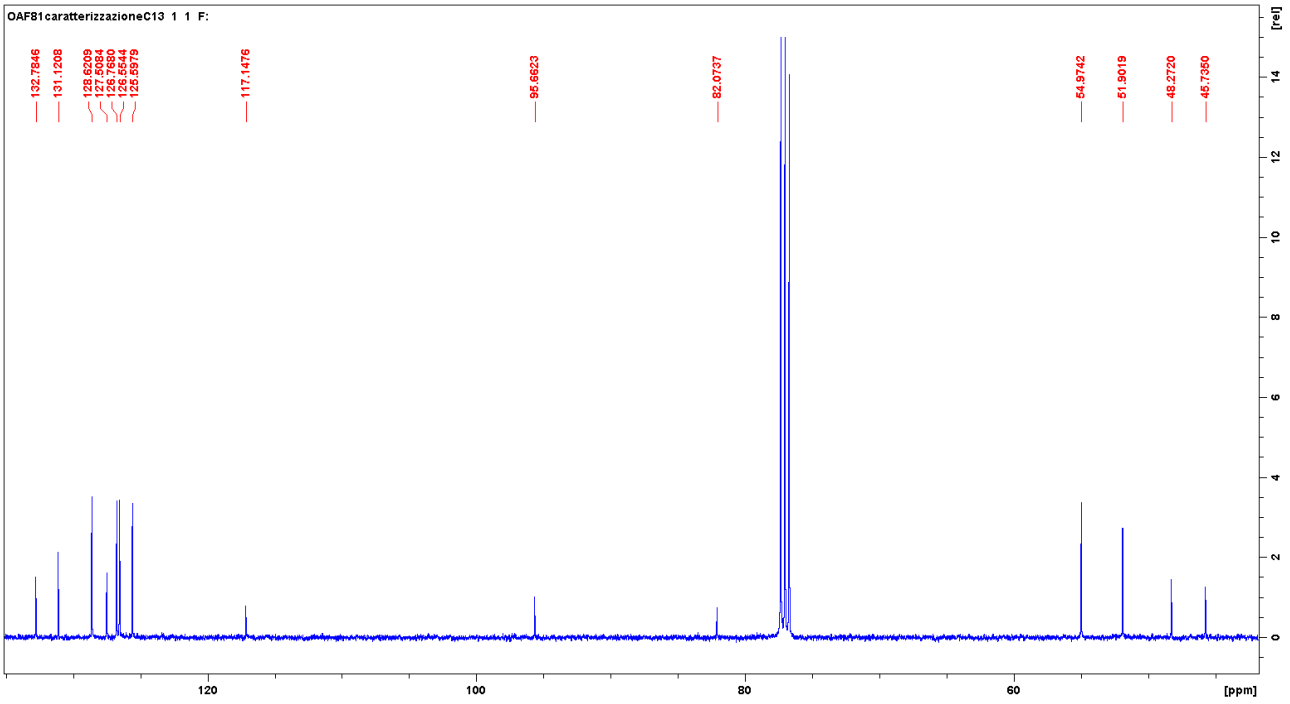


Figure S2.20. ^{13}C-NMR spectrum of compound **4e**.

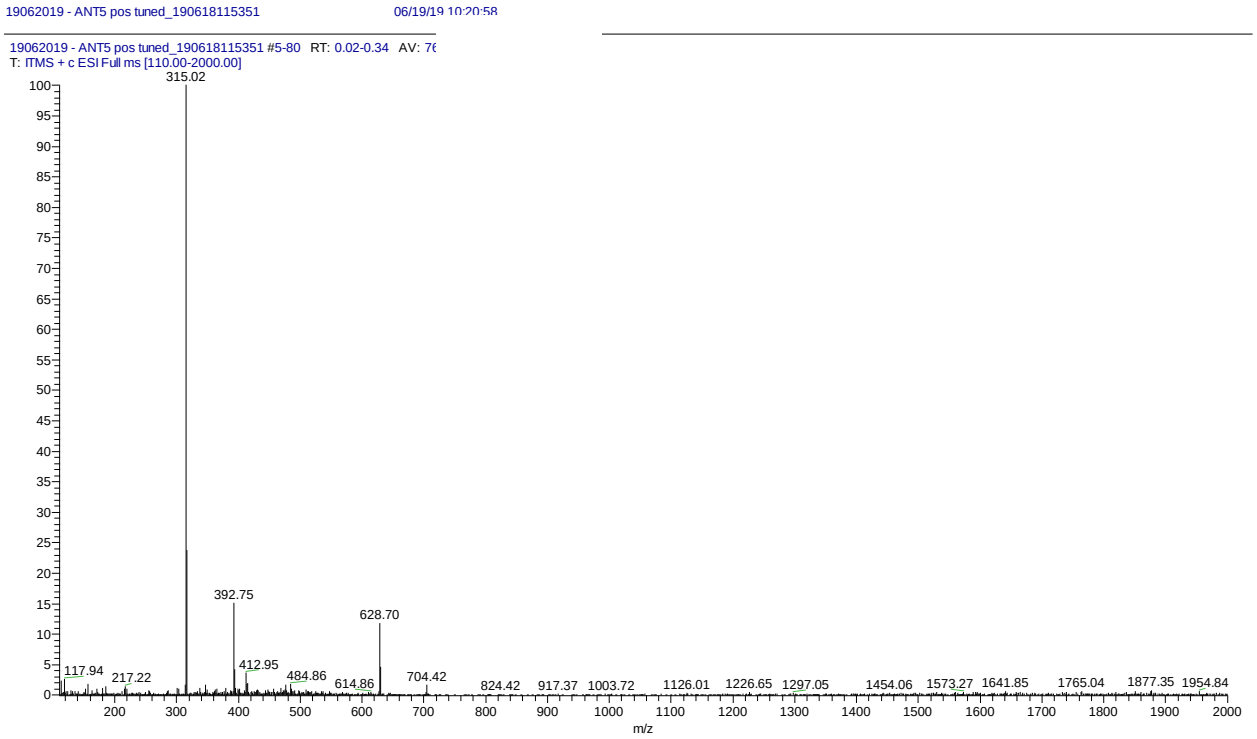


Figure S2.21. ESI-MS spectrum of compound **4e**. The molecular ion corresponding to a dimer can be observed.

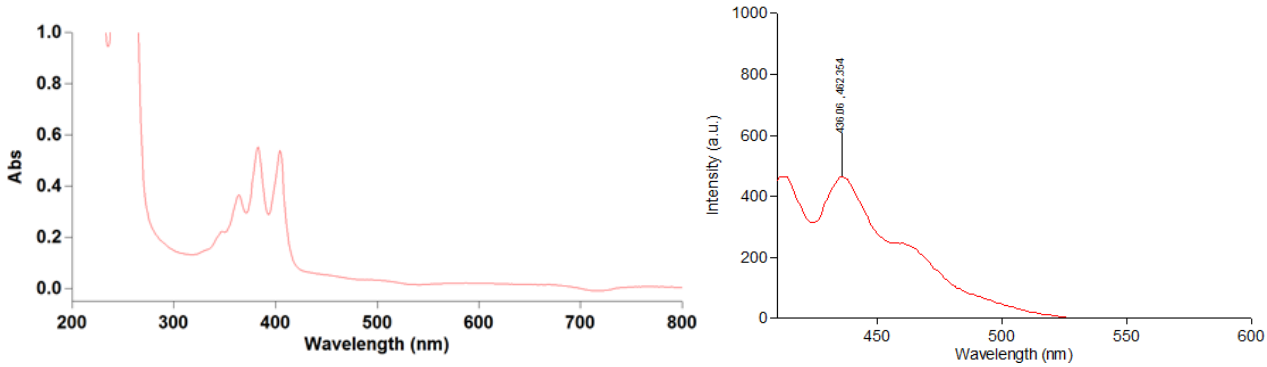


Figure S2.22. UV-Vis absorbance spectrum (left) and fluorescence emission spectrum (right).

*Characterization of compound **4f***

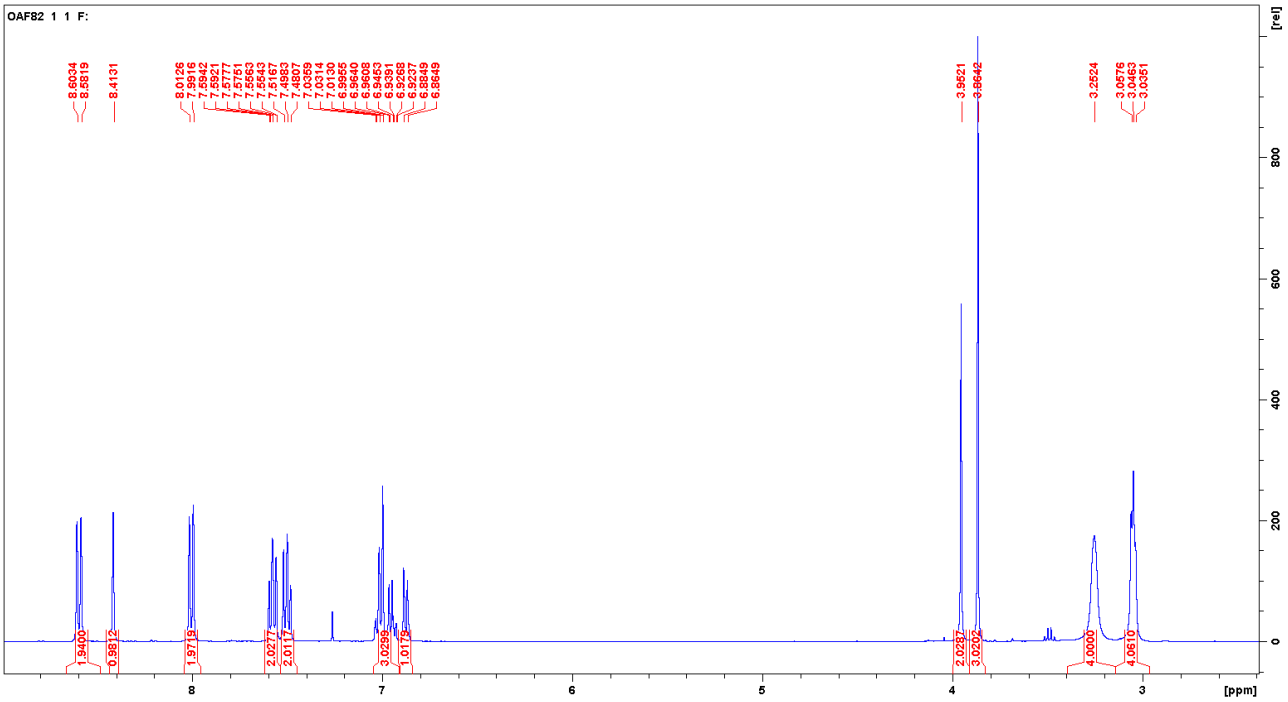


Figure S2.23. ^{1}H-NMR spectrum of compound **4f**.

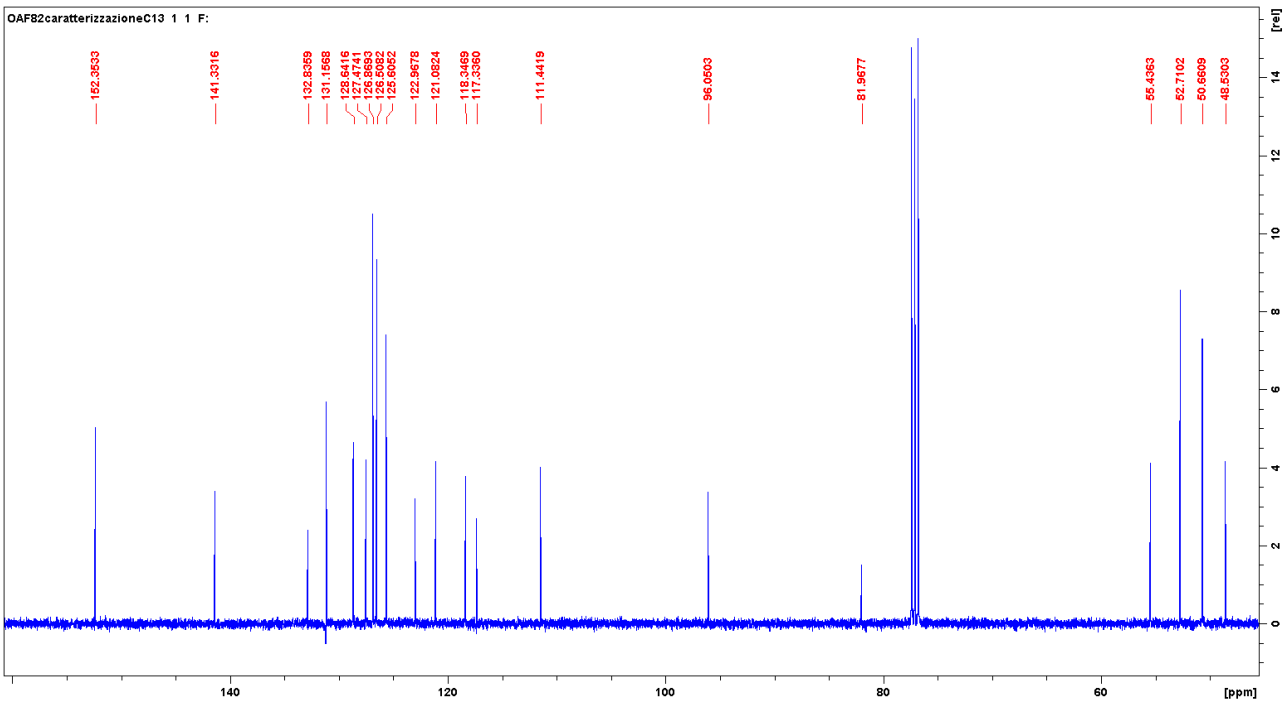


Figure S2.24. ^{13}C-NMR spectrum of compound **4f**.

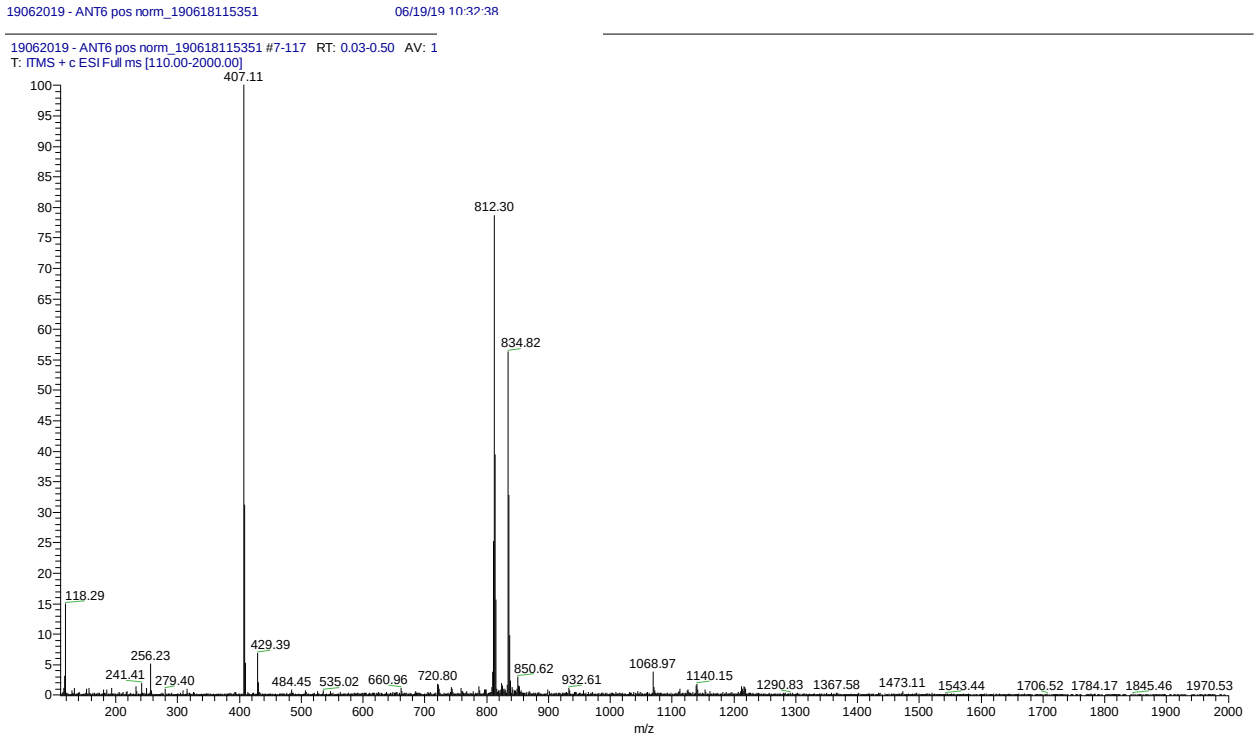


Figure S2.25. ESI-MS spectrum of compound **4f**. Molecular ions corresponding to a dimer and its sodium adduct can be observed.

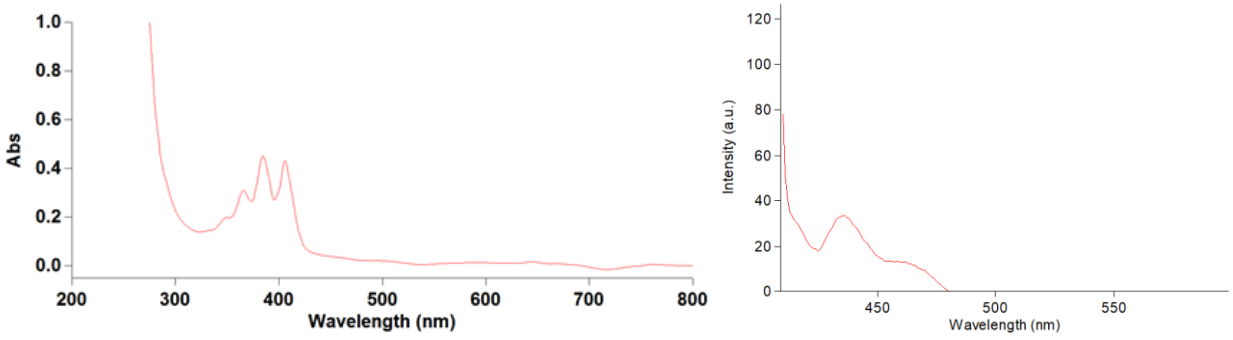


Figure S2.26. UV-Vis absorbance spectrum (left) and fluorescence emission spectrum (right).

Docking study

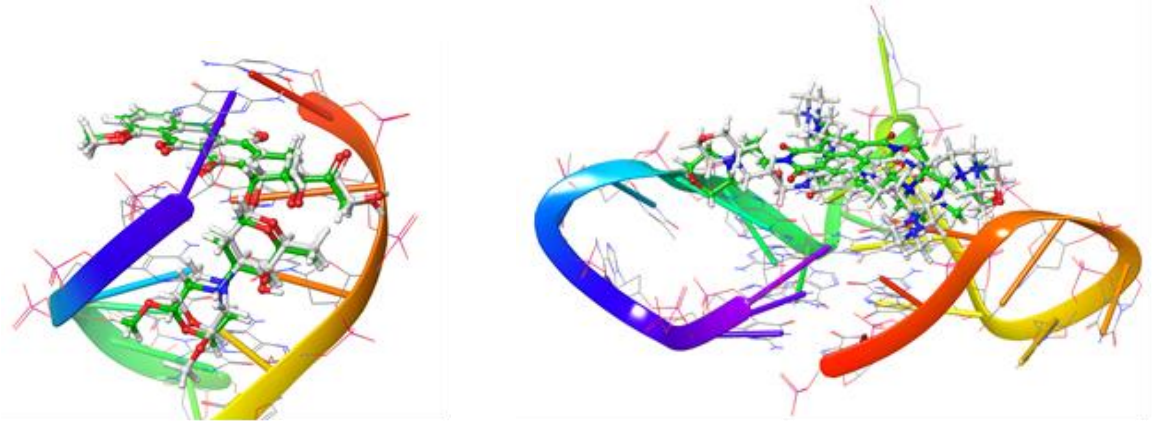


Figure S2.27. Superimpositions of the re-docked native ligands of the structures on their position in the crystal. Left: dsDNA (PDB ID: 2DES) with the superimposition of the ligand morpholino doxorubicin. Right: G4 DNA (PDB ID: 3UYH) with the superimposition of the ligand naphthalenediimide.

ESI-MS binding study

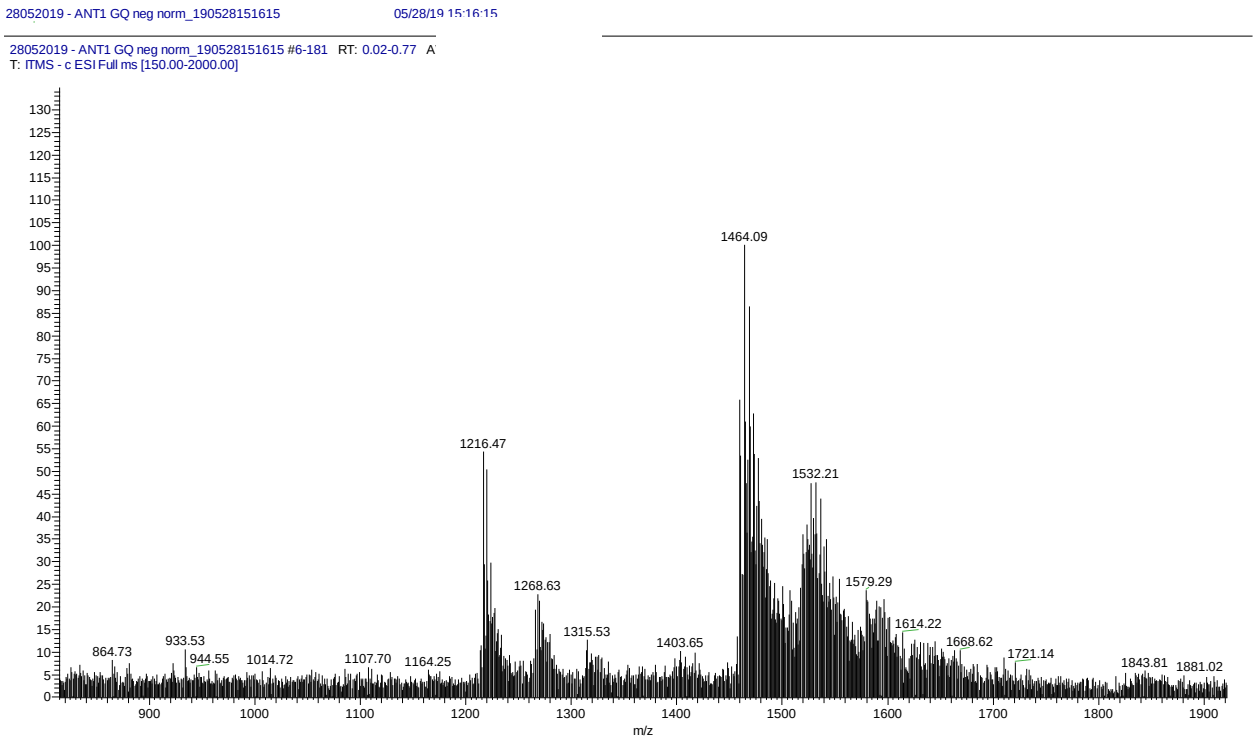


Figure S2.28. ESI-MS experiment showing the interaction of compound **4a** with G4 DNA (z = -5 and z= -6).

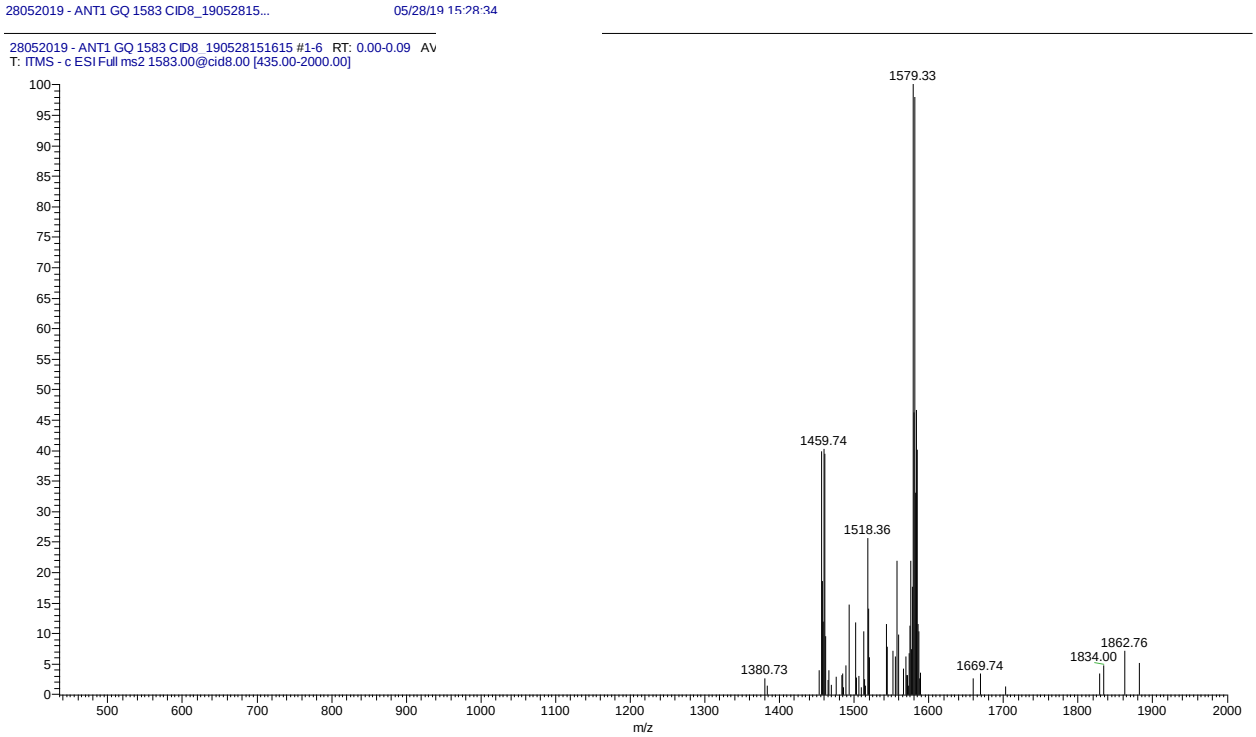


Figure S2.29. CID fragmentation spectrum of the 2:1 complex of compound **4a** with G4 DNA (z = -5, 33 eV).

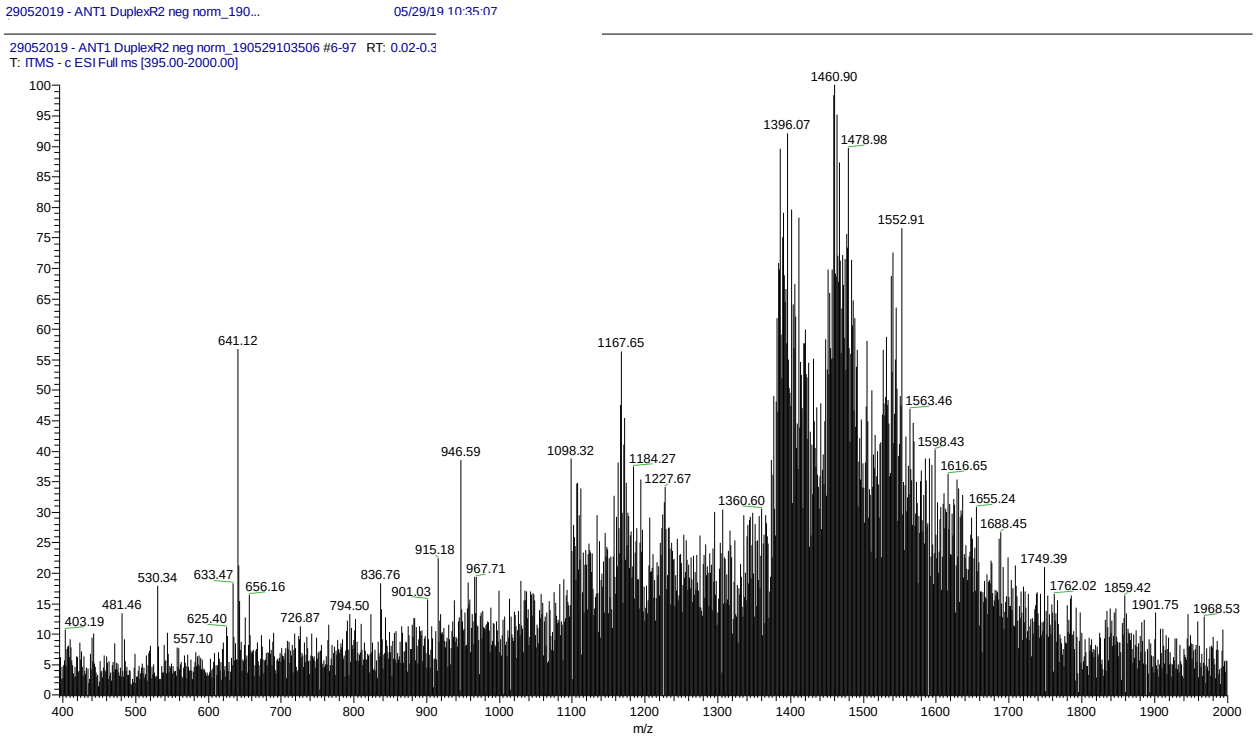


Figure S2.30. ESI-MS experiment showing the interaction of compound **4a** with dsDNA (z = -8).

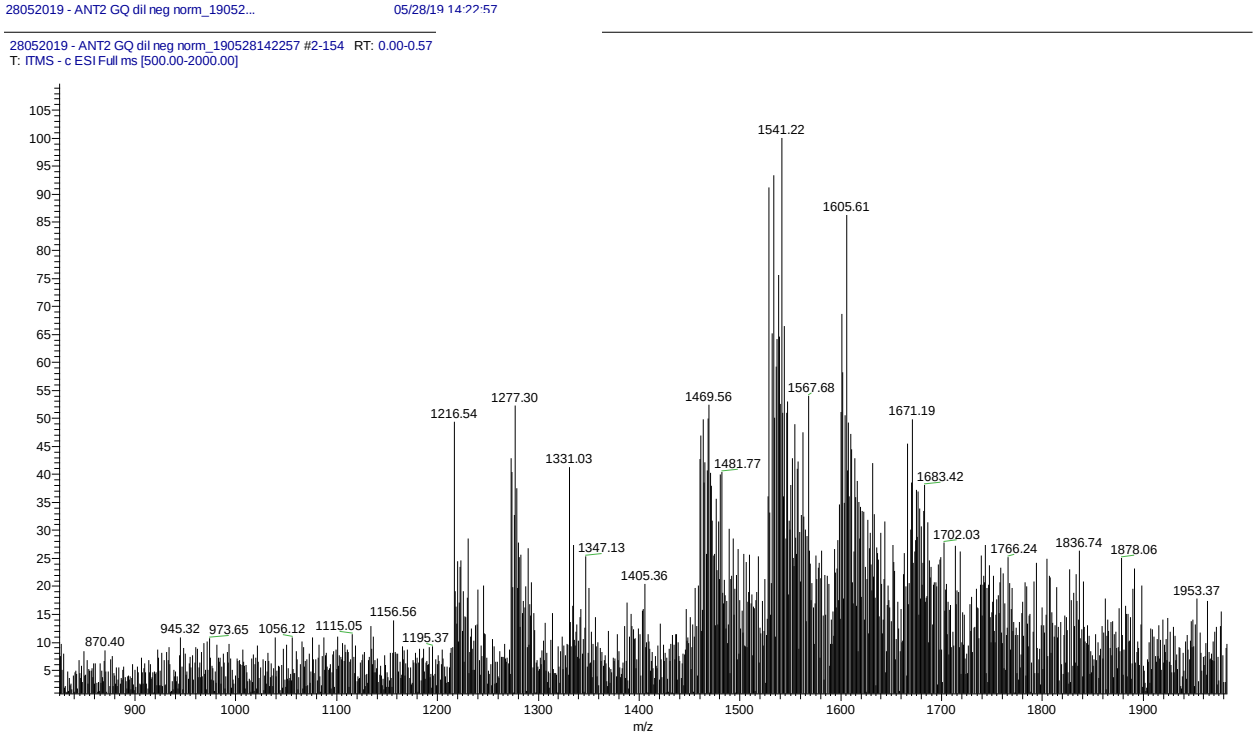


Figure S2.31. ESI-MS experiment showing the interaction of compound **4b** with G4 DNA (z = -5 and z= -6).

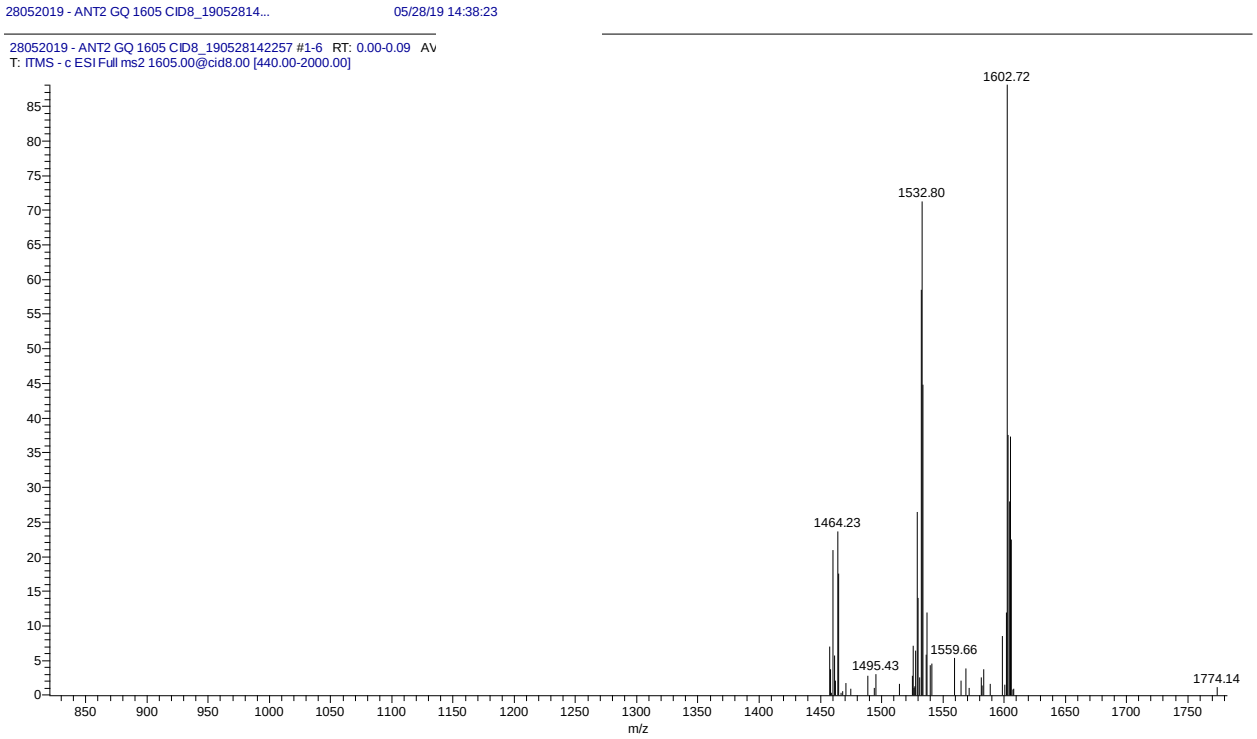


Figure S2.32. CID fragmentation spectrum of the 2:1 complex of compound **4b** with G4 DNA (z = -5, 34 eV).

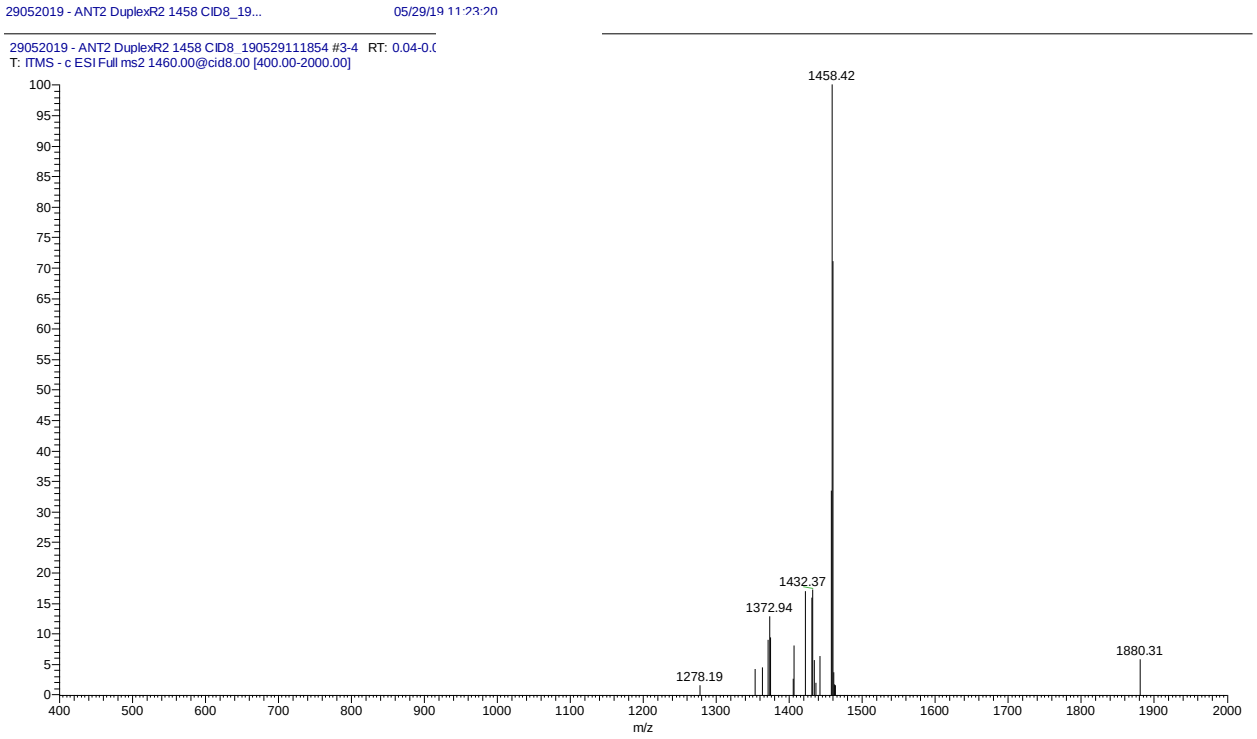


Figure S2.33. CID fragmentation spectrum of the 2:1 complex of compound **4b** with dsDNA (z = -8, 33 eV).

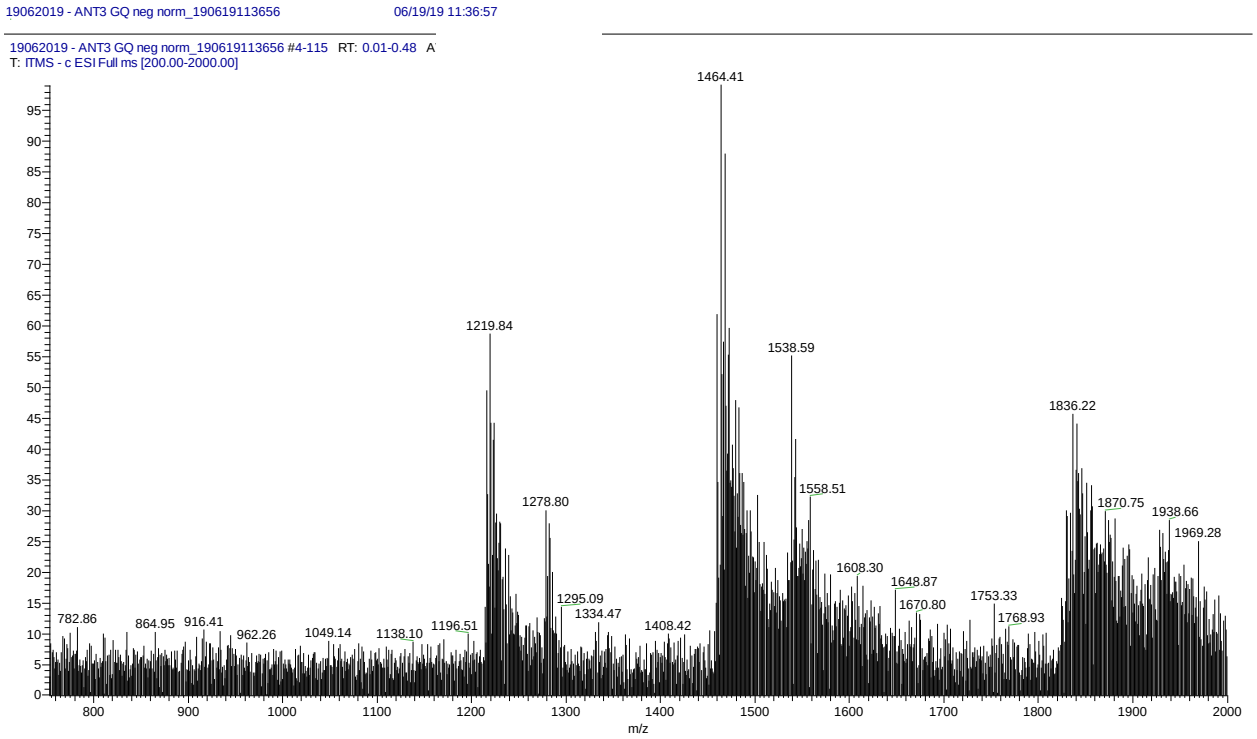


Figure S2.34. ESI-MS experiment showing the interaction of compound **4c** with G4 DNA (z = -4, z = -5 and z = -6).

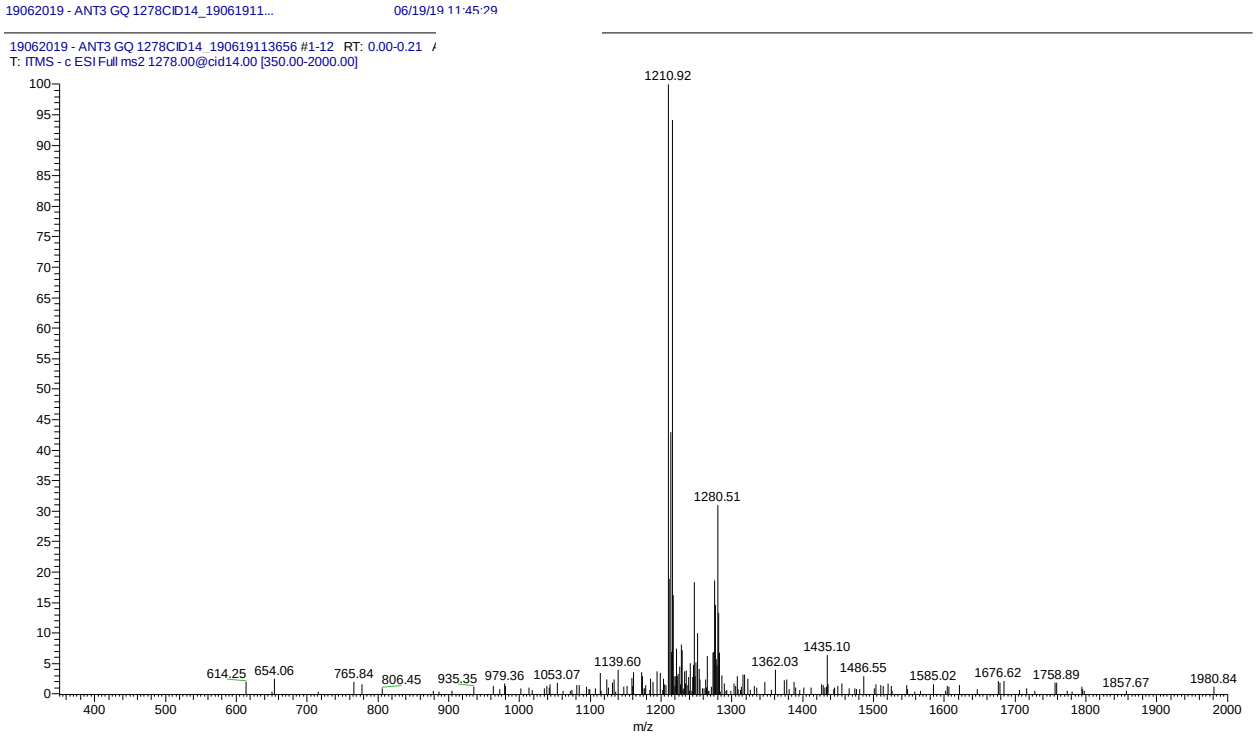


Figure S2.35. CID fragmentation spectrum of the 1:1 complex of compound **4c** with G4 DNA (z = -6, 48 eV).

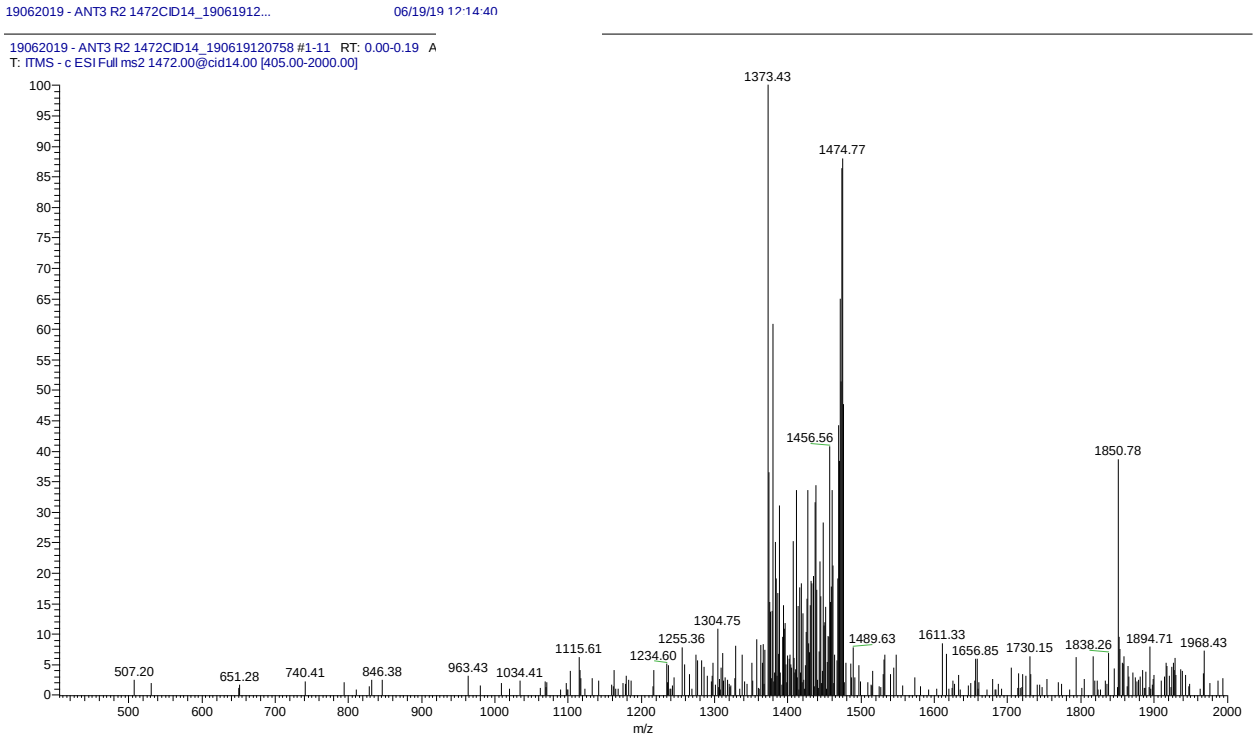


Figure S2.36. CID fragmentation spectrum of the 1:1 complex of compound **4c** with dsDNA (z = -8, 55 eV).

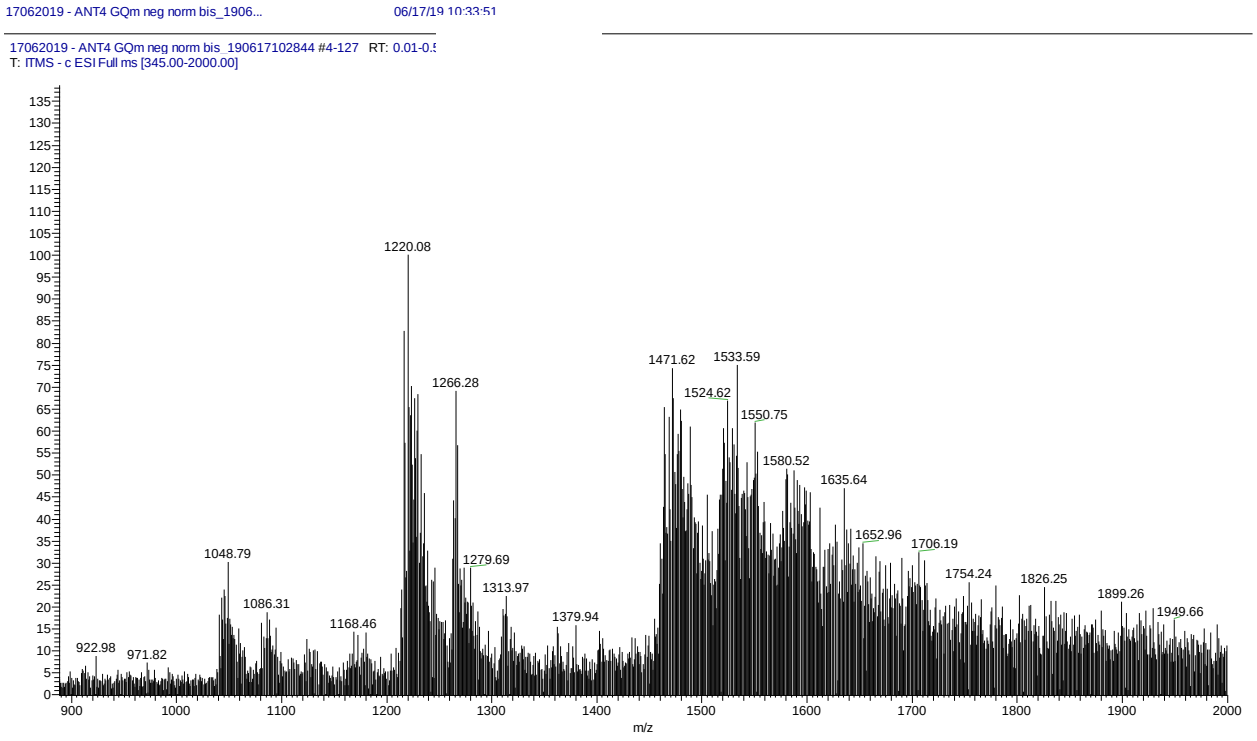


Figure S2.37. ESI-MS experiment showing the interaction of compound **4d** with G4 DNA (z = -5, z = -6 and z = -7).

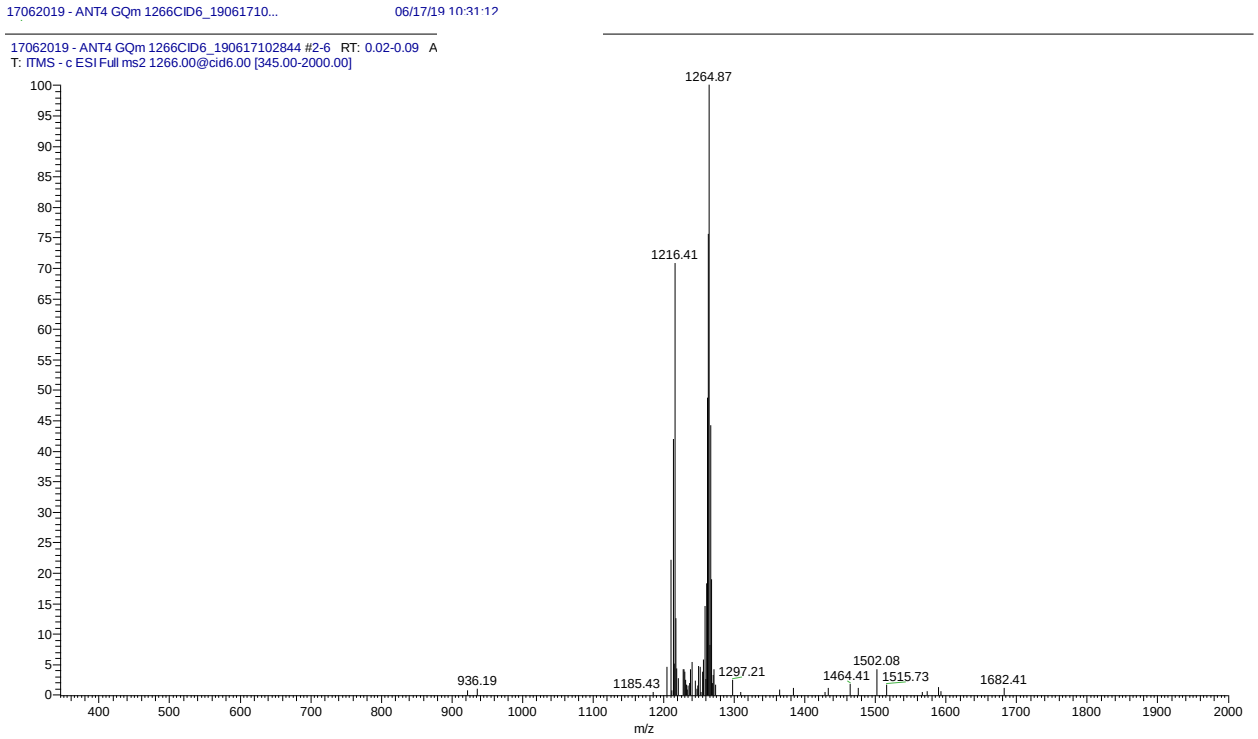


Figure S2.38. CID fragmentation spectrum of the 1:1 complex of compound **4d** with G4 DNA (z = -6, 21 eV).

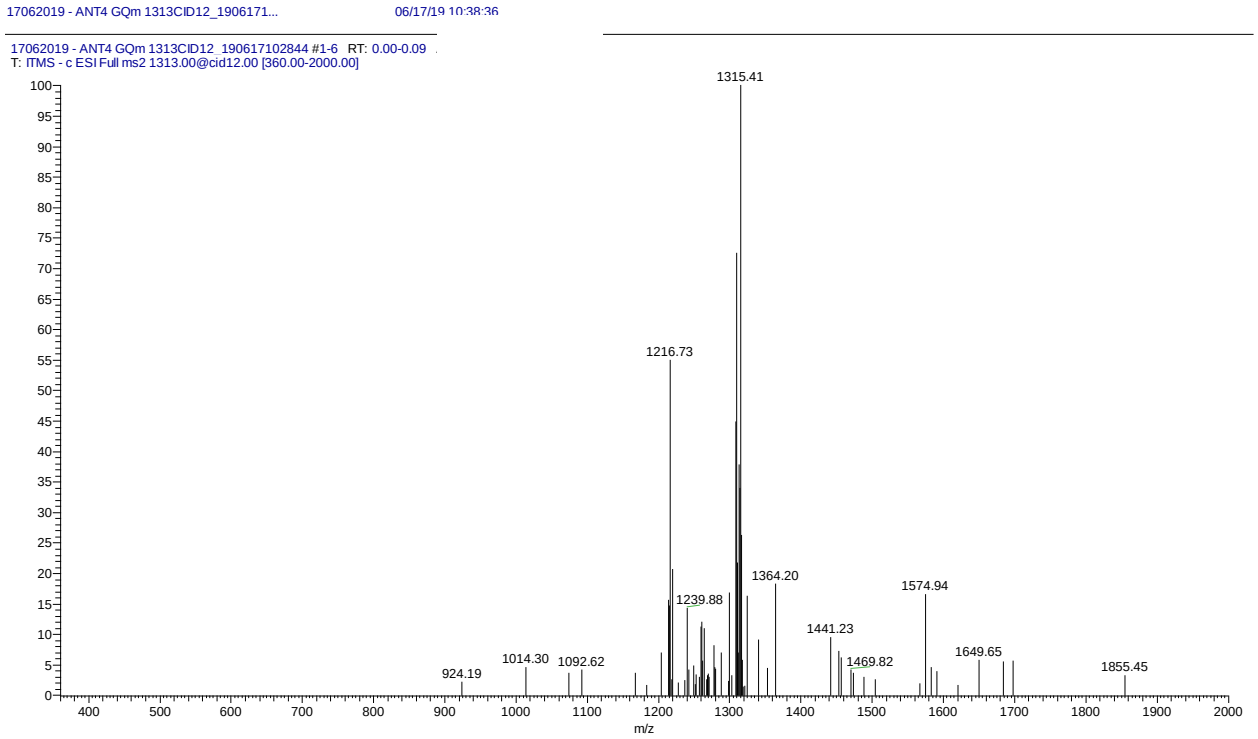


Figure S2.39. CID fragmentation spectrum of the 2:1 complex of compound **4d** with G4 DNA (z = -6, 42 eV).

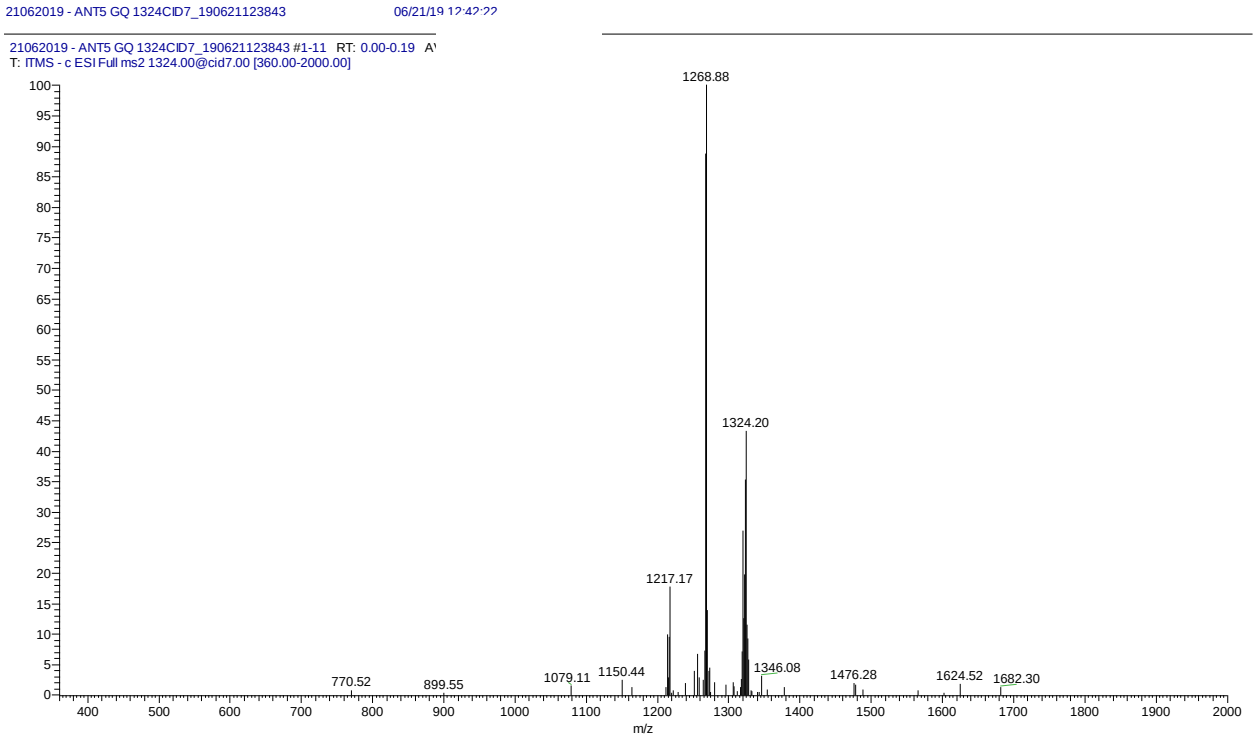


Figure S2.40. CID fragmentation spectrum of the 2:1 complex of compound **4e** with G4 DNA (z = -6, 25 eV).

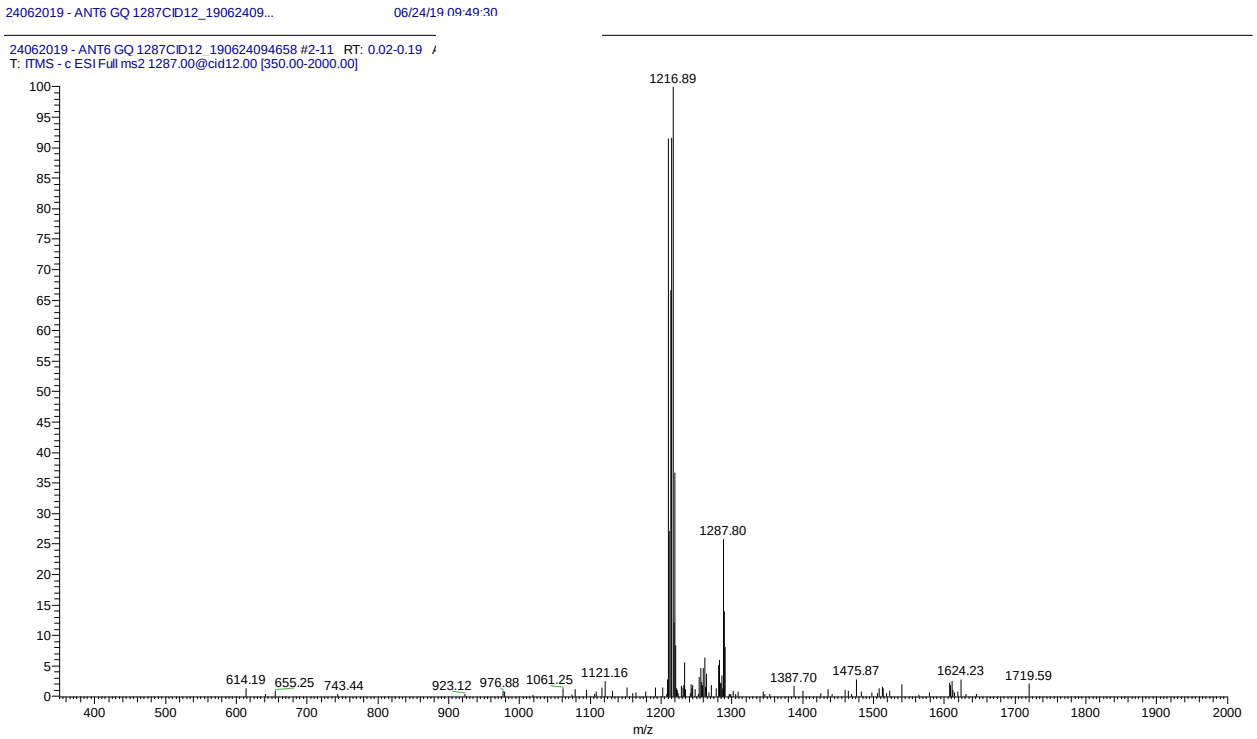


Figure S2.41. CID fragmentation spectrum of the 1:1 complex of compound **4f** with G4 DNA (z = -6, 41 eV).

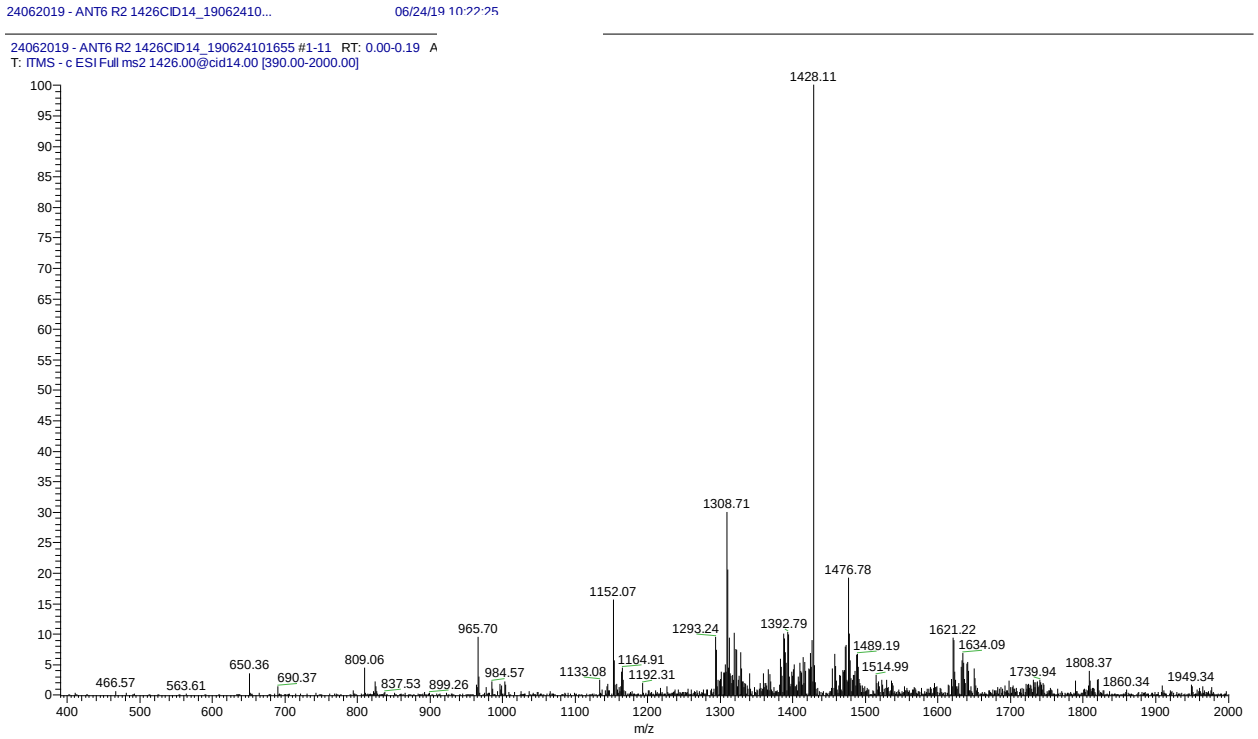


Figure S2.42. CID fragmentation spectrum of the 2:1 complex of compound **4f** with dsDNA (z = -8, 53 eV).

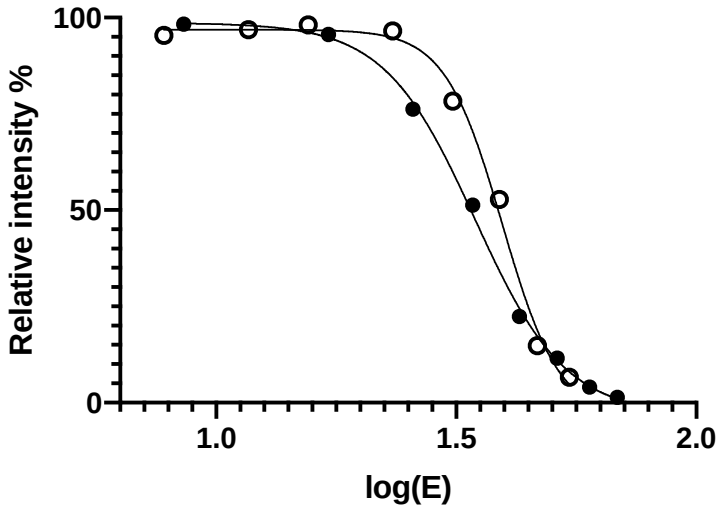


Figure S2.43. CID curve representing the fragmentation of the 2:1 compound-G4 complex (black dots) and 2:1 compound-dsDNA complex (white dots) for **4b**. Energy is expressed in eV.

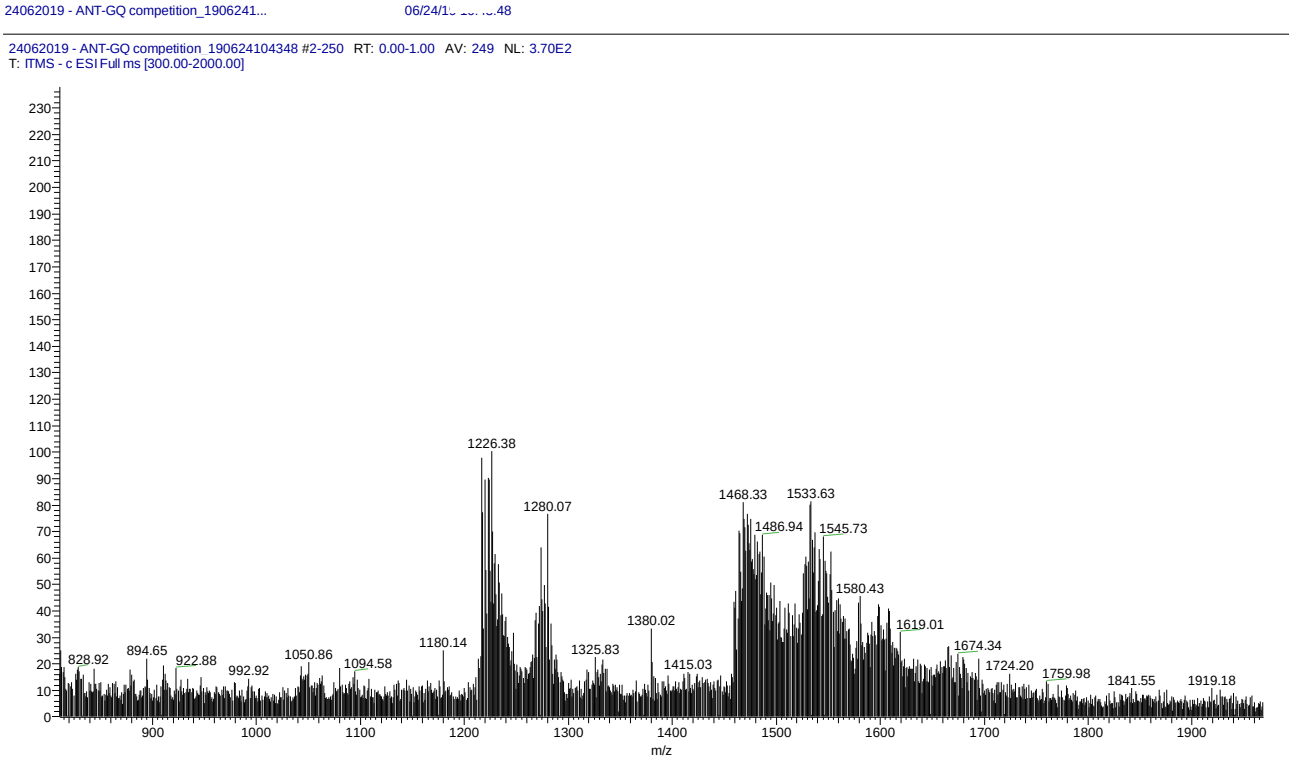


Figure S2.44. Competition ESI-MS binding experiment of **4b**, **4d** and **4a** towards G4 DNA (z = -5 and z = -6). The interaction peaks measured in the **4b**-G4 ESI-MS binding experiment were m/z = 1277, m/z = 1331 (z= -6, 1:1 and 2:1 complex, respectively) and m/z = 1541, m/z = 1605 (z= -5, 1:1 and 2:1 complex, respectively). The interaction peaks measured in the **4d**-G4 ESI-MS binding experiment were m/z = 1266, m/z = 1313 (z= -6, 1:1 and 2:1 complex, respectively) and m/z = 1524, m/z = 1580 (z= -5, 1:1 and 2:1 complex, respectively). The interaction peaks measured in the **4a**-G4 ESI-MS binding experiment were m/z = 1268, m/z = 1315 (z= -6, 1:1 and 2:1 complex, respectively) and m/z = 1532, m/z = 1579 (z= -5, 1:1 and 2:1 complex, respectively). By comparison with these results, and considering also 2:1 complexes, the here observed interaction peaks appear to resemble closely the pattern of compound **4b**-G4 ESI-MS binding study.

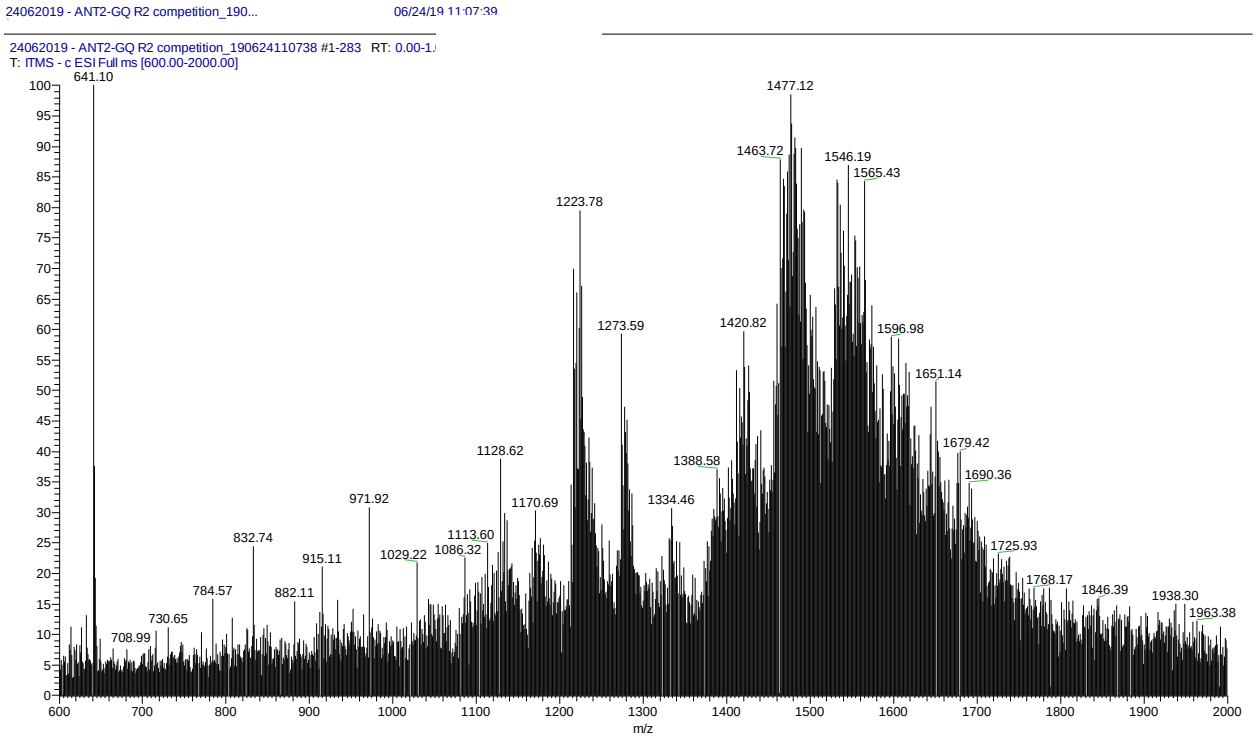


Figure S2.45. Competition ESI-MS binding experiment of **4b** towards G4 DNA (z = -6) and dsDNA (z = -8).

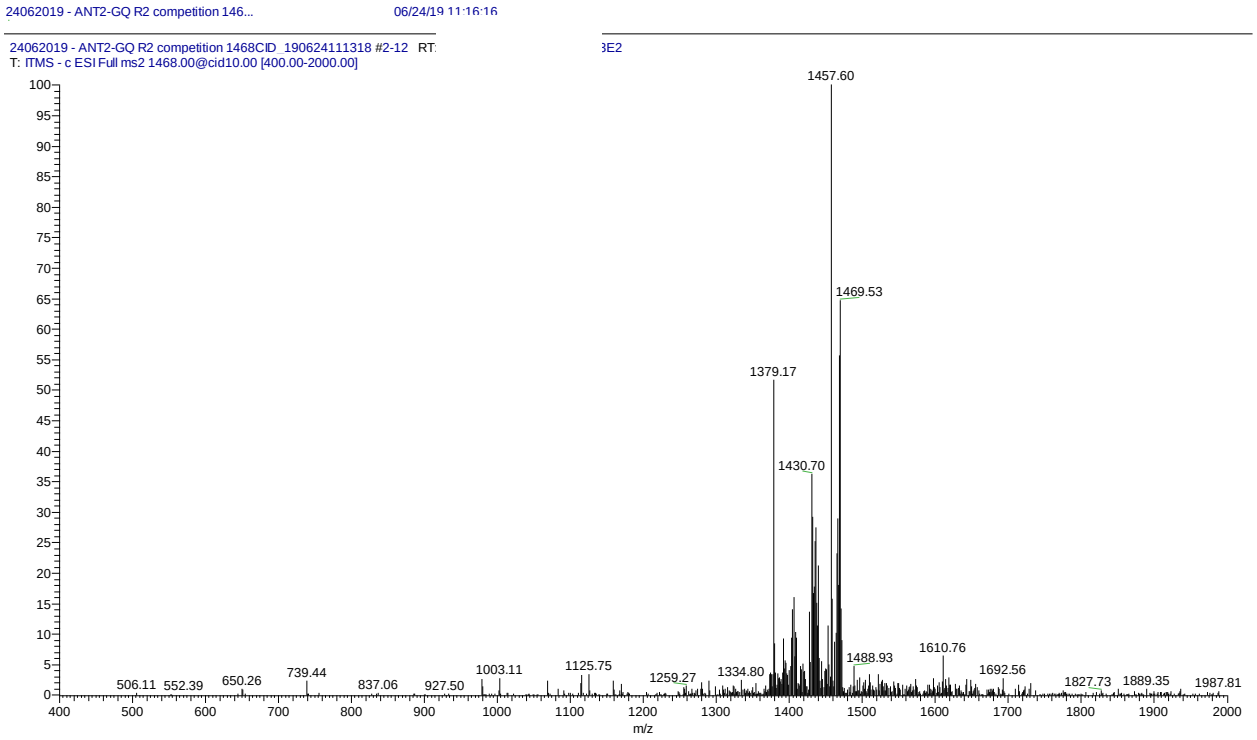


Figure S2.46. Competition ESI-MS binding experiment: CID fragmentation of 2:1 compound **4b**-dsDNA complex (z = -8, 24 eV).

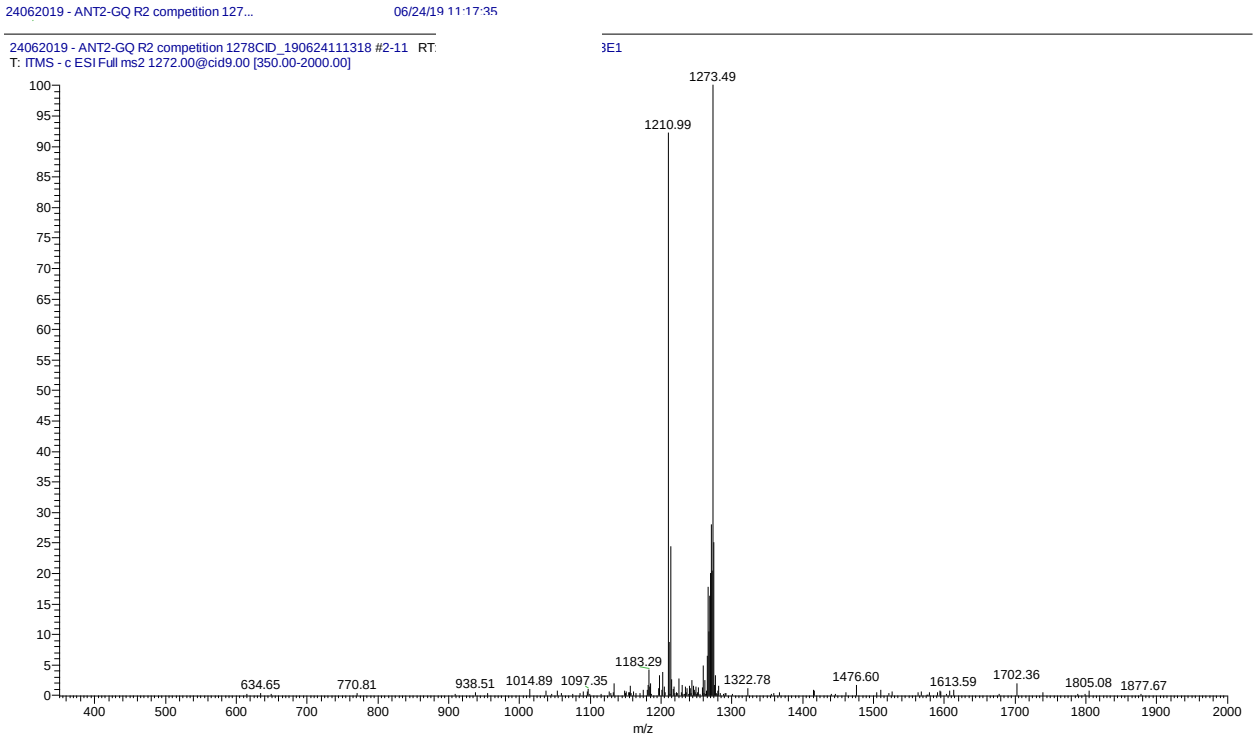


Figure S2.47. Competition ESI-MS binding experiment: CID fragmentation of compound **4b**-G4 DNA complex (z = -6, 28 eV).

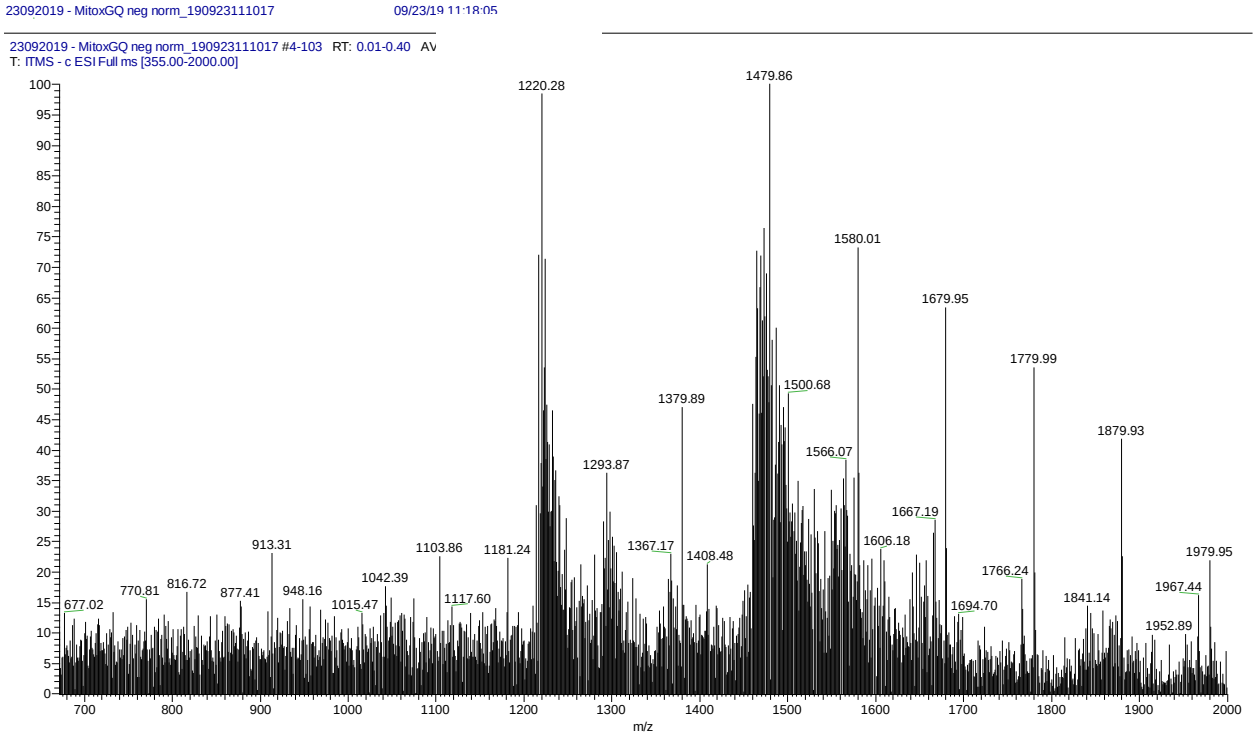


Figure S2.48. ESI-MS experiment showing the interaction of mitoxantrone with G4 DNA (z = -5 and z = -6).

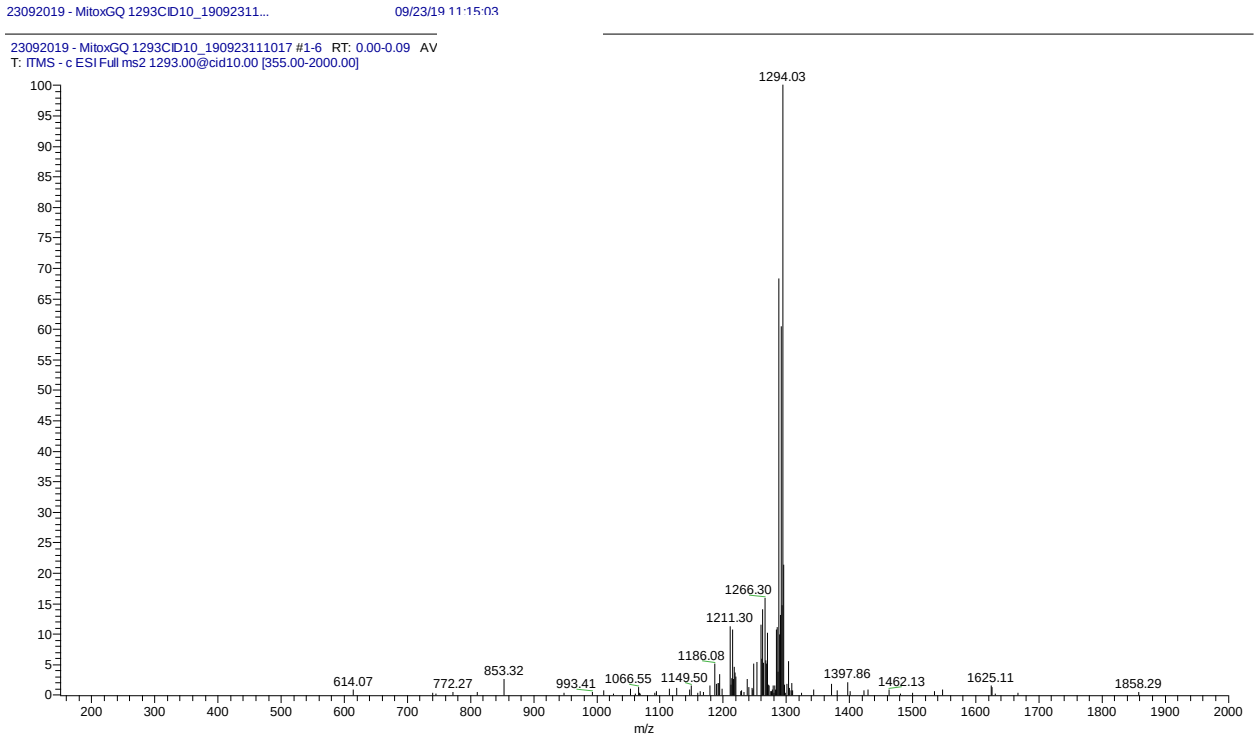


Figure S2.49. CID fragmentation spectrum of the 1:1 complex of mitoxantrone with G4 DNA (z = -6, 34 eV).

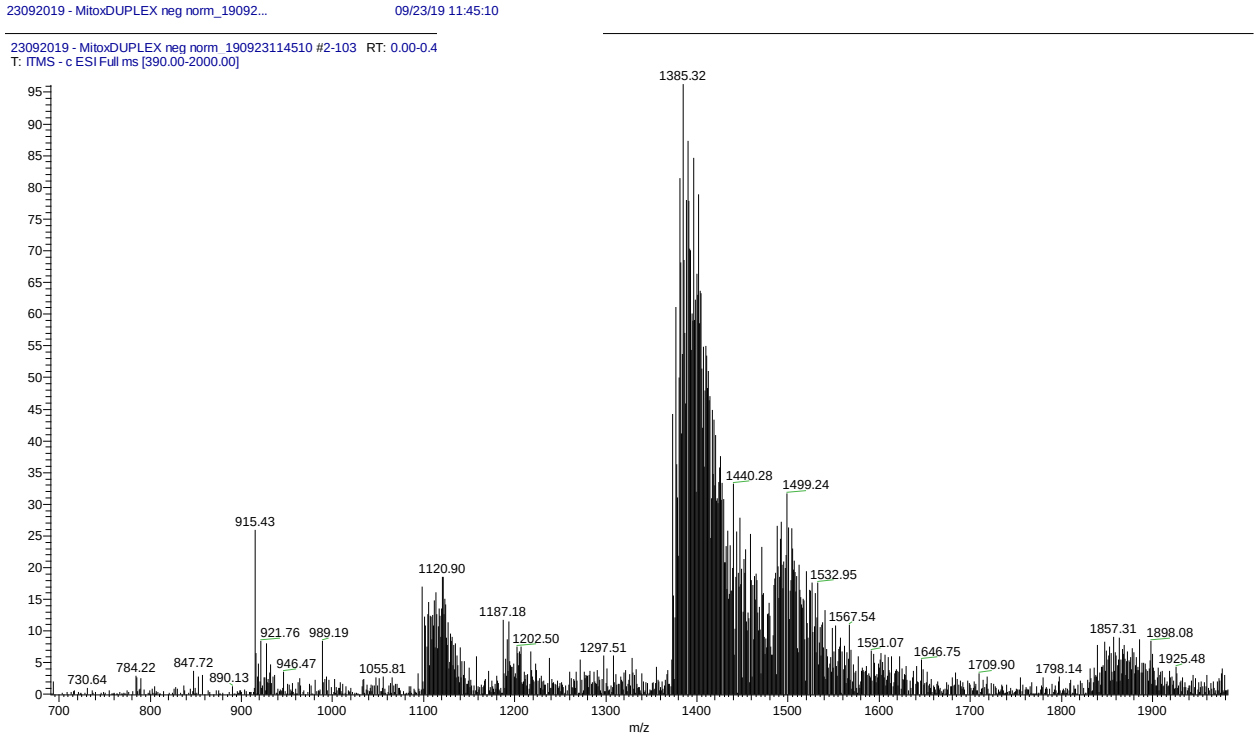


Figure S2.50. ESI-MS experiment showing the interaction of mitoxantrone with ds DNA (z = -8).

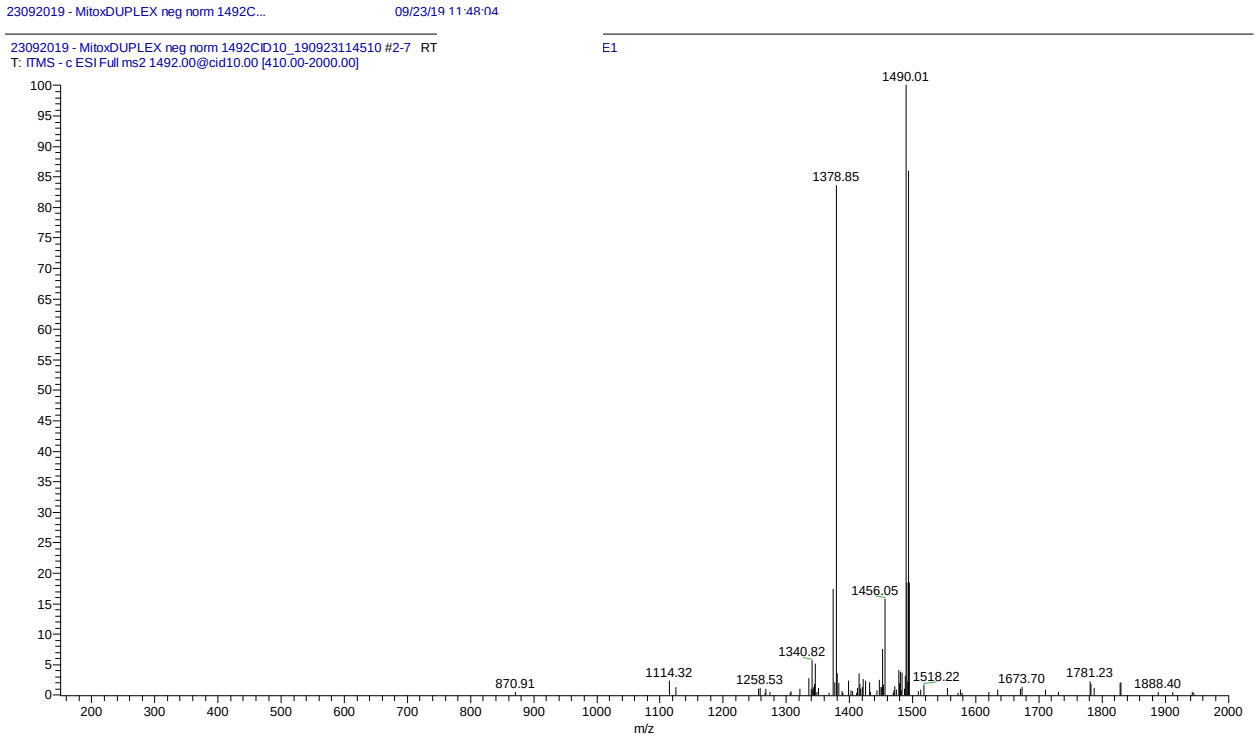


Figure S2.51. CID fragmentation spectrum of the 2:1 complex of mitoxantrone with dsDNA (z = -8, 40 eV).

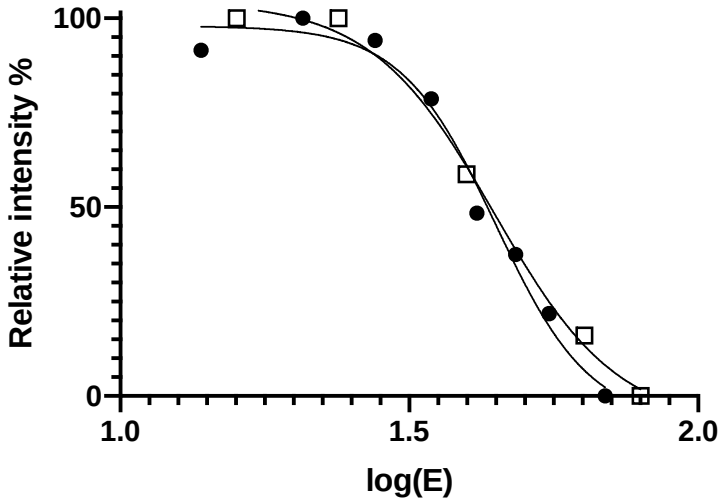


Figure S2.52. CID curve representing the fragmentation of the 1:1 compound-G4 complex (black dots) and 2:1 compound-dsDNA complex (white squares) for mitoxantrone. Energy is expressed in eV.

Fluorimetric titrations

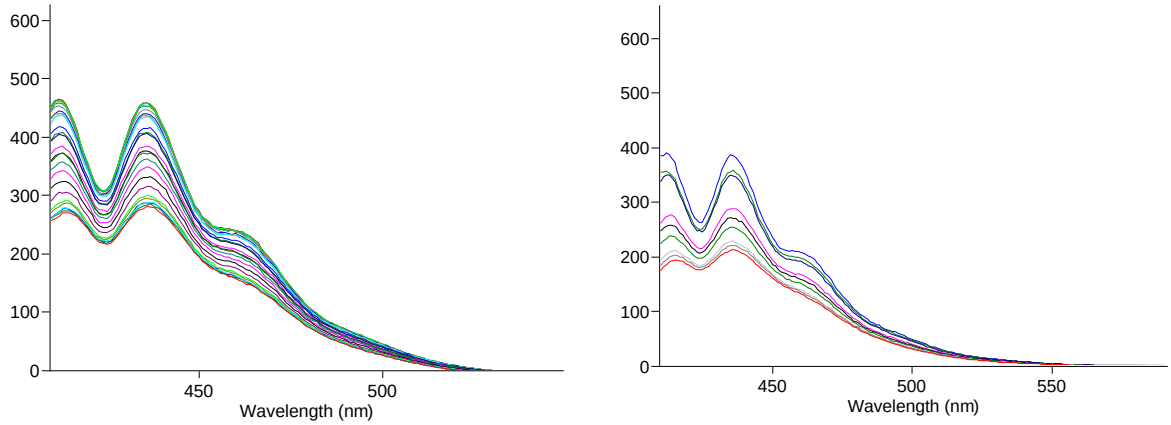


Figure S2.53. Fluorescence emission spectra corresponding to the **4b**-G4 (left) and **4b**-dsDNA titration experiment (right). Increasing amount of DNA promoted a decrease in fluorescence intensity.

3 9,10-Bis[(4-(2-hydroxyethyl)piperazine-1-yl)prop-2-yne-1-yl]anthracene: Synthesis and G4 Selectivity

3.1 Introduction

G4 DNA represents an attractive target for small molecules in the context of drug discovery.[322–324] These arrangements, constituted by stacked guanine nucleobases, can be formed by guanine-rich sequences, such as those that are present in telomeres, promoter genes and several viral genomes. The stabilization of G4 with small ligands is a strategy for interfering with gene expression,[297] uncontrolled cell proliferation,[325] and viral replication.[326–328].

Acridines, naphthalene diimides, anthracenes and anthraquinones are among the synthetic small molecules studied for their non-covalent interaction with dsDNA and G4 structures.[329–332] More specifically, G4 ligands generally bear an aromatic planar scaffold and charged side chains to efficiently bind the non-canonical nucleic acid arrangement.[305,333]

In the previous chapter, a series of monosubstituted anthracene-propargylamine derivatives that were obtained through the A3 coupling reaction were presented. Concerning DNA-binding properties, a set of ESI-MS experiments demonstrated a general preference for dsDNA over G4 and an overall lack of selectivity.[334] It will be here reported the preparation of a novel disubstituted anthracene-propargylamine derivative through a multi-step approach and its preliminary evaluation as a G4 ligand by ESI-MS, molecular docking and MD techniques.

3.2 Materials and methods

Molecular docking

G4 structure file was retrieved from the RCSB Protein Data Bank (PDB, www.rcsb.org) (PDB ID: 1KF1) and was prepared with the Protein Preparation Wizard included in the Schrödinger suite,[310] using default settings, i.e., adding hydrogens, removing waters further than 5 Å from the ligand, adjusting charges, capping termini, optimizing hydrogen bond clusters and performing a final minimization under the OPLS3e force field.[311] The central K^+ ions were retained, the outer one was eliminated. The ligands were prepared for docking with Ligprep tool under OPLS3e force field, using Epik ionizer all possible states in the 7±2 pH range were generated, including the possible metal binding states, allowing therefore the interaction with the potassium ions. The docking protocol consisted of rigid-receptor/flexible ligand docking, which was performed with the Schrödinger suite application XGlide, under the OPLS3e force field at XP precision, the grid center was set at the coordinate origin with the dimensions of 30 and 20 Å for the inner and the outer grid size. The other settings were set as default (*e.g*. 0.80 scaling factor for the ligand atom van der Waals radius).

Molecular dynamics

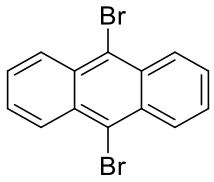
The structures of the G4-ligand complexes, obtained by the ensemble docking study, were prepared for the simulations with Schrödinger application System Builder as follows: the structures were solvated in a orthorhombic water box (10 Å buffer) composed by 4-site model $TIP4P_{ew}$ water molecules,[216] the total charge was neutralized by the addition of an adequate number of K^+ ions, and a 0.20 M KCl concentration was set. The system initially prepared for the simulation Schrödinger default force field OPLS3e was converted through Viparr application to the AMBER OL15 force field specific for nucleic acids,[212–214,288] the parameters set used for the ions were the ones developed by Joung and colleagues.[215,216] The ligands were parameterized with Antechamber application of AmberTools20 using GAFF2 parameters and AM1-BCC atomic charges, missing parameters were generated with Parmchk2. The simulations were conducted with Schrödinger Desmond for 200 ns each, recording 1 frame every 10 ps using the NPT ensemble at the temperature and pressure of 300 K and 1.01 bar, default settings were used for the RESPA integrator (2 fs for bonded interactions, 2 fs for near and 6 fs for far interactions), the system was relaxed prior the simulation using the default Desmond protocol.[335] The trajectories were analysed after the simulations with Simulation Event Analysis Schrödinger application, for the over-time ligand RMSD calculations the ligand atoms positions were fitted to the guanine planes.

Chemistry

General

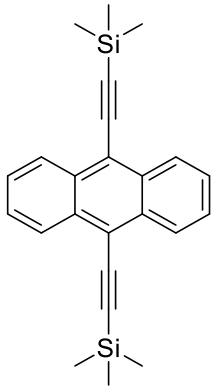
Commercially available chemicals were purchased from Sigma-Aldrich (Saint Louis, MO, USA) and used as received, unless otherwise stated. ^{1}H and $^{13}C\{1H\}$ NMR spectra were recorded on an Avance III 400 MHz spectrometer (Bruker, Billerica, MA, USA). All spectra were recorded at room temperature; the solvent for each spectrum is given in parentheses. Chemical shifts are reported in ppm and are relative to TMS internally referenced to the residual solvent peak. Datasets were edited with TopSpin (Bruker). The multiplicity of signals is reported as singlet (s), doublet (d), triplet (t), quartet (q), multiplet (m), broad (b), or a combination of any of these. Mass spectra were recorded by direct infusion ESI on LCQ Fleet (Thermo Fisher Scientific, Waltham, MA, USA) spectrometer. The purity profile (>96% unless otherwise stated) was assayed by HPLC using a Pro-Star system (Varian, Palo Alto, CA, USA) equipped with a 1706 UV–Vis detector (254 nm, Bio-rad, Hercules, CA, USA) and an C-18 column (5 µm, 4.6 × 250 mm, Agilent Technologies, Santa Clara, CA, USA). An appropriate ratio of water (A) and acetonitrile (B) was used as mobile phase with an overall flow rate of 1 mL/min; the general method for the analyses is reported here: 0 min (95% A–5% B), 5 min (95% A–5% B), 15 min (5% A–95% B), 20 min (5% A–95% B), and 22 min (95% A–5% B).

Synthesis of 9,10-Dibromoanthracene (**1**)



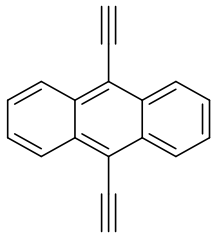
Anthracene (4.5 g, 25.0 mmol) was dissolved in $CHCl_3$ (50 mL) and a solution of bromine (3.0 mL, 58.2 mmol) in $CHCl_3$ (25 mL) was added dropwise, inducing the formation of a solid precipitate. The mixture was stirred at r.t. for 4 h and afterwards the solvent was removed with a stream of N_2. The solid residue was triturated with DCM, filtered and washed with small portions of DCM to give **1** in yellow needles (4.1 g, 36%). ^{1}H-NMR (400 MHz, $CDCl_3$): δ (ppm) = 8.57 (m, 4H), 7.62 (m, 4H), ^{13}C-NMR (100 MHz, $CDCl_3$): δ (ppm) = 131.2, 128.4, 127.6, 123.7.

Synthesis of 9,10-Bis((trimethylsilyl)ethynyl)anthracene (**2**)



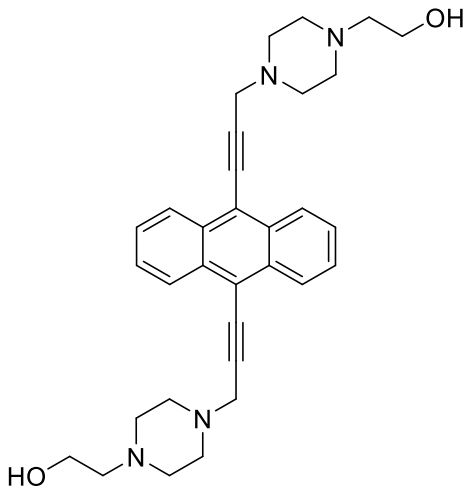
Compound **1** (236 mg, 0.70 mmol) and trimethylsilyl acetylene were dissolved in a mixture of toluene (37.0 mL) and triethylamine (8.4 mL). Tetrakis(triphenylphosphine)palladium(0) (17 mg, 0.015 mmol) and copper iodide (5.5 mg, 0.029 mmol) were then added. The mixture was purged with nitrogen and heated at reflux overnight. The solvent was removed at reduced pressure and the obtained mixture was poured in water and extracted with DCM, dried over anhydrous magnesium sulphate and evaporated under reduced pressure. The crude product was purified by column chromatography (silica gel, hexane) obtaining compound **2** as a red solid (86 mg, 36%). ^{1}H-NMR (400 MHz, $CDCl_3$): δ (ppm) = 8.61 (m, 4H), 7.63 (m, 4H), 0.47 (s, 18H),^{13}C-NMR (100 MHz, $CDCl_3$): δ (ppm) = 132.1, 127.2, 126.4, 117.5, 108.3, 101.7, 0.2.

Synthesis of 9,10-Diethynylanthracene (**3**)



Compound **2** (80 mg, 0.22 mmol) was dissolved in a mixture of MeOH (6.4 mL) and THF (3.2 mL), and 10% KOH was then added (1.2 mL). The solution was stirred at r.t. for 3 h and then the solvent was partially evaporated under reduced pressure. After that, diethyl ether (20 mL) was added, the obtained mixture was washed with water (3 × 10 mL), dried over magnesium sulphate and evaporated under reduced pressure obtaining a grey solid (45 mg, 91%). ^{1}H-NMR (400 MHz, $CDCl_3$): δ (ppm) = 8.61 (m, 4H), 7.62 (m, 4H), 4.07 (s, 2H),^{13}C-NMR (100 MHz, $CDCl_3$): δ (ppm) = 132.5, 127.2, 127.10, 117.5, 90.1, 80.3.

Synthesis of 9,10-Bis[(4-(2-hydroxyethyl)piperazine-1-yl)prop-2-yne-1-yl]anthracene (**4**)



Compound **3** (45 mg, 0.2 mmol), 37% formaldehyde (200 µL, 2.5 mmol), copper iodide (7 mg, 0.04 mmol) and 1-(2-Hydroxyethyl)piperazine (80 mg, 0.6 mmol) were dissolved in DMSO (1 mL). The solution was stirred overnight at r.t. and then DCM (30 mL) was added. The resulting mixture was washed with 0.1 M KOH (4 × 10 mL). The crude product was purified by column chromatography (silica gel, gradient of MeOH/TEA/DCM from 2,5/0,75/96,75 to 10/2/88) obtaining a solid yellow product (25 mg, 25%). ^{1}H-NMR (400 MHz, d-DMSO): δ (ppm) = 8.54 (m, 4H), 7.74 (m, 4H), 4.36 (t, 2H, J = 5.4 Hz), 3,90 (s, 4H), 3,50 (m, 4H), 2,72 (bs, 8H), 2,51 (bs, 8H), 2,41 (t, 4H, J = 6.2 Hz),^{13}C-NMR (100 MHz, $CDCl_3$): δ (ppm) = 132.8, 128.7, 127.2, 117.0, 95.4, 82.2, 59.5, 55.4, 52.9, 51.6, 48.2. ESI-MS m/z: Calcd for $C_{32}H_{39}N_4O_2^+$ $(M + H)^+$: 511.31, found 511.04.

ESI-MS binding studies

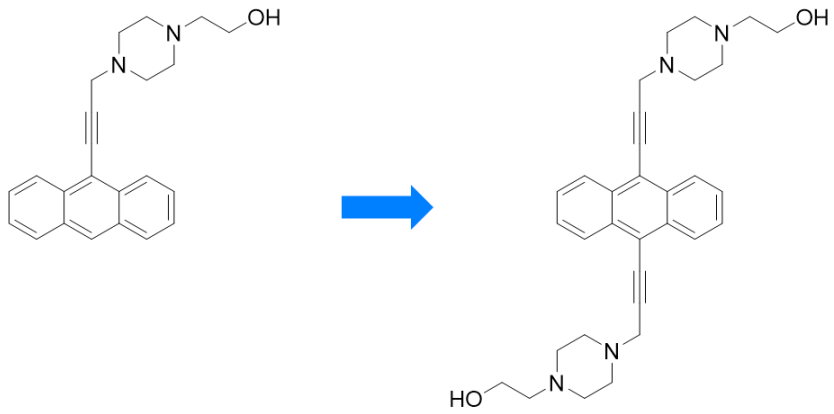
DNA sequences were obtained from Sigma-Aldrich (Saint Louis, MO, USA). Samples were heat-denatured and folded in 150 mM ammonium acetate before incubation with the ligand; ligand stock solution was prepared in methanol. The final concentration of the oligonucleotide was 5–10 µM in 150 mM ammonium acetate, with a 10:1 compound/oligo ratio. Samples were acquired after an equilibration time of 30 min and methanol was added to obtain a stable ESI signal. Mass spectra were recorded by direct infusion. The instrument was set in negative ionization mode with a 3.4 kV capillary voltage, 120 °C capillary temperature and a flow rate of 5 µL/min. Collision-induced dissociation (CID) experiments were performed on the complexes by isolating the precursor ion in the trap and increasing the "normalized collision energy" parameter. Exact mass for the G4 sequence (5'-AGGGTTAGGGTTAGGGTTAGGGT-3'): 7270.774 Da. Exact mass for dsDNA (5'-ACTATTTACGTATAATGA-3', 5'-TCATTATACGTAAATAGT-3',): 10987.922 Da. Exact mass for 9,10-bis[(4-(2-hydroxyethyl)piperazine-1-yl)prop-2-yne-1-yl]anthracene (**4**): 510.299 Da. Binding affinity was calculated according to the formula: BA= $(\Sigma DNA_{bound}/(\Sigma DNA_{free}+ \Sigma DNA_{bound})) \times 100$.[225,336]

In CID studies, relative intensity (%) of the DNA/ligand signal was plotted against the applied energy expressed in eV (logarithmic scale). Relative I (%) = ($I_{COMPLEX}$/($I_{COMPLEX}$+$I_{DISSOCIATION\ PRODUCTS}$)) x 100.

3.3 Results and discussion

In this work, the best G4 ligand (**4b**) of the series of the previously reported study (Chapter 2) was selected for a further development through its structural modification. The molecule is composed of an anthracene planar aromatic scaffold and by a hydroxy ethyl piperazine side chain, which can be easily protonated at physiological conditions and presents a primary alcohol moiety. These features are indeed crucial for the recognition of G4-DNA, however they can be also optimal for the interaction with the canonical dsDNA. This presumably happens by intercalation of the tricyclic ring in-between two base-pairs pointing the positively charged side chain towards the negatively charged DNA groove.

The addition of a second, identical, side chain, generates a symmetrical molecule (Figure 3.1), which is structural less congenial to the dsDNA due to the resulting evident impediment to the intercalation, but is more congenial to the G4 arrangement thanks to the possibility to stack on an external G-quartet while pointing the two side chains in two opposite grooves.



4b **4**

Figure 3.1. Comparison between the two mono- (**4b**) and the di-substituted (**4**) ligand generations.

Computational study

In order to evaluate the effects given by the addition of a second polar side chain to the anthracene mono-propargylamine scaffold a computational study was performed. In particular MD simulation of the apo receptor was followed by an ensemble docking and then by a MD simulation of the complexes.

As a model of the G4 arrangement, the X-ray crystal structure of a parallel G4 motif was taken in consideration (*apo*-form). The structure (PDB ID: 1KF1) was first prepared with Schrödinger and then subjected to a 200 ns molecular dynamics (MD) simulation under the OL15 force field (see for details the Materials and methods section) specific for simulation of G4-DNA arrangements.[212–214,216,288] A highly stable trajectory was obtained in terms both of over-time RMSD of the guanine atoms and also for the whole G4 structure (Figure 3.2), minor differences were observed from the crystallized one due to the relaxation of the structure in a simulated aqueous environment. This is indeed beneficial for the following interaction studies with the ligand, which is intended to mimic the cellular conditions.

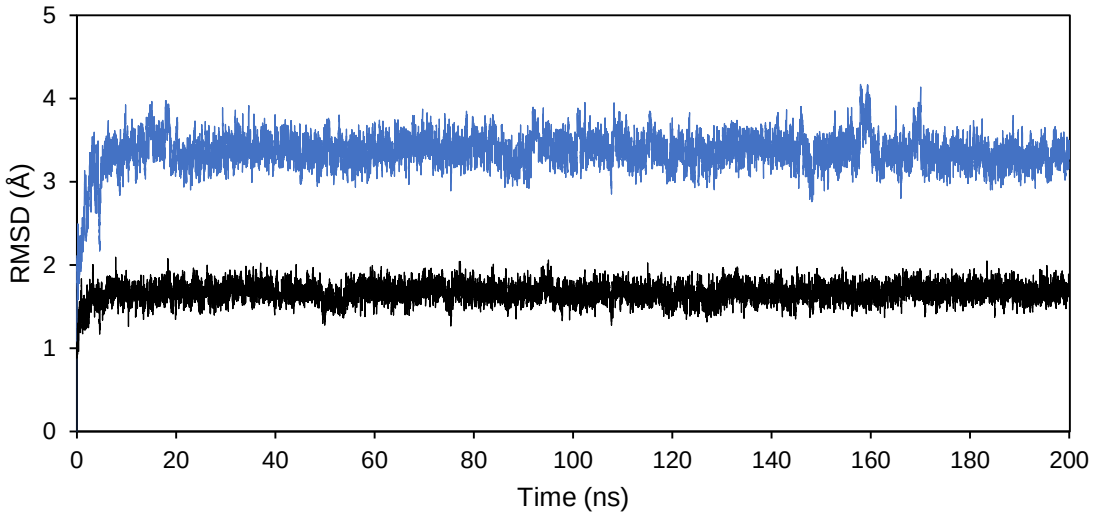


Figure 3.2. Graph showing the over-time RMSD values for the whole G4 structure (blue) and for the guanines (black) obtained from the 200 ns MD simulation trajectory of G4 *apo*-form.

From the resulting trajectory, a set of 10 frames was extracted and they were used as single G4 receptors for an ensemble docking with Glide at XP precision through the XGlide application present in the Schrödinger suite.

Then to investigate the stability of the complexes, for each of the two ligands, the best poses obtained in the docking stage were used as the starting structures for a MD simulation under the OL15 force field for 200 ns.

The combined computational study revealed a good interaction of both the compounds with the parallel G4 structure; the ligands formed a complex with the G4 receptor through an external binding interaction mode also pointing the charged side chains towards the G4 grooves (Figure 3.3). Moreover, the di-substituted compound **4** outperformed the mono-substituted **4b** in terms of the calculated energy of interaction, represented by the Glide Score values which were -10.3 and -13.0 kcal/mol for the mono- and for the di-substituted compounds respectively.

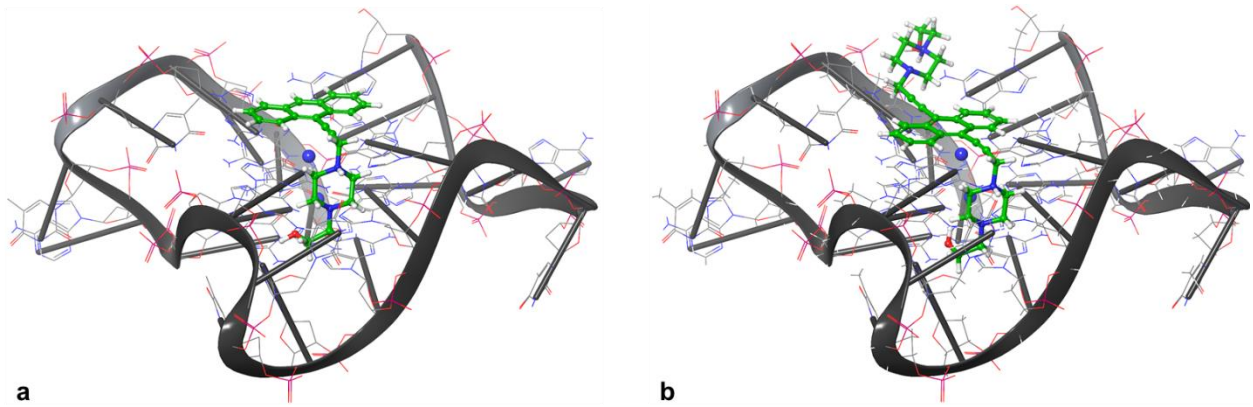


Figure 3.3. Best poses obtained through the ensemble docking study for the mono- (a) and the di-substituted (b) ligands with the G4 arrangement.

The analysis of the ligands-G4 interactions showed a similar pattern. **4b** developed with the G4 receptor 3 stacking interactions between its anthracene core and 2 guanines, 3 H-bonds between the polar groups of its side chain and 2 guanines and 1 thymine, and 1 ion pairing interaction between the charged amine and the phosphate of a guanine (Figure 3.4, a).Compound **4** formed with the G4 arrangement 4 stacking interactions between two anthracene rings and the two rings of a guanine, 1 H-bond interaction between its hydroxyl moiety and a guanine, 1 H-bond between one charged amine and a thymine, 2 ion paring interaction between the two charged amines and two phosphates, and 1 pi-cation between one charged amine and a guanine ring (Figure 3.4, b).

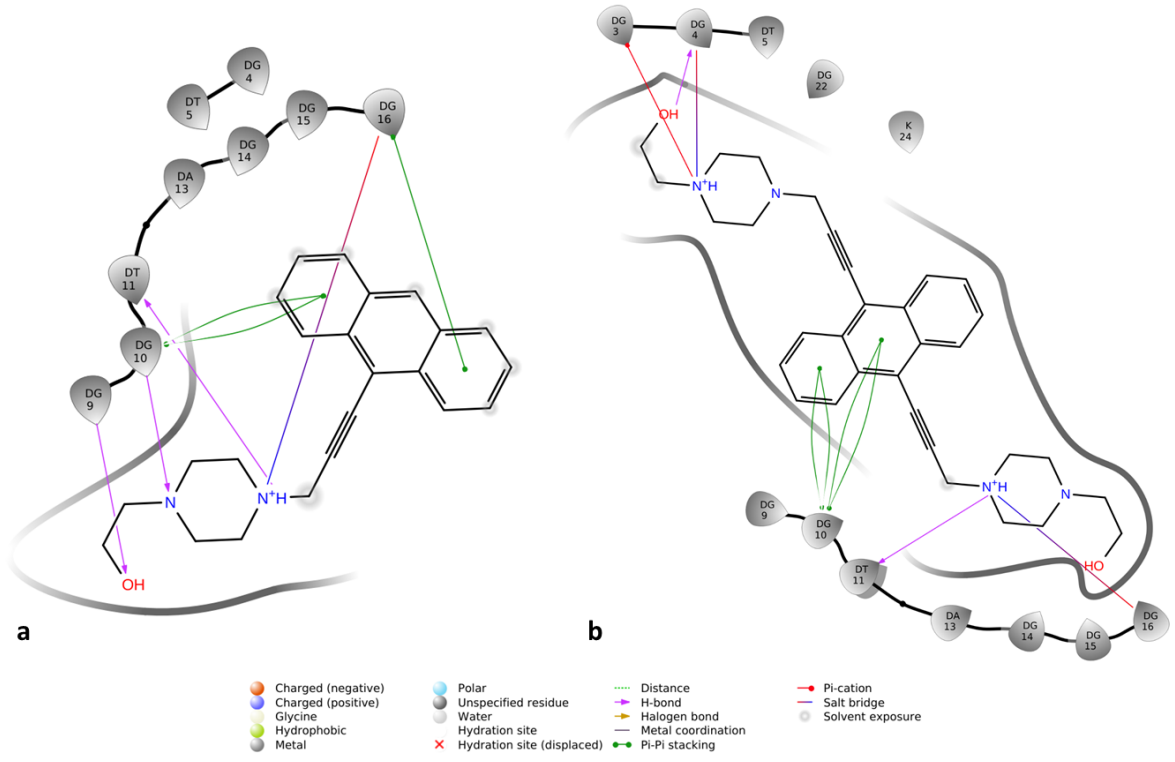


Figure 3.4. Interaction diagram for the G4 complex with the mono- (**a**) and di-substituted ligands (**b**).

The subsequent MD simulation demonstrated for both the ligands an overall stability of the complexes. Furthermore, the ligands remained for all the simulation time frame in the initial stacking positions obtained by docking. A minor fluctuation of one of the side-chains was recorded for the di-subtituted ligand. The trajectories of the MD simulations are reported in Figure 3.5.

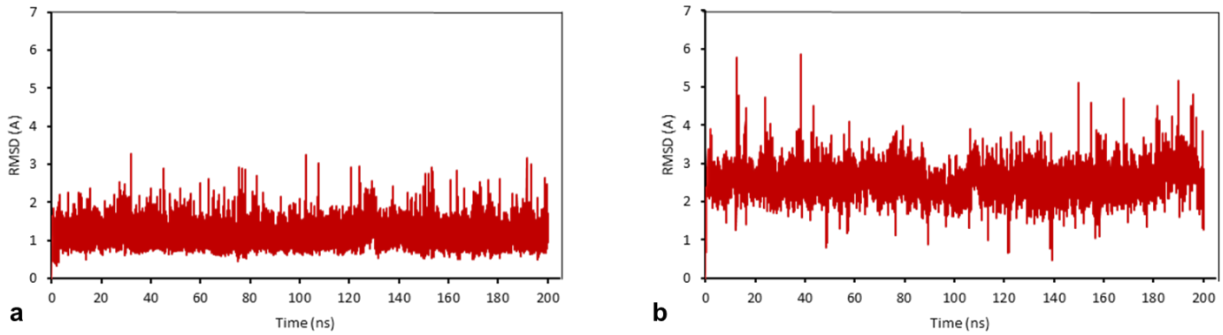
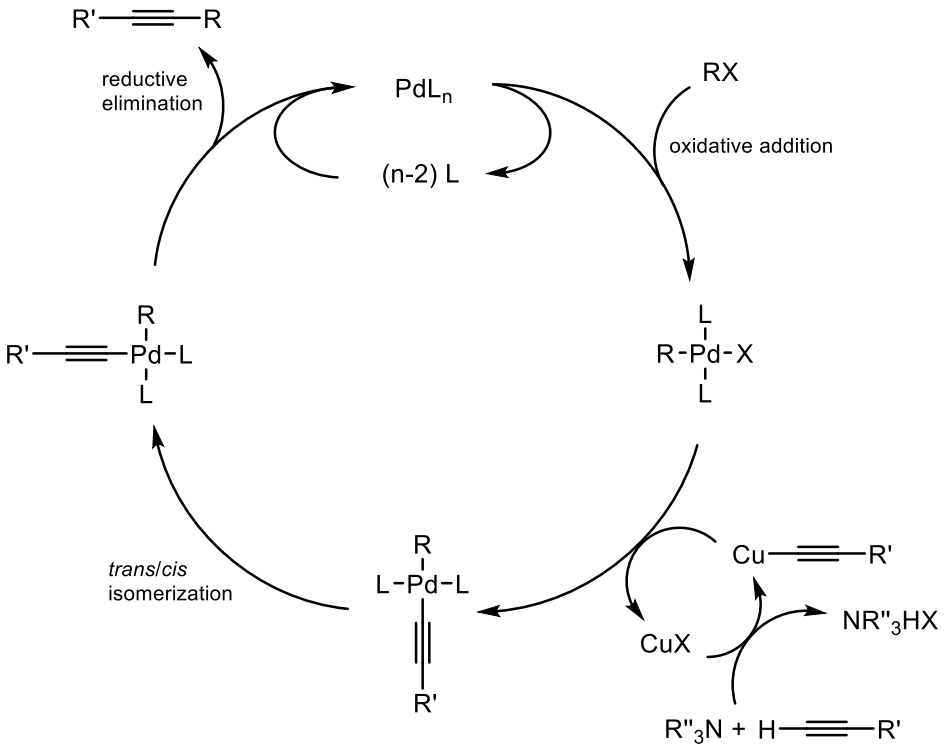


Figure 3.5. Graphs showing the over-time RMSD for the ligands from the 200 ns MD simulation trajectories for compound **4b** (**a**) and **4** (**b**).

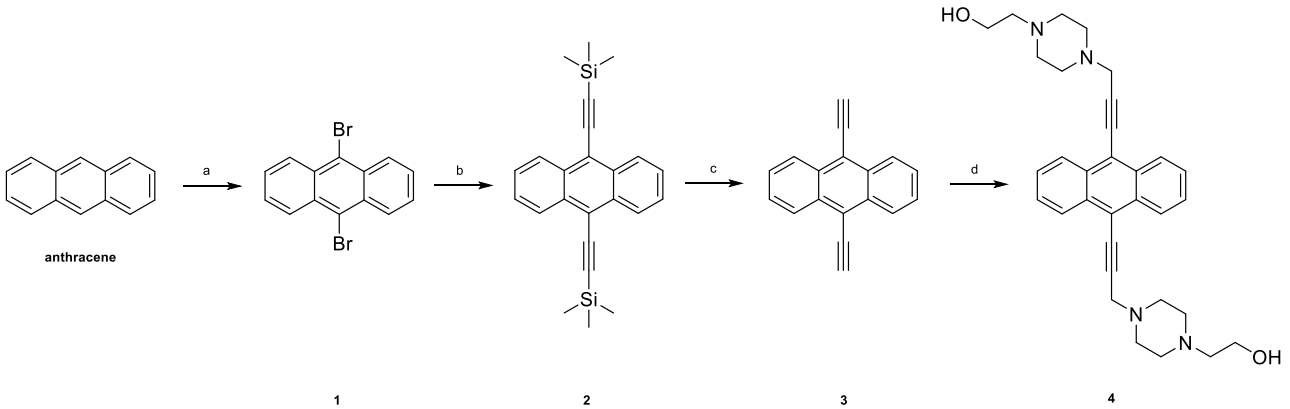
Chemistry

The target ligand **4** was synthesized by a multi-step procedure starting from anthracene. Specifically, anthracene was selectively di-brominated in positions 9,10 with elemental bromine producing compound **2**. In order to attach the two side chains to the scaffold, the intermediate was subjected to a double Sonogashira coupling reaction with trimethylsilylacetylene on the two bromides. More specifically the coupling reaction took place using Tetrakis(triphenylphosphine)palladium(0) as the Pd(0) source and as the main catalyst and CuI as the Cu(I) source and co-catalyst, an exemplificative mechanism is depicted in the following scheme (Scheme 3.1).[337]



Scheme 3.1. Proposed catalytic cycle for the Sonogashira reaction including Pd and Cu cycles.[337]

The trimethylsilylacetylene groups were then de-protected in basic conditions yielding compound **3** bearing the two alkynes in position 9,10. In conclusion the final compound **4** was obtained from **3** by the Cu(I) catalyzed A3 coupling reaction in the presence of formaldehyde and the secondary amine 1-(2-Hydroxyethyl)piperazine. The complete synthetic procedure is reported in the following scheme (Scheme 3.2).



Scheme 3.2. Multi-step procedure for the preparation of compound **4** starting from anthracene.

ESI-MS binding study

Analogous to the previous study on monosubstituted anthracenes,[334] ESI-MS was used to evaluate the binding of compound **4** to a telomeric 23-mer sequence containing the TTAGGG repeat, which forms a monomeric G4 (Figure 3.6).[305,336] To investigate sequence selectivity, a dsDNA oligonucleotide was also used as control in the experimental design. The efficiency of the interaction was calculated in terms of binding affinity (BA) considering signal intensities in the mass spectra.[320,321] Compound **4** interacts with both DNA structures, but it preferentially binds to the G4 arrangement. In fact, as resumed in Table 1, a higher BA value was recorded for the compound **4**/G4 complex. The results for the corresponding mono-substituted ligand **4b**, which demonstrated to be a non-selective dsDNA binder in the previous study, are also reported for comparison.[334] A G4/dsDNA selectivity ratio of 1.92 was calculated for compound **4**, demonstrating that the presence of two side chains in the scaffold is able to increase the selectivity of the ligand towards the G4 motif.[148] Furthermore, as previously observed for the corresponding monosubstituted derivative, a 2:1 binding stoichiometry was detected.[334]

Table 3.1. Binding affinity (BA) values for compounds **4** and **4b** towards G4 and dsDNA, as calculated from the ESI-MS binding experiment. The selectivity ratio was calculated dividing the G4 BA value by the dsDNA BA value.

Compound	BA G4	BA dsDNA	Selectivity ratio (G4/dsDNA)
4	80.3	41.8	1.92
4b	68.4	79.8	0.86

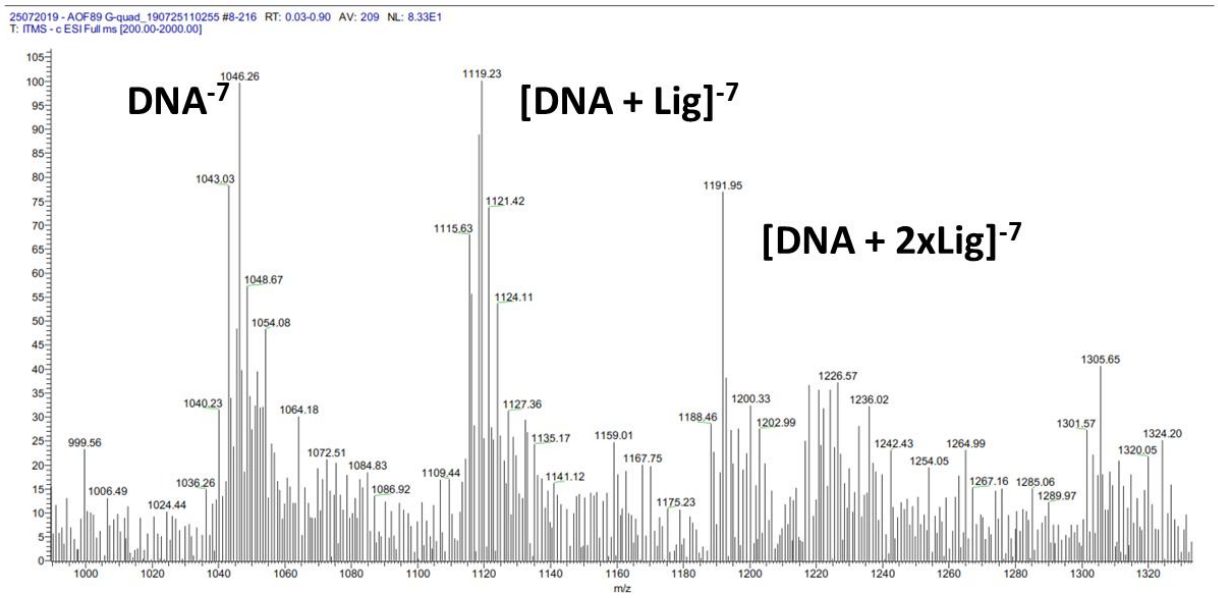


Figure 3.6. ESI-MS binding experiment of compound **4** with G4 DNA. Unbound G4 DNA (DNA^{-7}, m/z = 1046.26, z = -7), ligand/G4 complex with 1:1 stoichiometry ($[DNA+LIG]^{-7}$, m/z = 1119.23, z = -7) and ligand/G4 complex with 2:1 stoichiometry ($[DNA+2xLig]^{-7}$, m/z = 1191.95, z = -7) were detected.

Collision-induced dissociation (CID) experiments were used to investigate the gas-phase relative kinetic stability of small molecule-DNA complexes ($E_{COM}^{50\%}$), a parameter of biological relevance.[323,338]. Dissociation by loss of the small molecule was observed and complex stability was evaluated by increasing collision energy. The results reported in Table 3.2 show that compound **4** efficiently stabilizes G4 in the 2:1 stoichiometry.[338]

Table 3.2. Relative gas-phase kinetic stability ($E_{COM}^{50\%}$) investigated by collision-induced dissociation (CID).

Compound	$E_{COM}^{50\%}$ (eV) 2:1 Ligand/G4	$E_{COM}^{50\%}$ (eV) 1:1 Ligand/G4	$E_{COM}^{50\%}$ (eV) dsDNA
4	43.58	34.39	33.42
4b	40.07	34.57	39.43

3.4 Conclusion

In this study, the anthracene-propargylamine scaffold **4b**, developed for the interaction with DNA in the previous work, was optimized for increasing its affinity for the G4 arrangement. For this reason, a di-substituted ligand **4**, characterized by two symmetrical propargylamine side chains was designed. An initial computational study, including both the original mono and the di-substituted ligands, was performed by means of ensemble docking and MD experiments using as a model G4 receptor the structure of a parallel G4 motif. The study predicted for both the ligands a preferential stacking binding mode on top of the external G-quartet. In the case of the di-subtituted ligand **4**, an increased interaction energy was observed respect to the mono-substituted **4b**. Moreover, an overall stability of the resulting complex was assessed by MD validating the docking poses. The molecule was then prepared through a four-step synthesis including a *Sonogashira* coupling reaction followed by an A3 coupling reaction. In conclusion, the ligand was tested for its ability to bind to a telomeric G4 DNA through an ESI-MS binding assay which demonstrated for the newly prepared ligand, respect to the former one, a higher BA and selectivity for the G4 DNA.

3.5 Supporting material

*Characterization of the final compound **4***

ESI-MS binding study

*Characterization of the final compound **4***

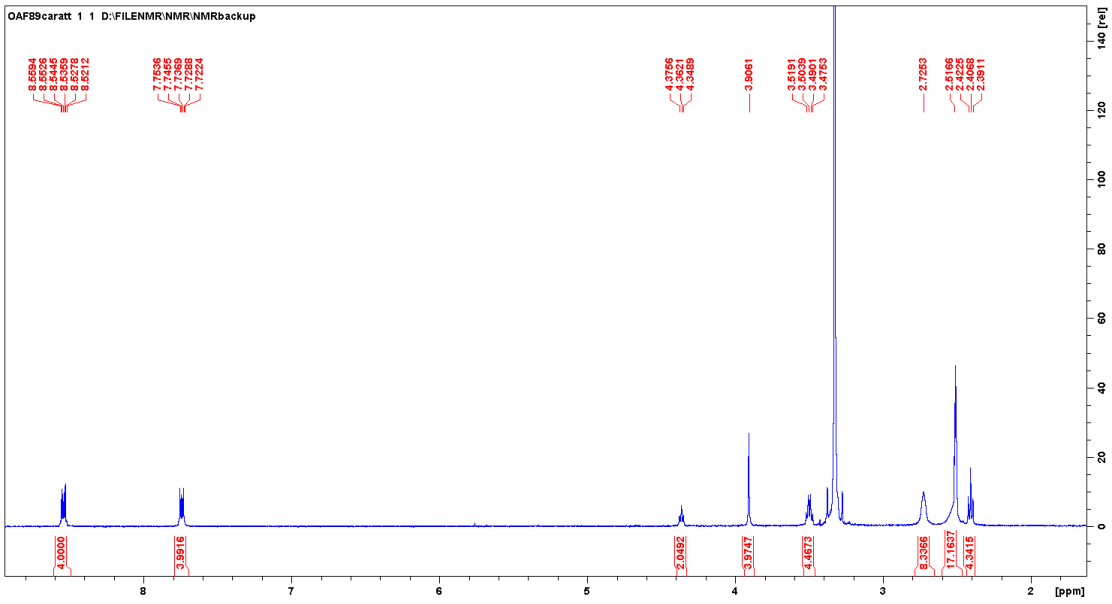


Figure S3.1. ^{1}H-NMR spectrum of compound **4**. Note: the signal of 4 x CH_2 protons (piperazine, bs) partially overlap with the signal of the solvent (d-DMSO).

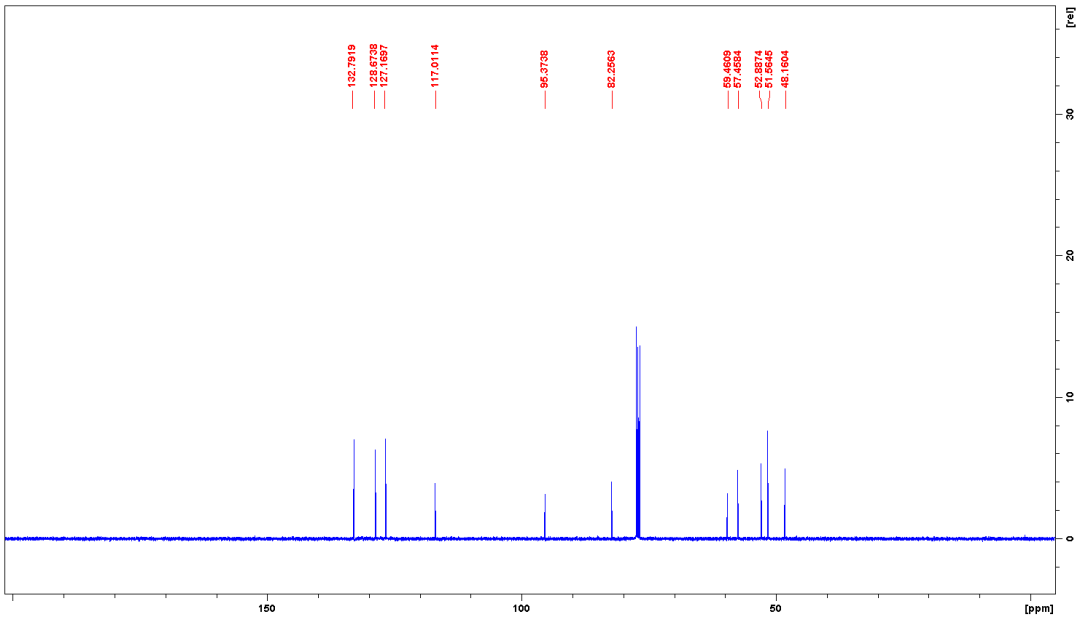


Figure S3.2. ^{13}C-NMR spectrum of compound **4** (d-DMSO).

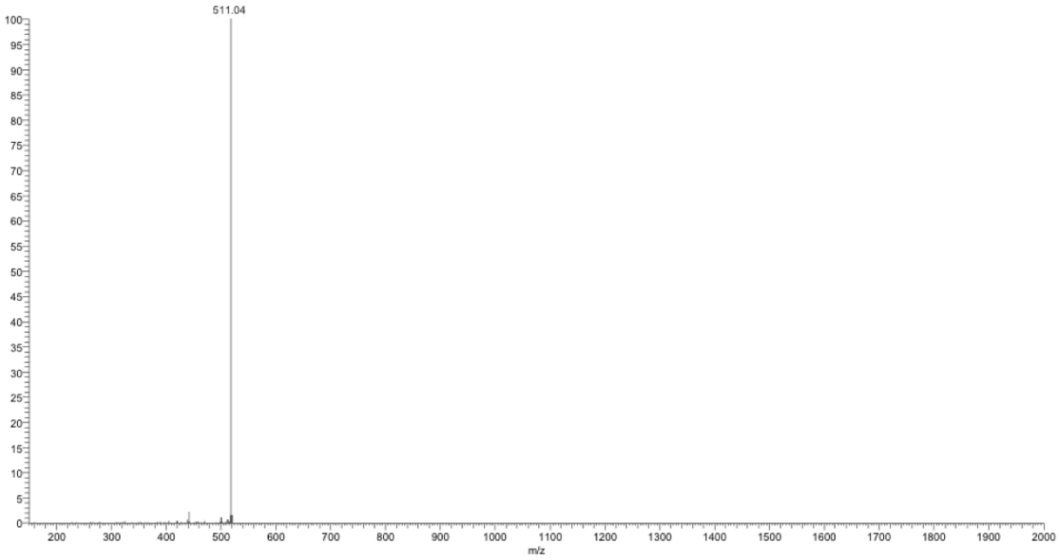


Figure S3.3. ESI-MS spectrum of compound **4** (positive ionization mode).

ESI-MS binding studies

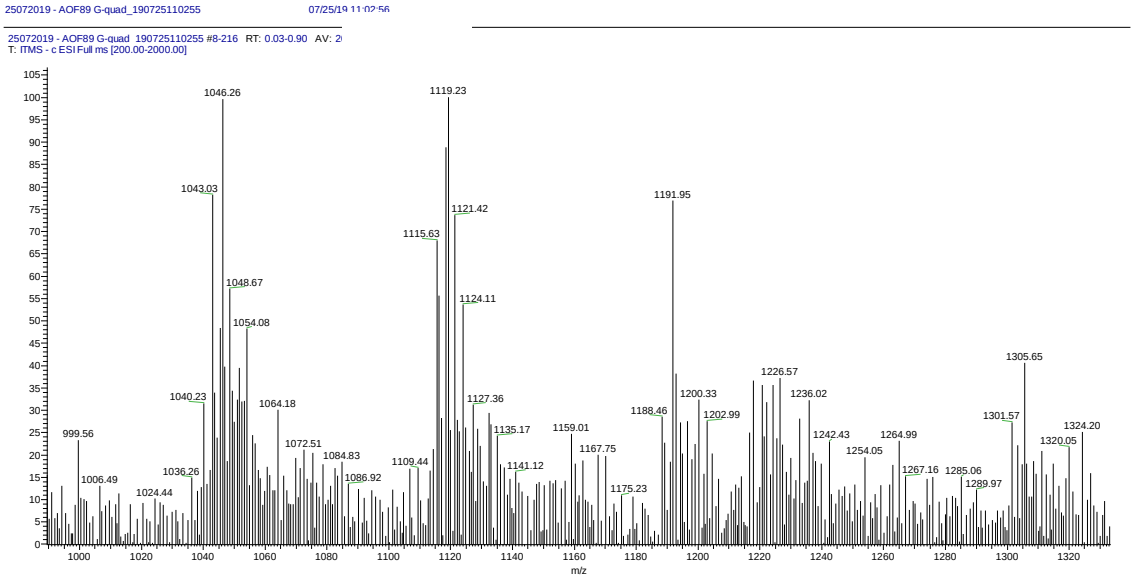


Figure S3.4. ESI-MS binding experiment of compound **4** with G4 DNA. Unbound G4 DNA (m/z = 1046.26, z = -7), ligand/G4 complex with 1:1 stoichiometry (m/z = 1119.23, z = -7) and ligand/G4 complex with 2:1 stoichiometry (m/z = 1191.95, z = -7) were detected.

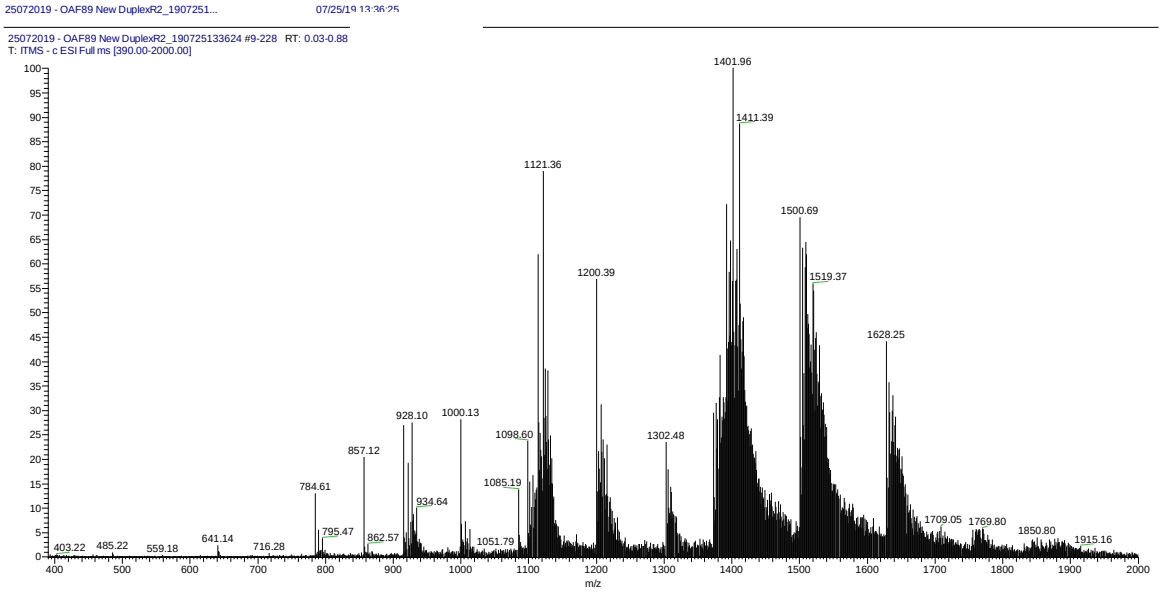


Figure S3.5. ESI-MS binding experiment of compound **4** with double-stranded DNA. Unbound dsDNA (m/z = 1121.36, z = -9), ligand/dsDNA complex with 1:1 stoichiometry (m/z = 1200.39, z = -9) and ligand/dsDNA complex with 2:1 stoichiometry (m/z = 1302.48, z = -9) were detected. The same behavior was observed in the z = -8 charge state.

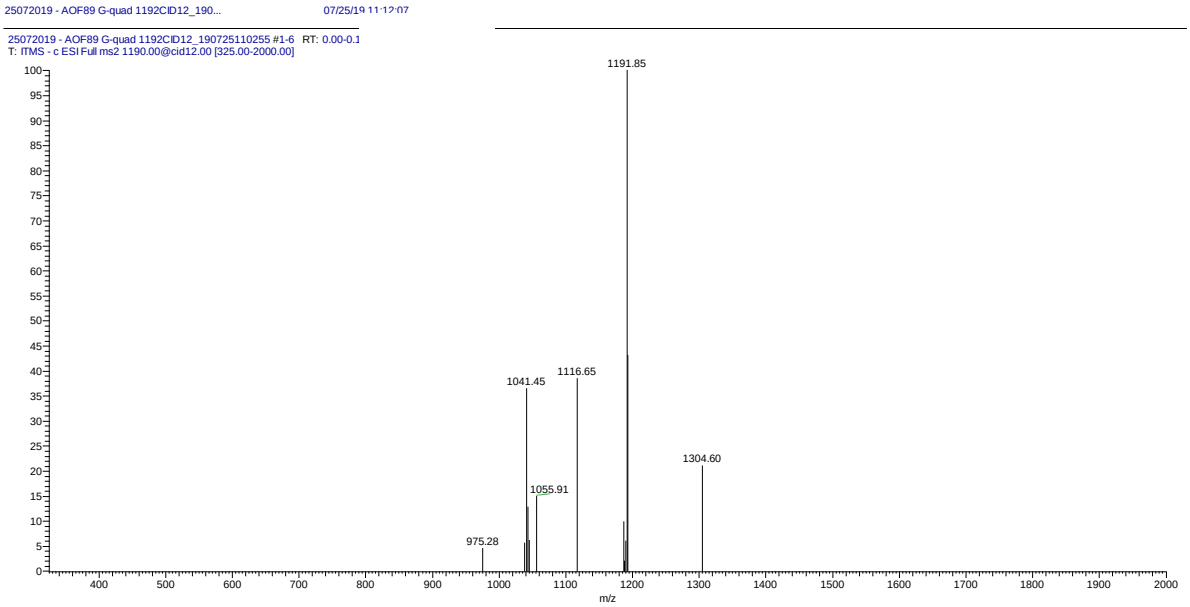


Figure S3.7. CID fragmentation spectrum of the ligand/G4 complex with 2:1 stoichiometry (normalized collision energy = 12). The fragmentation of the 2:1 ligand/G4 complex (m/z = 1191.85, z = -7) promoted the formation of the peaks corresponding to the 1:1 complex (m/z = 1116.65, z = -7) and to the unbound DNA (m/z = 1041.45, z = -7). Please note that the difference in m/z value for fragmentation peaks with respect to Figure S4 is due to instrumental resolution.

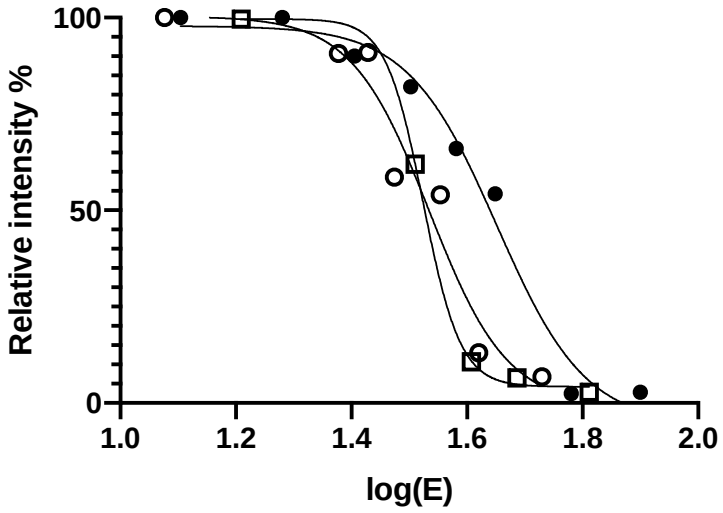


Figure S3.8. CID dissociation curves of the complexes. Black dots: ligand/G4 complex with 2:1 stoichiometry, white dots: ligand/G4 complex with 1:1 stoichiometry, white squares: ligand/dsDNA complex with 2:1 stoichiometry.

4 Amino acid-anthraquinone click chemistry conjugates selectively target human telomeric G-quadruplex

4.1 Introduction

The canonical structure of DNA is known as B-form double helix. However, guanine-rich sequences can fold into the non-canonical arrangement called G-quadruplex (G4). These motifs are constructed via pi-pi stacking of planar G-quartets, formed by four guanines in a single layer. The latter arrangement is maintained by Hoogsteen-type base-pair hydrogen bonds between N1H and O6, and between N2H and N7.[58] Moreover, the coordination of various mono-, di- and even tri-valent cations at G-quartets is crucial for G4 stabilization. Different topologies are possible for G4s depending on the considered sequences and on the environmental conditions. For both intramolecular and intermolecular sequences, parallel, antiparallel and hybrid arrangements have been reported. More specifically, parallel topology is prevalent in crowded conditions as it occurs in the cell.[72] G4 forming sequences are present in the promoter region of several oncogenes such as c-myc, vegf, bcl2, k-ras and c-kit. Moreover, induction of these structures represents an interesting strategy for the selective inhibition of their transcription.[339] G4s can be also detected at the telomeres, the terminal domains of chromosomes whose length regulates the life cycle of the cells. In addition, telomerase, the enzyme that elongates telomeres by adding telomeric repeats (TTAGGG in mammals), is almost absent in non-immortalized cells but it is expressed at significant levels in most spontaneously immortalized cell and human cancer cells.[179]

Different G4 stabilizing molecules such as anthraquinones (AQs),[225,332,340] porphyrins, perylenes, acridines, anthracenes,[334,341] and also natural compounds,[324,342] have been shown to inhibit the telomere elongation operated by telomerase, and several other derivatives are under investigation as potential anticancer drugs. AQs are well-known nucleic acids ligands.[225,332,340,343] More in detail, it has been observed that the position of side chains on the central scaffold is crucial for the selectivity for specific DNA arrangements: the 1,4-substitution is preferred for the interaction with the duplex whereas the insertion of side chains at position 2,6 or 2,7 favours the recognition of triplex and G4 motifs (see chapter 1.5.1 of the Introduction for a more exhaustive treatment of anthraquinones as G4 ligands).[228] In this connection, we previously reported a series of 2,6- and 2,7-bis-substituted anthraquinone-amino acid conjugates that act as G4 selective binders and stabilizers. In such compounds, different amino acids or short peptides were introduced as side chains directly bound to the central core, to allow them to fit within the G4 grooves. In particular, 2,6-AQ derivatives bearing *β*-Ala, *β*-Ala-Lys and *β*-Ala-Arg side chains, endowed with positive charges, demonstrated a marked ability to stabilize telomeric G4s. Nevertheless, complete selectivity for G4 over duplex DNA was not fully achieved.

Taking the step from these previous findings, with the current work we aimed at enhancing the 2,6-AQ–amino acid conjugate scaffold employing a new triazole moiety as the linker for attaching amino acids and dipeptides to the central tricyclic core. More in detail, the 1,2,3-triazole, generated by the click chemistry Copper(I)- catalysed Azide-Alkyne Cycloaddition (CuAAC) reaction, is a polar aromatic ring capable of developing hydrogen bond and pi-interactions, thus it represents an ideal moiety to be included in a side chain designed to interact with nucleic acids.[344,345] Moreover, the CuAAC reaction is a valuable tool for the development of a library of G4 ligands, as it empowers a flexible approach for the prompt derivatization of the scaffold with side chains of different size or chemical nature.

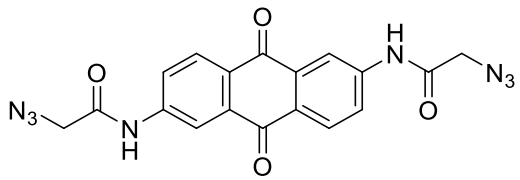
The newly synthesized compounds were subjected to a preliminary binding assay to telomeric G4 and double stranded DNA (dsDNA) by means of electrospray mass spectrometry (ESI-MS). Then, fluorescence melting and circular dichroism (CD) studies were performed to gain more insights on the physico-chemical properties of the interaction and provide preliminary information on sequence selectivity. Additionally, a combined computational study composed by molecular docking and molecular dynamics simulations was performed in order to predict, for the various G4 topologies, the structure of ligand-G4 complexes and the involved intermolecular interactions.

4.2 Materials and methods

Chemistry

Commercially available chemicals were purchased from Sigma-Aldrich and used without any further purification if not specified elsewhere. NMR experiments were performed on a Bruker Avance III 400 spectrometer (frequencies: 400.13 and 100.62 MHz for ^{1}H and ^{13}C nuclei, respectively) equipped with a multinuclear inverse z-field gradient probe head (5 mm). For data processing, TopSpin 4.0.8 software was used, and the spectra were calibrated using solvent signal (^{1}H-NMR, δ_H = 7.26 ppm for $CDCl_3$, δ_H = 3.31 ppm for CD_3OD, δ_H = 2.50 ppm for d-DMSO; ^{13}C-NMR, δ_C = 77.16 ppm for $CDCl_3$, 39.52 ppm for d-DMSO, δ_C = 49.00 for CD_3OD). Multiplicities are reported as follows: s, singlet; d, doublet; t, triplet; q, quartet; m, multiplet; b, broad; dd, doublet of doublets. Mass spectra were recorded by direct infusion ESI on a LCQ Fleet ion trap (Thermo Fisher Scientific, Waltham, MA, USA) mass spectrometer. For data processing, Qual Browser Thermo Xcalibur 4.0.27.13 software was used. ESI parameters for samples acquired in positive ionization mode: spray voltage 3.2 kV, capillary temperature 160 °C, capillary voltage 43 V. ESI parameters for samples acquired in negative ionization mode: spray voltage 5.0 kV, capillary temperature 160 °C, capillary voltage −8 V.

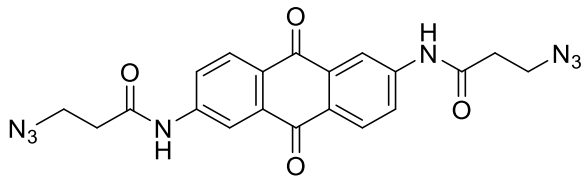
Synthesis of N,N'-(9,10-dioxo-9,10-dihydroanthracene-2,6-diyl)bis(2-azidoacetamide) (**1**)



2,6-Diaminoanthraquinone (1 equiv., 500 mg, 2.10 mmol) was dissolved in chloroacetyl chloride (24 equiv., mL, 5.70 g, 50 mmol) and the mixture was heated at 110 °C for 4 h until a colour change from violet to yellow was observed. The mixture was cooled to r.t. and diethyl ether was added to precipitate the solid product that was filtered and washed with diethyl ether. Without further purification the solid was dissolved in DMF and NaN_3 was added (6 equiv.), the mixture was stirred at r.t. for 16 h and then water was added. The resulting precipitate was filtered, washed with water and evaporated at reduced pressure to afford a yellow solid (680 mg, 80%).

^{1}H-NMR (d-DMSO, 400 MHz) δ 10.78 (s, 2H), 8.44 (d, 2H, J = 2.2 Hz), 8.19 (d, 2H, J = 8.5 Hz), 8.05 (dd, Ja = 8.5 Hz, Jb = 2.2 Hz, 2H), 4.15 (s, 4H); ^{13}C-NMR (d-DMSO, 101 MHz) δ 182.12, 170.43, 144.14, 135.03, 128.97, 128.50, 124.16, 116.42, 45.71. ESI-MS m/z: Calcd for $C_{18}H_{13}N_8O_4^+$ $(M+H)^+$: 405.10, found: 405.51.

Synthesis of N,N'-(9,10-dioxo-9,10-dihydroanthracene-2,6-diyl)bis(3-azidopropanamide) (**2**)



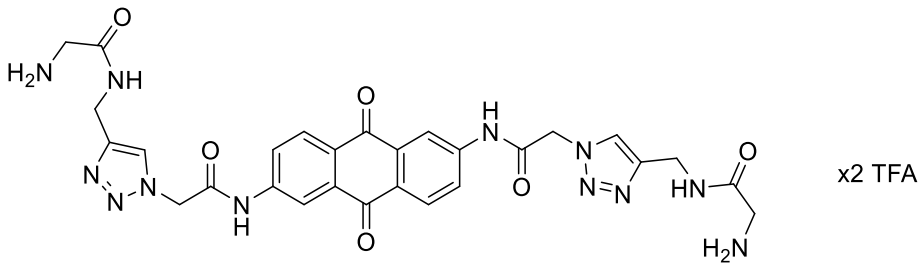
2,6-Diaminoanthraquinone (1 equiv., 200 mg, 0.84 mmol) was refluxed in 3-azidopropionyl chloride (23 equiv., 2.60 g, 19.5 mmol) at 100 °C for 4 hours. The solution changed from violet to yellow in colour. After 4 hours the mixture was cooled down to r.t. and diethyl ether (30 mL) was added to precipitate the product as a solid. The mixture was filtered under vacuum, washed with diethyl ether, and left to dry obtaining a yellow solid (258 mg, 71%).

^{1}H-NMR (d-DMSO, 400 MHz) δ 10.73 (s, 2H), 8.46 (d, 2H, J = 1.8 Hz), 8.19 (d, 2H, J = 8.5 Hz), 8.05 (dd, Ja = 8.5 Hz, Jb = 1.8 Hz, 2H), 3.66 (t, 2H, J = 6.3 Hz), 2.73 (t, 2H, J = 6.3 Hz); ^{13}C-NMR (d-DMSO, 101 MHz) δ 181.79, 170.22, 144.93, 134.87, 129.08, 128.61, 123.93, 116,31, 47.07, 36.23. ESI-MS m/z: Calcd for $C_{20}H_{17}N_8O_4^+$ $(M+H)^+$: 433.14, found: 433.62.

General synthesis of the final compounds (**3-9**)

Propargylamide amino acid derivatives were prepared as reported in literature.[346–349] Compounds **1** or **2** (1 equiv.) and the opportune propargylamide amino acid derivative (4 equiv.) were dissolved in DMSO (1 mL). In a separate vial, sodium ascorbate (1.2 equiv.), anhydrous $CuSO_4$ (0.5 equiv.) and K_2CO_3 (0.3 equiv.) were dissolved in distilled water (200 µL). The salts solution was added to the reaction mixture under an N_2 stream, stoppered, and left to stir overnight at room temperature. Water was added and the mixture left to stir for 10 min in an ice bath. The mixture was filtered using a Buchner funnel and the brown solid obtained were washed with water and ethyl acetate and dried to obtain a brown crystal solid. Without further purifications, the solid was dissolved in an equal amount of DCM and TFA, the solution was left to stir for 2 hours. The solvents were evaporated under reduced pressure and left to dry overnight in vacuum affording a crystalline solid.

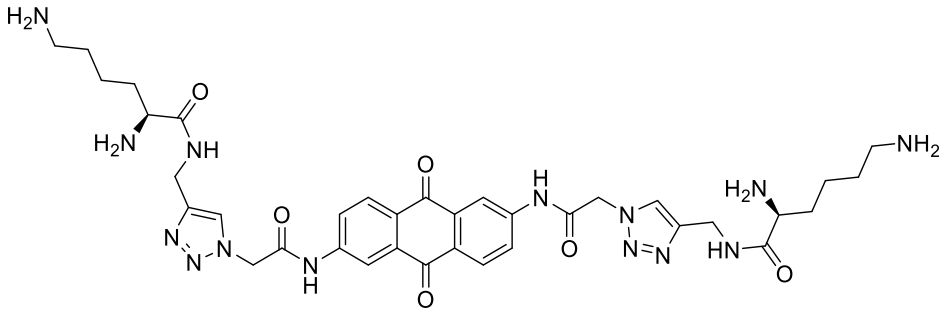
Synthesis of compound **3**



The product was obtained with the procedure described above yielding a brown solid (364 mg, 86%).

^{1}H-NMR (d-DMSO, 400 MHz) δ 11.20 (s, 2H), 8.87 (brs, 2H), 8.48 (d, 2H, J = 2.1 Hz), 8.21 (d, Ja = 8.6 Hz, 2H), 8.11-7.97 (m, 9H), 5.44 (s, 4H), 4.45 (d, 4H, J = 4.2 Hz), 3.59 (m, 4H); ^{13}C-NMR (d-DMSO, 101 MHz) δ 181.07, 165.66, 165.20, 143.76, 134.27, 128.54, 128.36, 124.36, 123.55, 115.88. ESI-MS m/z: Calcd for $C_{28}H_{29}N_{12}O_6^+$ $(M+H)^+$: 629.23, found: 629.37; Calcd for $C_{28}H_{30}N_{12}O_6^{2+}$ $(M+2H)^{2+}$: 315.12, found: 315.23.

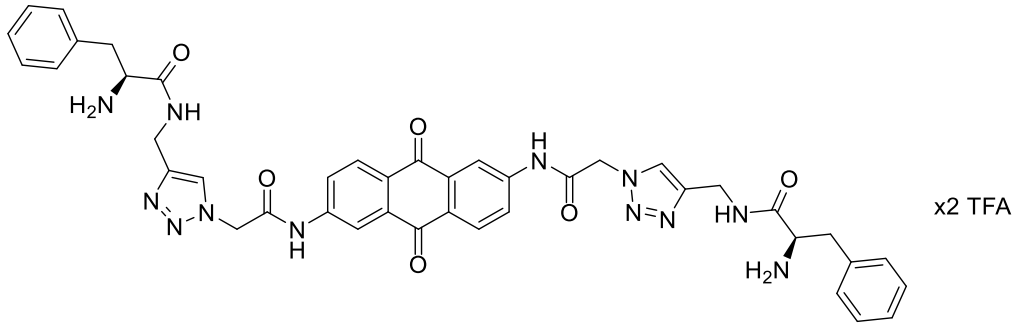
Synthesis of compound **4**



The product was obtained with the procedure described above yielding a brown solid (558 mg, 92%).

^{1}H-NMR (d-DMSO, 400 MHz) δ 11.21 (s, 2H), 8.97 (m, 2H), 8.47 (d, 2H, J = 2.1 Hz), 8.29-8.12 (m, 7H), 8.04 (s, 2H), 8.00 (dd, 2H, Ja = 8.6 Hz, Jb = 2.1 Hz), 7.73 (brs, 5H), 5.42 (s, 4H), 4.45 (m, 2H), 2.76 (m, 4H), 1.72 (m, 4H), 1.53 (m, 4H), 1.32 (m, 4H); ^{13}C-NMR (d-DMSO, 101 MHz) δ 181.46, 168.52, 165.50, 144.01, 134.56, 128.90, 128.70, 124.95, 123.99, 116.24, 52.40, 52.22, 38.69, 34.39, 30.61, 26.65, 21.28. ESI-MS m/z: Calcd for $C_{36}H_{47}N_{14}O_6^+$ $(M+H)^+$: 771.38, found: 771.57; calcd for $C_{36}H_{48}N_{14}O_6^{2+}$ $(M+2H)^{2+}$: 386.19, found: 386.41.

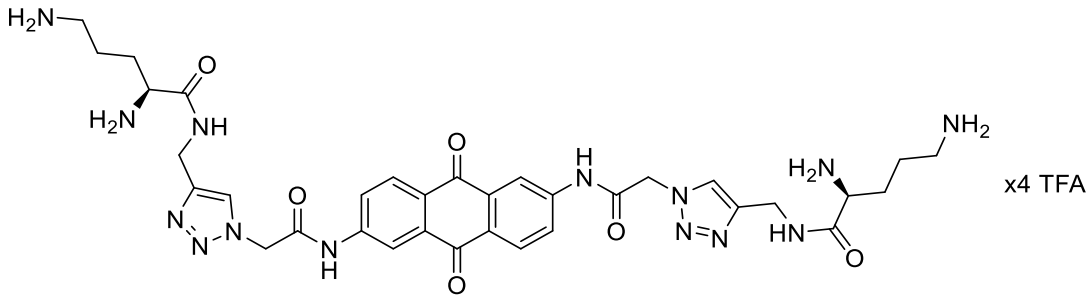
Synthesis of compound **5**



The product was obtained with the procedure described above yielding a brown solid (456 mg, 89%).

^{1}H-NMR (d-DMSO, 400 MHz) δ 11.22 (s, 2H), 8.96 (brs, 2H), 8.49 (d, 2H, J = 2.1 Hz), 8.29-8.16 (m, 7H), 8.04 (s, 2H, Ja = 8.6 Hz, Jb = 2.1 Hz), 7.39-7.11 (m, 10H), 5.43 (s, 4H), 4.40 (brs, 4H), 4.01 (m, 2H), 3.08 (m, 1H), 2.99 (m, 1H); ^{13}C-NMR (d-DMSO, 101 MHz) δ 181.08, 167.66, 165.17, 143.78, 134.69, 134.28, 129.33, 128.55, 128.38, 124.50, 123.56, 115.90, 53.34, 52.15, 36.90, 34.08. ESI-MS m/z: Calcd for $C_{42}H_{41}N_{12}O_6^+$ $(M+H)^+$: 809.33, found: 809.46; calcd for $C_{42}H_{42}N_{12}O_{62}^+$ $(M+2H)^{2+}$: 405.17, found: 405.33.

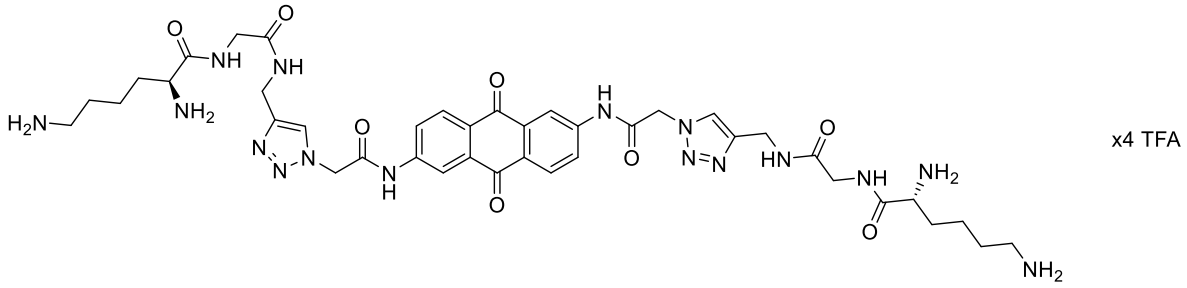
Synthesis of compound **6**



The product was obtained with the procedure described above yielding a brown solid (447 mg, 93%).

^{1}H-NMR (d-DMSO, 400 MHz) δ 11.23 (s, 2H), 9.01 (t, 2H, J = 5.2 Hz), 8.47 (d, 2H, J = 2.1 Hz), 8.33-8.16 (m, 8H), 8.07 (s, 2H), 8.02 (dd, 2H, 8.7 Hz, 2.1 Hz), 7.80 (brs, 6H), 5.44 (s, 4H), 4.52 (dd, 2H, Ja = 15.0 Hz, Jb = 5.5 Hz),), 4.39 (dd, 2H, Ja = 15.0 Hz, Jb = 5.5 Hz), 3.81 (m, 2H), 2.81 (m, 4H), 1.76 (m, 4H), 1.59 (m, 4H); ^{13}C-NMR (d-DMSO, 101 MHz) δ 181.12, 168.10, 165.28, 143.80, 134.32, 128.60, 128.41, 124.68, 123.61, 115.94, 52.17, 51.55, 38.18, 34.20, 28.06, 22.49. ESI-MS m/z: Calcd for $C_{34}H_{43}N_{14}O_6^+$ $(M+H)^+$: 743.35, found: 743.62, calcd for $C_{34}H_{44}N_{14}O_6^{2+}$ $(M+2H)^{2+}$: 372.18, found: 372.52.

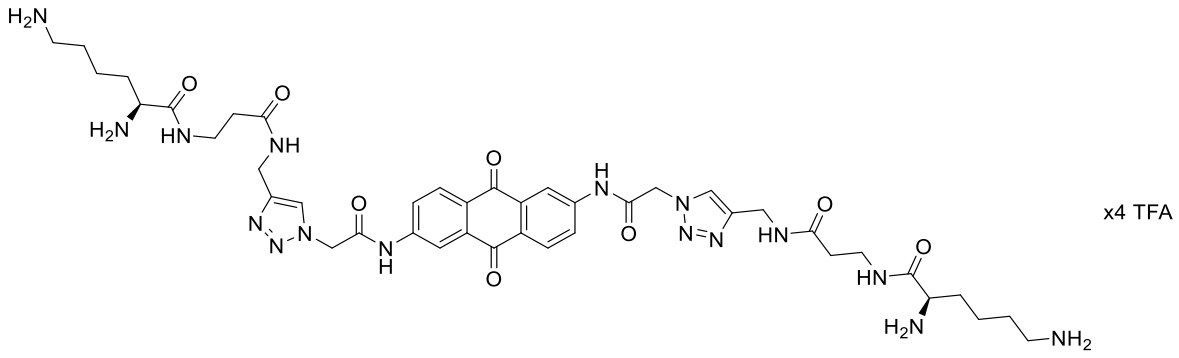
Synthesis of compound **7**



The product was obtained with the procedure described above yielding a brown solid (584 mg, 88%).

^{1}H-NMR (d-DMSO, 400 MHz) δ 11.27 (s, 2H), 8.77 (brs, 2H), 8.57 (brs, 2H), 8.47 (brs, 2H), 8.31-8.12 (m, 8H), 8.02 (brs, 4H), 7.83 (brs, 6H), 5.43 (s, 4H), 4.40 (m, 2H), 3.83 (m, 8H), 2.76 (m, 4H), 1.72 (m, 4H), 1.52 (m, 4H), 1.38 (m, 4H); ^{13}C-NMR (d-DMSO, 101 MHz) δ 181.07, 168.68, 167.98, 143.80, 134.26, 128.51, 128.35, 124.60, 123.57, 115.90, 52.12, 51.83, 41.68, 38.31, 34.00, 30.22, 26.30, 20.84. ESI-MS m/z: Calcd for $C_{40}H_{54}N_{16}O_8^{2+}$ $(M+2H)^{2+}$: 443.21 found 443.65.

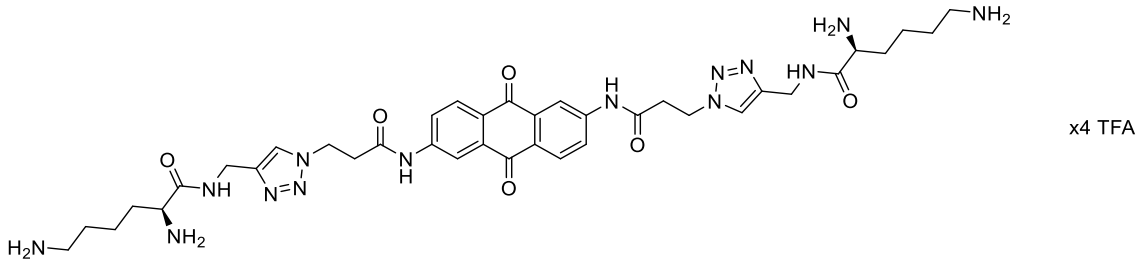
Synthesis of compound **8**



The product was obtained with the procedure described above yielding a brown solid (616 mg, 91%).

^{1}H-NMR (d-DMSO, 400 MHz) δ 11.24 (s, 2H), 8.53-8.45 (m, 5H), 8.27-8.09 (m, 8H), 8.07-7.96 (m, 4H), 7.89-7.71 (m, 6H), 5.42 (s, 4H), 4.35 (brs, 4H), 3.39 (m, 1H), 3.31 (m, 1H), 2.75 (brs, 4H), 2.35 (brs, 4H), 1.66 (m, 4H), 1.50 (m, 4H), 1.29 (m, 4H); ^{13}C-NMR (d-DMSO, 101 MHz) δ 181.69, 170.55, 168.77, 165.86, 144.41, 134.88, 129.13, 128.96, 124.18, 116.51, 52.75, 52.42, 38.93, 35.84, 34.63, 30.93, 26.99, 21.56. ESI-MS m/z: Calcd for $C_{42}H_{58}N_{16}O_8^{2+}$ $(M+2H)^{2+}$: 457.23, found: 457.55.

Synthesis of compound **9**



The product was obtained with the procedure described above yielding a brown solid (546 mg, 94%).

^{1}H-NMR (d-DMSO, 400 MHz) δ 10.81 (s, 2H), 8.93 (brs, 2H), 8.45 (d, 2H, J = 1.9 Hz), 8.25-8.09 (m, 8H), 8.06-7.96 (m, 4H), 7.83-7.71 (brs, 5H), 4.67 (t, 4H, J = 6.2 Hz), 4.38 (m, 4H), 3.73 (m, 2H), 3.11 (t, 4H, J = 6.2 Hz), 2.74 (m, 4H), 1.70 (m, 4H), 1.50 (m, 4H), 1.30 (m, 4H); ^{13}C-NMR (d-DMSO, 101 MHz) δ 181.26, 169.18, 168.25, 144.38, 134.32, 128.53, 128.07, 123.42, 115.77, 51.92, 45.10, 38.41, 36.22, 34.22, 30.36, 26.41, 21.08. ESI-MS m/z: Calcd for $C_{38}H_{52}N_{14}O_6^{2+}$ $(M+H)^{2+}$: 400.21, found: 400.60.

ESI-MS binding studies

DNA sequences were obtained from Sigma-Aldrich (Milan, Italy). Samples were heat-denatured and folded in 150 mM ammonium acetate before incubation with the ligands. Stock solutions of the compounds were prepared in methanol. The final concentration of the oligonucleotide was 5- 10 μM in 150 mM ammonium acetate, with a 10:1 compound/oligo ratio. Samples were acquired after an equilibration time of 30 minutes and methanol was added to the samples to obtain a stable ESI signal. Mass spectra were recorded by direct infusion ESI on a Thermo Fisher Scientific (Waltham, MA) LCQ Fleet ion trap mass spectrometer. The instrument was set in negative ionization mode with a 3.4 kV capillary voltage, 120 °C capillary temperature and a flow rate of 5 μL/min. Collision-induced dissociation (CID) experiments were performed on the complexes by isolating the precursor ion in the trap. The fragmentation was promoted by increasing the "normalized collision energy" parameter. For data processing, Qual Browser Thermo Xcalibur 4.0.27.13 software was used. Binding affinity (BA) was calculated according to the formula: $BA= (\Sigma DNA_{bound}/(\Sigma DNA_{free}+ \Sigma DNA_{bound})) \times 100$. Relative I (%) = $(I_{COMPLEX}/(I_{COMPLEX}+I_{DISSOCIATION\ PRODUCTS})) \times 100$.

Exact mass for the G4 sequence (5'-AGGGTTAGGGTTAGGGTTAGGGT-3'): 7270.774 Da. Exact mass for dsDNA (5'-ACTATTTACGTATAATGA-3', 5'- TCATTATACGTAAATAGT-3',): 10987.922 Da.

Fluorescence melting

Oligonucleotides were purchased lyophilized from Eurogentec (Seraing, Belgium) and used without further purification. The oligonucleotides were resuspended in water to obtain 100 μM stock solutions, which were freshly diluted in the proper buffer for further analyses. Reactions were prepared in a 96 well plates. Nucleic acid solutions (0.25 μM) were properly annealed in LiP buffer (40 mM $Li_3(PO)_4$, pH 7.5), containing 5 (bcl2) or 50 mM KCl in the presence of increasing amounts of the tested compounds (0-20 μM). For double stranded DNA solutions contained equimolar amount of the two complementary strands. Plates were read in a Roche LightCycler, using an excitation source at 488 nm and recording the fluorescence emission at 520 nm. Recordings were taken during melting and annealing (1 °C/min) to check for hysteresis. Melting temperatures (T_m) were determined from the first derivatives of the melting profiles using the Roche LightCycler software. Each data was repeated at least three times and errors were ± 0.4 °C. ΔT_m were calculated by subtracting the T_m value recorded in the presence of the ligand from the one corresponding to the free oligonucleotides.

Circular dichroism (CD)

DNA sequences were the same utilized for fluorescence melting assay. CD spectra were recorded at 25 °C in 10 mM Tris-HCl, 1 mM EDTA, 50 mM KCl pH 7.5 on a Jasco J-810 spectropolarimeter equipped with a Peltier-type temperature control system using a 10 mm path length cell. Before data acquisition, the human telomeric sequence Tel24 (4 μM) was heated at 95 °C for 5 min and left to cool at room temperature overnight. Spectra were then acquired in the absence and in the presence of increasing concentration of tested ligands. The reported spectrum of each sample represents the average of 3 scans recorded with 1 nm step resolution. Observed ellipticities were converted to mean residue ellipticity [θ] = deg x cm^2 x $dmol^{-1}$ (Mol. Ellip.). The variation of the CD signal recorded at 290 nm as a function of ligand concentration have been analysed according to a one binding site model.

Molecular docking

Atomic coordinate files were retrieved from the RCSB Protein Data Bank (PDB, www.rcsb.org): the selected 3D structure are 1KF1,[350] 143D,[67] 2HY9,[61] and 2JPZ,[62] which are all telomeric arrangements formed by the sequence [$G_3(T_2AG_3)_3$], all head and tail caps were removed resulting in a 21mer. Two central K^+ ions were placed between the three G-tetrads for the hybrid-1, hybrid-2 and antiparallel NMR resolved structures (PDB ID: 2HY9, 2JPZ and 143D), in the parallel structure (PDB ID: 1KF1) they were already present in the X-ray diffraction resolved structure. The structures were prepared with the Protein Preparation Wizard included in the Schrödinger suite,[351] using default settings, i.e. adding hydrogens, removing waters further than 5 Å from the ligand, adjusting charges, capping termini, optimizing hydrogen bond clusters. The ligands were prepared for docking with Ligprep tool under OPLS4 force field, using Epik ionizer all possible state in the 7 pH range were

generated. The docking protocol consisted of rigid-receptor/flexible ligand docking, which was performed with Glide under the OPLS4 force field,[352] with default settings (*i.e.*, a 0.80 scaling factor for the ligand atom van der Waals radius and 0.15 for the partial charge cut-off, with generation of 1 pose per ligand). For the docking to the different intramolecular apo G4s (parallel, hybrid-1, hybrid-2 and antiparallel) two cubic grids were generated, with Receptor Grid Generation app included in Schrödinger suite, selecting either the upper or the lower external quartets as the centre, the dimensions of each side of the grids were 15 Å for the internal and 35 Å for the external. The precision for the docking was set to XP.

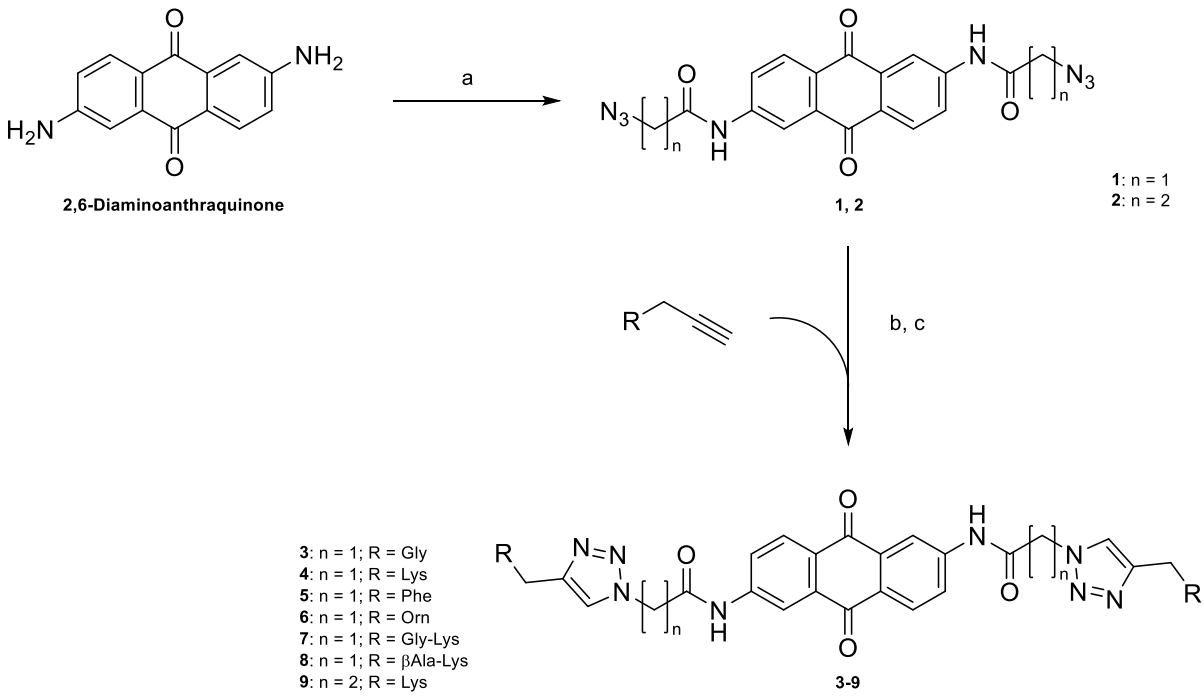
Molecular dynamics

The structures of the G4-ligand complexes, obtained by docking study, were prepared for the simulations with Schrödinger application System Builder as follows: the structures were solvated in a triclinic water box (10 Å buffer) composed by 4-site model TIP4P$_{ew}$ water molecules,[216] the total charge was neutralized by the addition of an adequate number of K^+ ions, and a 0.20 M KCl concentration was set. The system initially prepared for the simulation Schrödinger default force field OPLS4 was converted through Viparr application to the AMBER OL15 force field specific for nucleic acids,[212–214,288] the parameters set used for the ions were the ones developed by Joung and colleagues.[215,216] The ligands were parameterized with Antechamber application of AmberTools20 using GAFF2 parameters and AM1-BCC atomic charges, missing parameters were generated with Parmchk2. The simulations were conducted with Schrödinger Desmond for 200 ns each, recording 1 frame every 10 ps using the NPT ensemble at the temperature and pressure of 300 K and 1.01 bar, default settings were used for the RESPA integrator (2 fs for bonded interactions, 2 fs for near and 6 fs for far interactions), the system was relaxed prior the simulation using the default Desmond protocol. The trajectories were analysed after the simulations with Simulation Event Analysis Schrödinger application, for the over-time ligand RMSD calculations the ligand atoms positions were fitted to the guanine planes.

4.3 Results and discussion

Chemistry

In this study, a series of 2,6-bis-substituted anthraquinone-amino acid conjugates (2,6-AQs) were designed and prepared following a convergent synthetic scheme (Scheme 4.1). The synthesis included the preparation of two 2,6-AQ diamido-azide derivatives, serving as central scaffolds, and of 6 propargylamide derivatives of 4 single amino acids (L-glycine, phenylalanine, lysine, ornithine) and 2 dipeptides (Gly-Lys, *β*-Ala-Lys) as side chains.

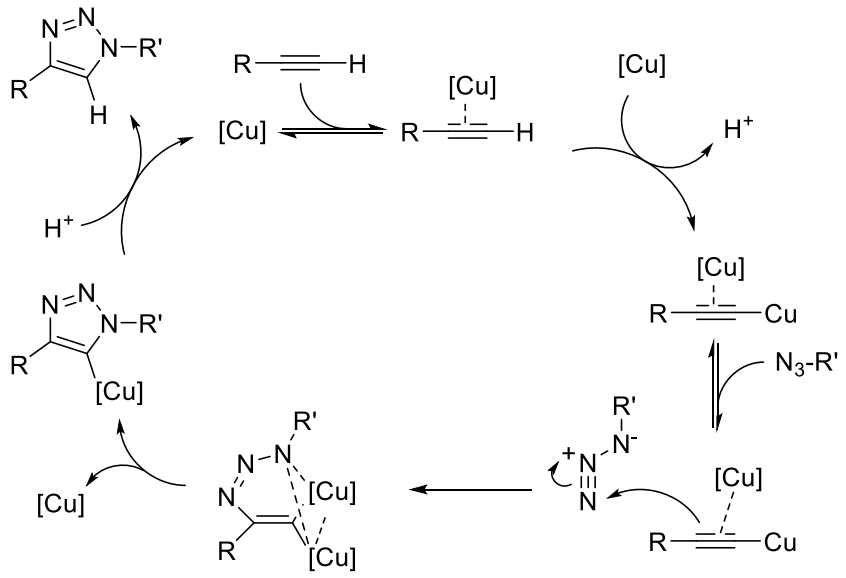


Scheme 4.1. General scheme for the synthesis of the set of 2,6-bis-substituted anthraquinone-amino acid conjugates **3**-**9**. a: chloro acetyl chloride, NaN_3, DMF for 1, 3-azido propionyl chloride for 2; b: sodium ascorbate, CuSO4, K2CO3, DMSO, water; c: TFA, DCM.

More in detail, the choice of the amino acids was based on the variation of two parameters, which are crucial for the interaction with the anion-phosphate backbone within the G4 grooves, namely the overall length of the side chains and the number of cationic charges resulting from the presence of 1 or 2 primary amines. A final CuAAC reaction allowed the combination of the scaffold with the side chains, yielding the desired final compounds (**3**-**9**) after removal of the protecting groups.

More specifically, the synthesis of **1**, namely the bis(2-azidoacetamide) 2,6-AQ derivative, comprehended di-amidation of the 2,6-diaminoanthraquinone which was carried out using an excess of chloroacetyl chloride. The resulting intermediate was then directly subjected to a nucleophilic substitution reaction with NaN_3. The bis(3-azido propanamide) 2,6-AQ derivative (**2**) was prepared

similarly, as the 2,6-diaminoanthraquinone was directly reacted with 3-azidopropanoyl chloride prepared from 3-bromopropanoic acid.[353] The amino acid or dipeptide propargylamide derivatives were obtained by liquid phase coupling procedures known in the literature from the mono- or di-BOC-protected amino acids glycine, phenylalanine, lysine and ornithine.[349,354–356] Eventually, the final compounds **3**-**9** were obtained through the Copper-Catalyzed Azide-Alkyne Cycloaddition (CuAAC) reaction between the central scaffolds and the side chains which was performed using a catalytic amount a pre-formed Cu(I) catalyst, obtained by reduction in water of Cu(II)SO_4 with sodium ascorbate. Moreover, alkaline conditions were assured by the presence of a sub-stoichiometric amount of K_2CO_3. In fact, neutral conditions did not allow the reaction to proceed, probably due to required deprotonation step of the alkyne C-H, which is crucial for the catalytic cycle, a model for the mechanism of the CuAAC, proposed by Fokin in 2013 is reported in the scheme below (Scheme 4.2).[332,357,358]



Scheme 4.2. Mechanism of the catalytic cycle composing the CuAAC reaction between a terminal alkyne and an azide, the scheme includes the alkyne deprotonation and formation of a reactive copper acetylide which reacts with the triple bond in a [4+2] fashion that generates the aromatic triazole ring.[358]

ESI-MS binding studies

ESI-MS binding studies were performed to assess the ability of compounds **3**-**9** to interact with a telomeric G4. Native ESI-MS represents a valuable screening tool for drug discovery and, in particular, for studying ligand-G4 complexes. In fact, under specific experimental conditions, the G4 arrangement is retained in the gas phase. Moreover, the reliability of this ligand screening method has been previously assessed by correlating the results with those derived from other experimental techniques.[145,147,149,359,360] Compounds **3**-**9** were tested against a G4 23-mer sequence containing the telomeric TTAGGG repeat and the selectivity for G4 was investigated using a dsDNA

oligonucleotide as control. In both instances, solutions containing 10:1 molar ratio of ligand:DNA were analysed. By assuming that the electrospray response factor of the DNA is constant upon ligand binding, the relative intensities (I) of free (unbound DNA) or bound DNA in the mass spectra should be proportional to their concentration in solution. Therefore, to evaluate the magnitude of the binding event in the MS experiments, the binding affinity (BA) parameter was determined as the fraction of bound DNA.

While no interaction with G4 was detected for **3** and **5**, all the other compounds showed a marked selectivity for G4 over dsDNA, since no binding with the latter was measured (see Figures S22, S23 and S25 in the Supporting Information for representative ESI-MS spectra). Furthermore, under our experimental conditions only complexes corresponding to 1:1 (and no 2:1) stoichiometry were observed. In terms of BA, compounds **7** (42.5) and **4** (37.1) were highlighted as the most promising ligands. As expected, longer and more positively charged amino acids, such as Lys, possess a greater aptitude to interact with G4, probably due to an increased ability to intercept a higher number of negative charges present in the nucleic acid grooves.

In order to investigate the efficiency of ligands in stabilizing the macromolecular arrangement, Collision-Induced Dissociation (CID) experiments were also performed.[359] In particular, the relative gas-phase kinetic stability was investigated by subjecting the complexes to increasing collisional energies, which upon conversion to internal energy causes bonds breakage, molecular ion fragmentation into smaller fragments and ligand loss (see Figure S24 in the Supporting Information for representative CID spectrum). In these experiments, the estimated parameter is $E_{COM}^{50\%}$, which is expressed in eV and represents the collision energy needed to promote the dissociation of the complex to its relative half-intensity (Figure 4.1) In terms of $E_{COM}^{50\%}$, ligands **8** (46.2 eV) and **9** (41.8 eV), bearing *β*-Ala-Lys and Lys (n = 2) as side chains, showed the highest values. Limited stability was observed for compound **6**, and this observation is in accordance with the low BA value detected (Table 4.1).

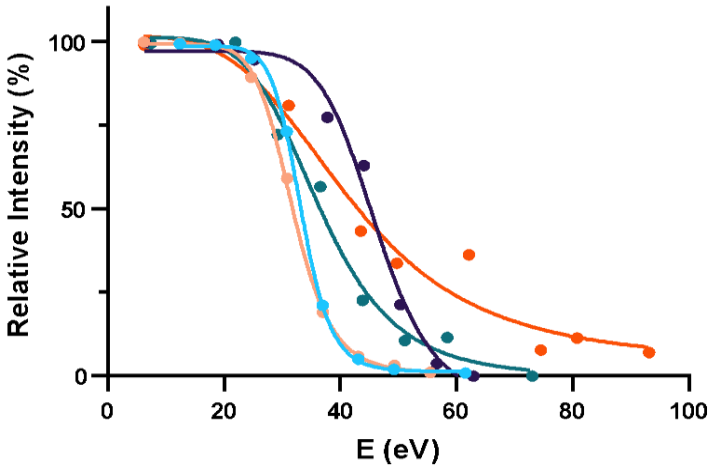


Figure 4.1. Plot of relative intensity of the complex G4-ligand against collision energy (eV) applied for the different compounds: • **4**; • **6**; • **7**; • **8**; • **9**.

Table 4.1. BA and $E_{COM}^{50\%}$ (eV) values determined by ESI-MS interaction assays between compounds **3**-**9** and the 23-mer G4 telomeric or double stranded DNA sequences.

Compounds	BA G4	$E_{COM}^{50\%}$ 1:1 (eV)	BA dsDNA
3	-	-	-
4	37.1	33.2	-
5	-	-	-
6	14.5	31.8	-
7	42.5	36.4	-
8	30.0	46.2	-
9	22.8	41.8	-

Fluorescence melting

In fluorescence melting analyses, target DNA sequences are properly designed to bear a fluorophore (6-FAM) and a quencher (Dabcyl) in close proximity when the nucleic acid is folded. This results in the quench of FAM fluorescence signal. Conversely, upon DNA denaturation, the two labelling groups fall apart and an enhancement of fluorescence is observed.[361,362] The increase in melting temperature (ΔT_m) is measured upon ligand addition to estimate the stabilization induced by the formation of the complex. This experiment was performed using different G4 sequences, namely Tel23, Tel24, kit* and bcl2, and a dsDNA sequence as control. The resulting ligand stabilization data

are summarized in Figure 4.2: tested compounds stabilize G4s with different efficiencies and, consequently, can be divided into three main groups. Overall, compounds **4**, **7** and **9** are the best stabilizers, **6** and **8** showed intermediate efficiency while **3** and **5** have the lowest stabilizing effect. This observation agreed with ESI-MS experiments, which highlighted compounds **4**, **7**, **8** and **9** as good binders and suggested that **3** and **5** do not interact with G4. Importantly, by fluorescence melting we address that this ranking order was not significantly affected by the preferred topology assumed by the oligonucleotides in our working conditions. Nevertheless, as expected, the maximal ΔT_m detected can vary according to the sequences and the compounds seems to best perform in the presence of DNA arranged into parallel G4 (*i.e.* Tel24 in presence of 20% PEG or bcl2). It should be underlined that, whether our compounds extensively increase the stability of the tested G4s, an excess of ligand was always required. Under these conditions it was demanding to monitor their behaviour towards the canonical double helix. When tested in the presence of a dsDNA, the tested compounds were similarly ranked. Noteworthy, the stabilizing effect was much more reduced and only compound **7** significantly incremented the dsDNA melting temperature at concentration in the low micromolar range (≈ 4 °C at 5 μM). This points to a preferential recognition of the G4 by compounds **4** and **9**.

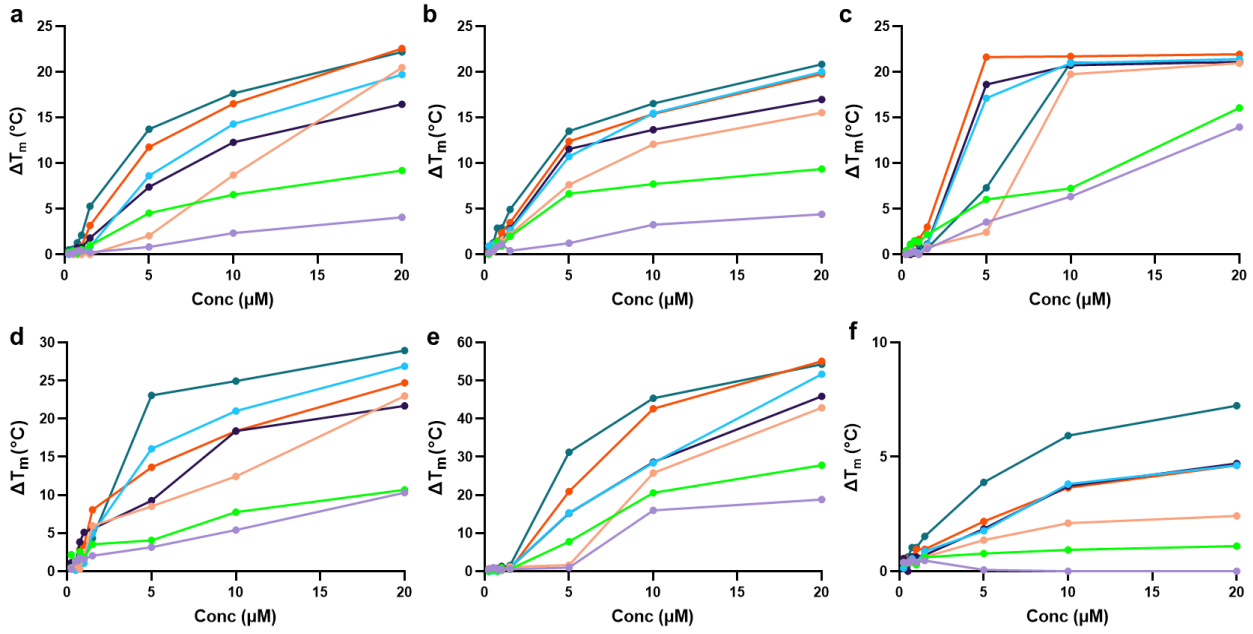


Figure 4.2. Values of ΔT_m of different DNA sequences with increasing concentration of ligands: Tel23 (50 mM of K^+, T_m DNA = 57.6 °C) (a), Tel24 (50 mM of K^+, T_m DNA = 60.4 °C) (b), Tel24 PEG 20% (5 mM of K^+, T_m DNA = 68.9 °C) (c), kit*(F53) (50 mM of K^+, T_m DNA = 58.5 °C) (d), bcl2 (F10) (5 mM of K^+, T_m = 35.1 °C) (e), dsDNA (T_m DNA = 64.1 °C) (f). Different colours indicate different ligands: • **3**; • **4**; • **5**; • **6**; • **7**; • **8**; • **9**.

Circular dichroism (CD)

Based on data reported above, derivatives **4** and **9** were considered to evaluate in more details the G4 recognition process. At first, we assessed if the environmental conditions impact on the G4-ligand binding. For this goal, we used the target sequence Tel23 that folds into the hybrid-2 topology. Our data showed that in buffer solution containing either K^+ or NH_4^+, the addition of the ligand forces G4 to a preferential antiparallel arrangement as evidenced by the increment of the positive contribution at 290 nm. As evidenced in the following figure (Figure 4.3), the shape of the final complex is comparable in all conditions with both tested compounds, thus suggesting the formation of complexes with similar arrangements also under the conditions used for MS experiments which involve the ammonium acetate buffer. Noteworthy, similar behaviour was monitored also with Tel24 that folds into hybrid-2 topology and show a stronger parallel contribution.

The variation of the CD signal was eventually used to derive the dissociation constant of the binding process, which are reported in Table 4.2. Similar values, in the range of ≈ 10 μM, were observed both in KCl and ammonium acetate buffers. However, it must be pointed out that modest changes in CD signals were observed using the latter, thus leading to a reduced reliability of data analyses.

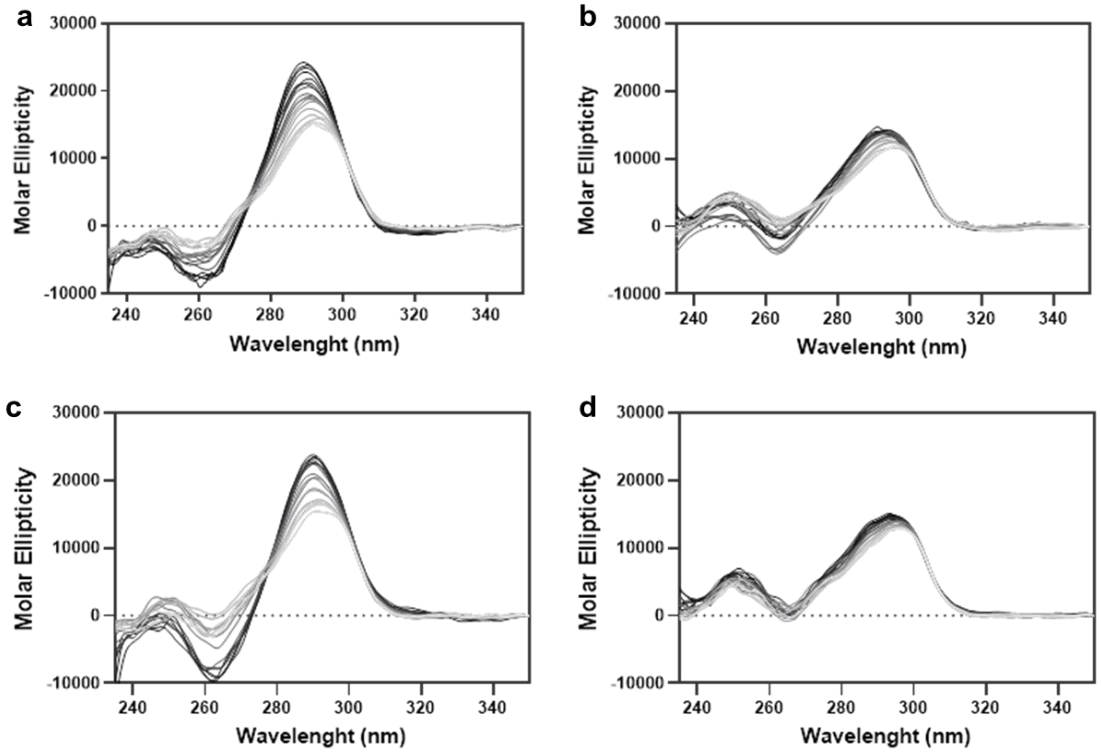


Figure 4.3. CD spectra of 4 μM Tel24 in 10 mM Tris-HCl (pH 7.4), 50 mM KCl at increasing ligand/Tel24 molar ratios (the darker the line colour the higher molar ratio); **4** in K^+ (a), **4** in NH_4^+ (b), **9** in K^+ (c) and **9** in NH_4^+ (d).

Table 4.2. Equilibrium dissociation constants obtained from CD experiments performed in 10 mM Tris-HCl, 1 mM EDTA, at 1:1 ligand/DNA molar ratio.

Sequence	Buffer	Ligand	K_d (μM)
Tel24	50 mM KCl	**9**	6.6 ± 1.2
Tel23	50 mM KCl	**9**	8.8 ± 1.4
Tel23	50 mM KCl	**4**	13.64 ± 1.2
Tel23	150 mM NH_4Ac	**9**	12.04 ± 2.26
Tel23	150 mM NH_4Ac	**4**	9.91 ± 1.93

Molecular docking

In order to investigate the binding mode of the 2,6-AQ most promising derivatives with the telomeric G4 arrangement, a computational study was performed. In particular, compounds **4** and **9** were docked to four G4 structures, one for each topology, namely parallel, antiparallel, hybrid-1 and hybrid-2 (PDB IDs: 1KF1,[350] 143D,[67] 2HY9,[61] 2JPZ,[62] respectively), following sequence preparation and docking protocol of a previous work.[193] The docking-study was performed with Schrödinger Glide,[363] using the telomeric G4 structures as receptors.

According to the predicted models for the parallel arrangement, which is the prevalent topology in the cellular crowded environment,[72,200] 2,6-AQs interact with G4 by quasi-external stacking on the upper guanine tetrad, elongating the side chains in the grooves. The models obtained for compounds **4** and **9** are depicted in the pictures reported below (Figure 4.4a and 4.4b). More specifically, pi-pi stacking interactions between the AQs and the guanines, electrostatic interactions between the charged amines of the side chains and the phosphate anions in the grooves and H-bond networks were detected in the interaction pattern. Hybrid and antiparallel G4s showed a more complex binding mode, probably due to the rigidity of the receptor in the docking procedure. According to CD experiments, these ligands are able to induce a conformational change of the G4 arrangement (from hybrid to anti-parallel) that can't be taken in account with molecular docking, therefore the poses obtained for ligand **4** (Figure 4.4c, 4.4e, 4.4g) and **9** (Figure 4.4d, 4.4f, 4.4h), may represent a starting point for a subsequent rearrangement. The binding energies obtained in the computation and the poses obtained for the ligands **4** and **9** are reported in the following table (Table 4.3).

Table 4.3. Calculated binding energy (kcal/mol) for the ligands **4** and **9** obtained by molecular docking.

Ligand	Parallel G4 (PDB ID: 1KF1)	Antiparallel G4 (PDB ID: 143D)	Hybrid 1 (PDB ID: 2HY9)	Hybrid 2 (PDB ID: 2JPZ)
4	-12.1	-8.9	-11.9	-9.6
9	-12.8	-9.0	-11.5	-10.3

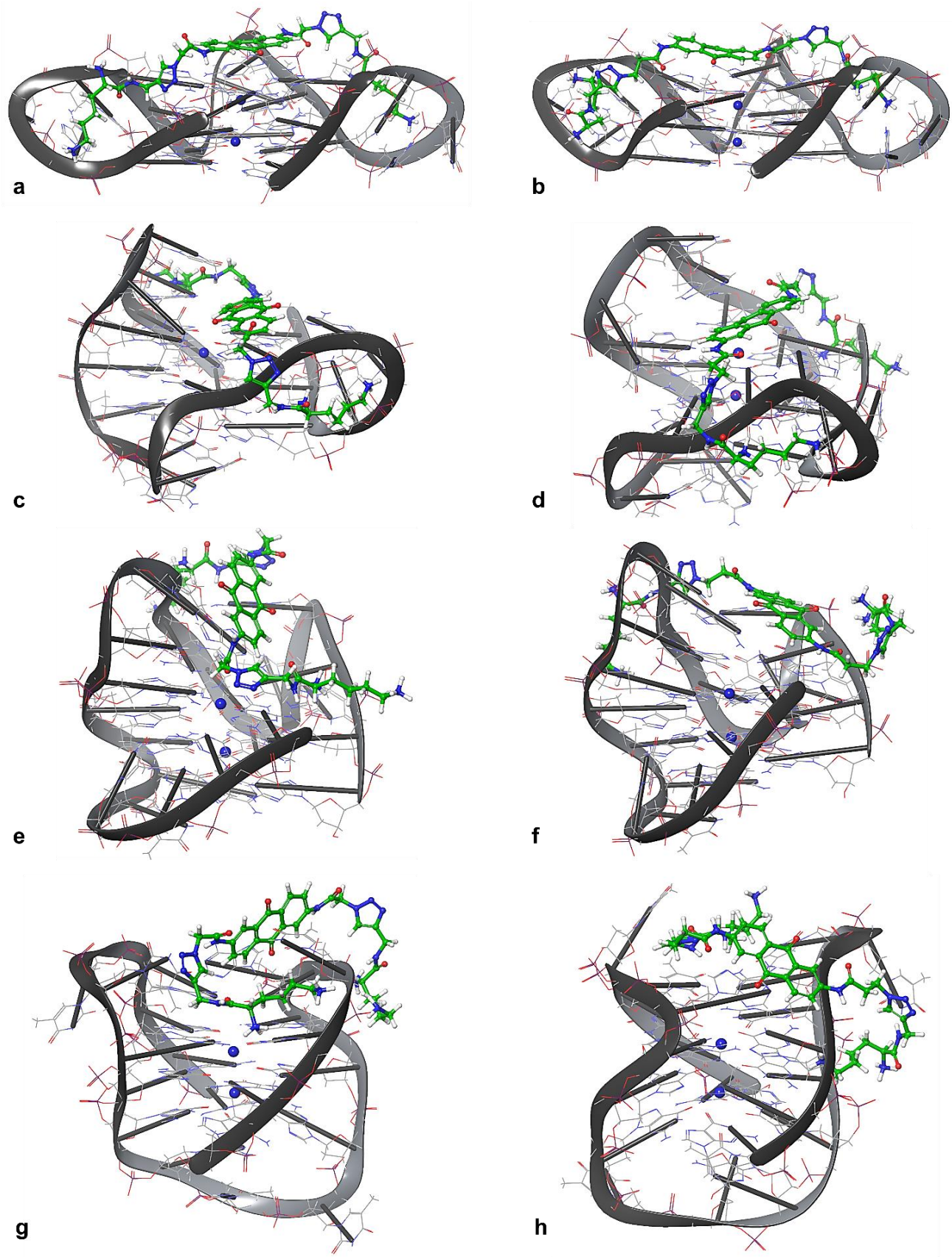


Figure 4.4. Predicted binding modes for the ligands with the four G4 topologies. Ligand **4**: parallel (PDB ID: 1KF1) (a), hybrid-1 (PDB ID: 2HY9) (c), hybrid-2 (PDB ID: 2JPZ) (e), antiparallel (PDB ID: 143D) (g). Ligand **9**: parallel (PDB ID: 1KF1) (b), hybrid-1 (PDB ID: 2HY9) (d), hybrid-2 (PDB ID: 2JPZ) (f), antiparallel (PDB ID: 143D) (h).

Molecular dynamics

To better investigate the ability of the 2,6-AQs ligands to interact with the G4 motif, a series of MD calculations were conducted for the best ligands **4** and **9** in complex with the four different G4 topologies already explored by molecular docking.

Prior to the runs with the ligands, the apo forms of the G4 receptors were also subjected to MD simulations of 200 ns each, in order to assess their stability under the simulation conditions, namely the OL15 force field specific for simulation of DNA arrangements,[212,213,216,288] the TIP4P$_{ew}$ four sites water model and the JC ions parameters (see for details the Materials and methods section). All the G4s demonstrated a good stability under the simulation conditions (basing on the G-tetrads' guanines RMSDs) (Figure 4.5) and complete retention of the central K^+ ions (manually placed in the starting structure for the non-parallel G4s as described in materials and methods section).

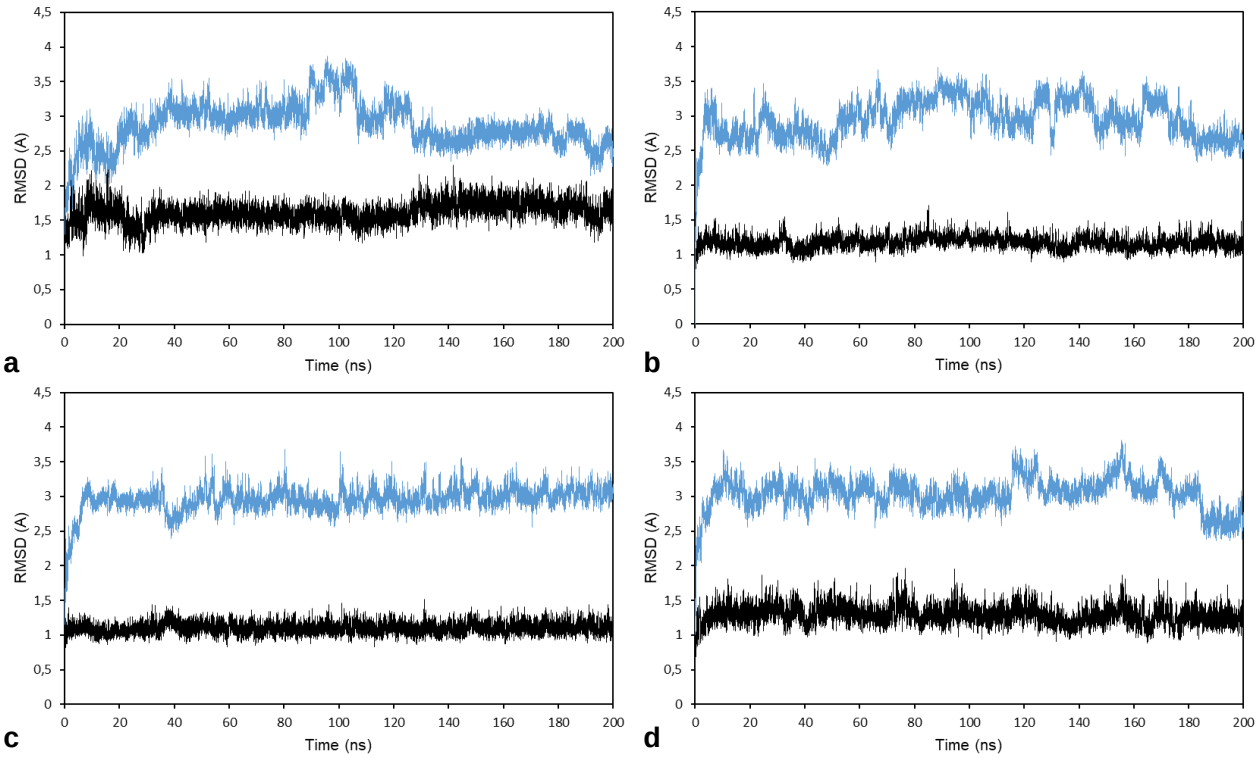


Figure 4.5. Graphs showings the over-time RMSD values for the whole G4 structure (blue) and for the guanines (black) obtained from the 200 ns MD simulation trajectory of G4 apo-forms of the parallel (a), the hybrid-1 (b), the hybrid-2 (c) and the antiparallel (d) topology.

After the assessment of the stability of the apo structures under the simulation's conditions, the *holo* forms were simulated using the same methodology and parameters. In particular, a total of 8 poses obtained in the docking stage, corresponding to 4 topologies (PDB IDs: 1KF1, 2HY9, 2JPZ and 143D) for each of the two ligands **4** and **9**, were used as the starting structures for MD simulations of 200 ns each. For all the simulations similar features were observed, namely an overall stability of

the G4 structure (basing on the G-tetrads' guanines RMSDs) (Figure 4.6 black trace), the motion of the ligands' anthraquinone core from the docked position due to the reaching of a lower energy position and/or due to G4-ligand complex mutual rearrangement (Figure 4.6, red trace) and a high mobility of the ligands' side chains (Figure 4.6, grey trace). Before going more in the detail about the description of the simulations, it has to be pointed out that the length and the complexity of that two ligands' side chains, greatly influence the whole-ligand RMSD parameter, for this reason the central anthraquinone core RMSD parameter has been computed and will be also used to establish the ligands' position and stability.

For the parallel topology, ligands **4** and **9** mostly maintained the initial positions of their tricyclic cores on top of the G-quartet, for **9** a small movement towards the G-quartet center was observed as evidenced by a variation in the trajectory at 70 ns (Figure 4.6, b), moreover the initial opposite-grooves hugging interaction of the side chains was completely conserved for **9** whereas for **4** a higher mobility was observed as showed by the last simulation frame (Figure 4.7, b). For what concerns the interaction with the hybrid-1 G4, ligand **4** reached its position of stability in about 50 ns (Figure 4.6, c) which not consisted in a proper stacking on the G-quartet but instead comprised the stacking of the anthraquinone core on an adenine of one loop and of a triazole on a guanine on the opposite loop. On the other hand, ligand **9** reached in the last 40 ns a stable interaction with the G4 (Figure 4.6, d) in particular by stacking on G-quartet and embracing with its side chains one of the loops (Figure 4.7, d). The interaction of the two ligands with the hybrid-2 topology included a rearrangement of the G4 receptor. In particular, two synchronic events were observed, namely the retreat and the lifting of two adenines covering the tetrad, and the reaching of the G-quartet stacking position by the ligands (Figure 4.7, e and f). Lastly, with the antiparallel topology, both the ligands showed a similar behaviour producing a final binding mode consisting in a stable stacking interaction (Figure 4.7, g and h) on the top of a mixed adenine/thymine quartet, which arouse in the first ns of the simulation from uncoupled bases, moreover the side chains embraced a lateral loop with low level of fluctuation (Figure 4.6, g and h). On a final note to the trajectories obtained, it is worth to say that from a close observation of the different frames, an explanation to the significant mobility of the side chains, can be derived from the propension of these latter to disengage from the interaction with the G4 to interact with the waters of the solvent. This particular behaviour is strictly correlated with the high polarity of the side chains, but could be mitigated in a second generation of the ligands using tertiary amines instead of primary ones decreasing therefore the number of H-bond donors.

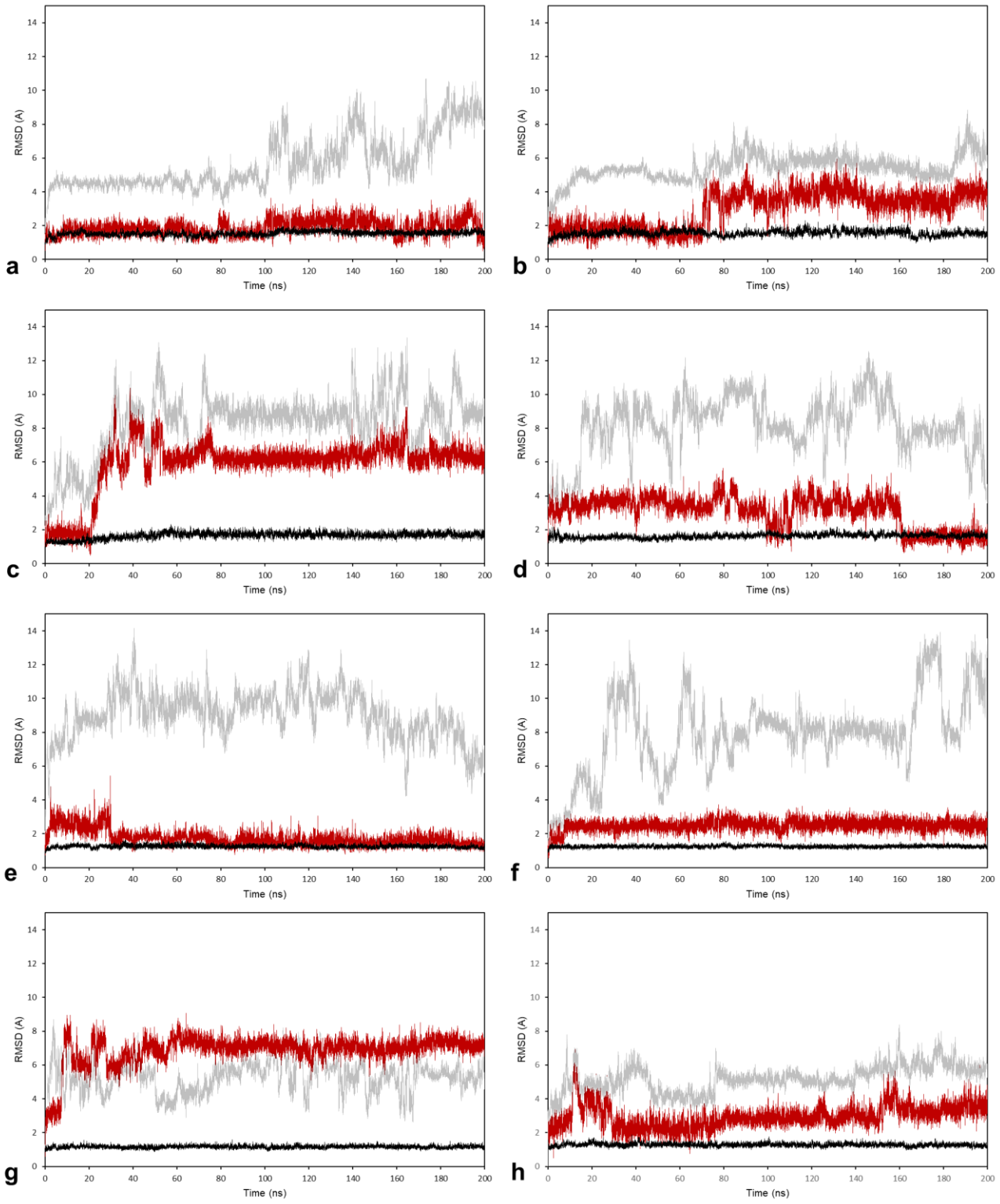


Figure 4.6. Graphs showings the over-time RMSD values for the guanines (black), the ligands anthraquinone core (red) and of the whole ligands (light-grey) obtained from the 200 ns MD simulation trajectories of the G4-ligands **4** and **9** complexes (*holo* forms) of the parallel (a), the hybrid-1 (b), the hybrid-2 (c) and the antiparallel (d) topology.

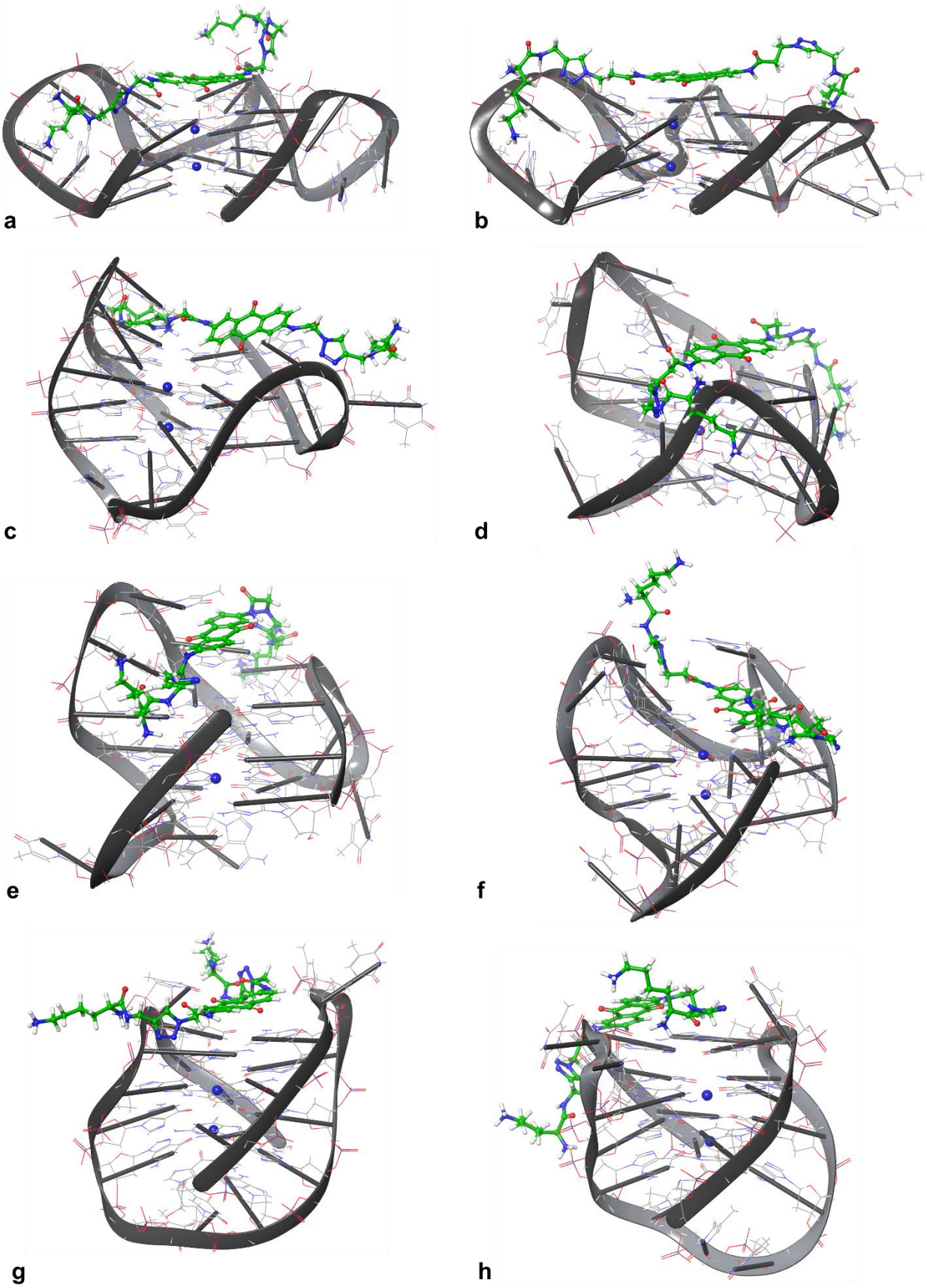


Figure 4.7. Last frame of the MD simulations of the ligand-G4 complexes. Ligand **4**: parallel (PDB ID: 1KF1) (a), hybrid-1 (PDB ID: 2HY9) (c), hybrid-2 (PDB ID: 2JPZ) (e), antiparallel (PDB ID: 143D) (g). Ligand **9**: parallel (PDB ID: 1KF1) (b), hybrid-1 (PDB ID: 2HY9) (d), hybrid-2 (PDB ID: 2JPZ) (f), antiparallel (PDB ID: 143D) (h).

4.5 Conclusion

The identification of novel, selective G4 ligands is an attractive aim for developing antiproliferative and antiviral agents. In this work, we described the preparation of a novel class of AQ-based G4 ligands. The adopted synthetic scheme is flexible and allows the prompt derivatization of the planar core with a variety of peptide side chains or other moieties thanks to the click chemistry approach. Moreover, the presence of the resulting triazole ring further extends the aromatic region and may be used to optimize ligand-target recognition.

Overall, the compounds able to interact with G4 sequences showed marked selectivity over dsDNA. Most importantly, interaction data retrieved from ESI-MS and spectroscopic studies showed the same trend and consistently highlighted the same pool of most promising binders. The best ligands **4** and **9** demonstrated the same activity towards the Tel23 sequence, which, upon titration, changed its starting hybrid topology to an antiparallel, moreover it must be noted that high ΔT_m values were observed in the thermal stabilization study upon ligands binding to bcl2 sequences.

The final computational experiments conducted for the two best ligands showed a good ability of the ligands to develop stable complexes with the different possible topologies and in particular with the hybrid-2 and the antiparallel ones which appears consistent with the *in vitro* findings.

Taken together, these results pave the way for a further investigation of 2,6-AQ derivatives as a novel class of synthetically accessible, efficient and selective G4 ligands.

4.6 Supporting material

*Characterization of the final compounds **3**-**9***

ESI-MS binding study

Sequences used for fluorescence melting and circular dichroism

*Characterization of the final compounds **3**-**9***

Compound **3**

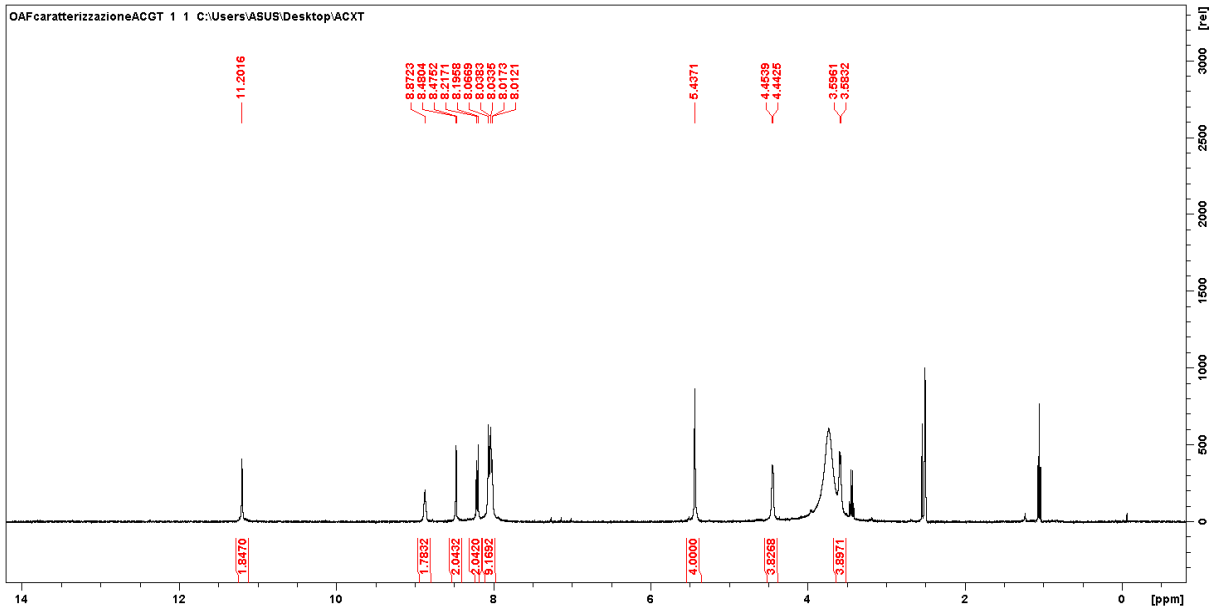


Figure S4.1. ^{1}H-NMR spectrum of compound **3**.

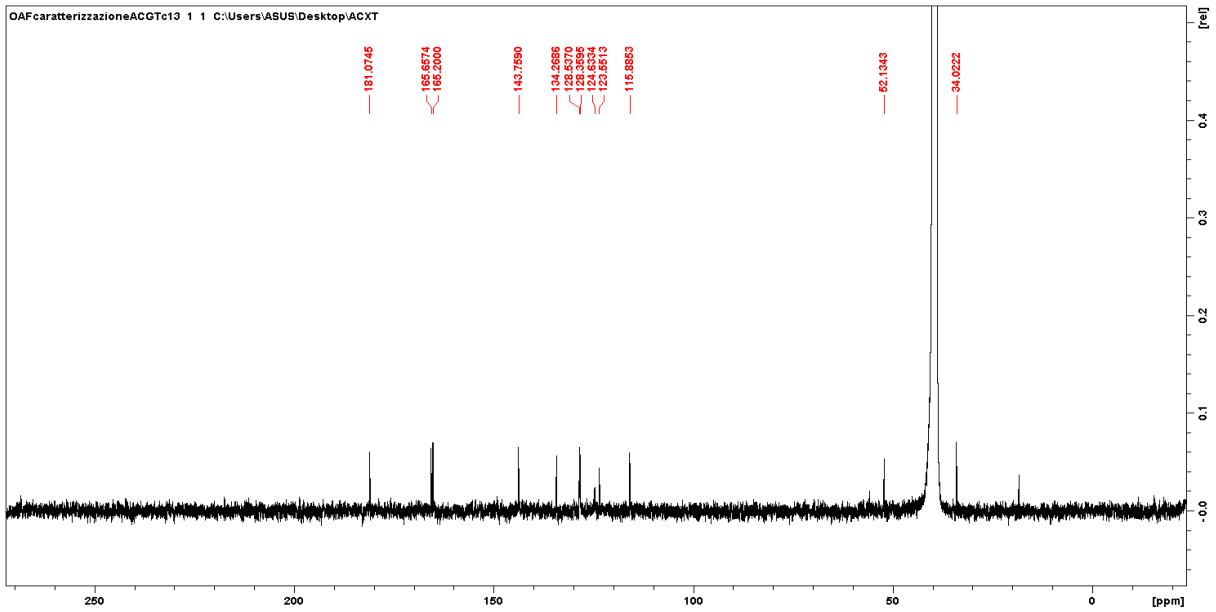


Figure S4.2. ^{13}C-NMR spectrum of compound **3**.

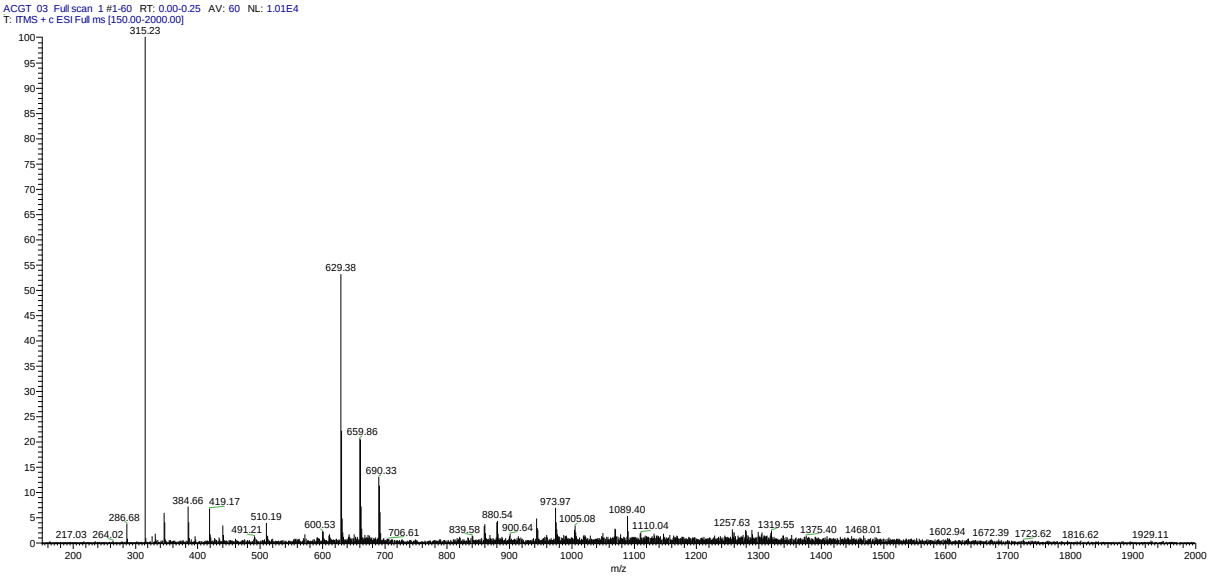


Figure S4.3. ESI-MS spectrum of compound **3**.

Compound **4**

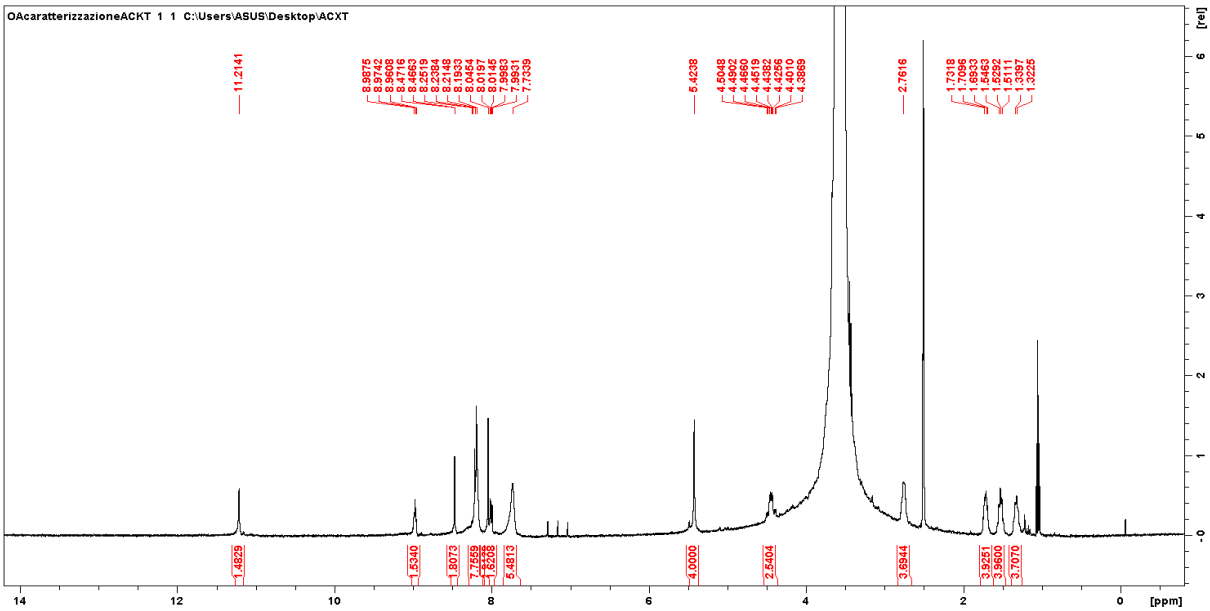


Figure S4.4. ^{1}H-NMR spectrum of compound **4**.

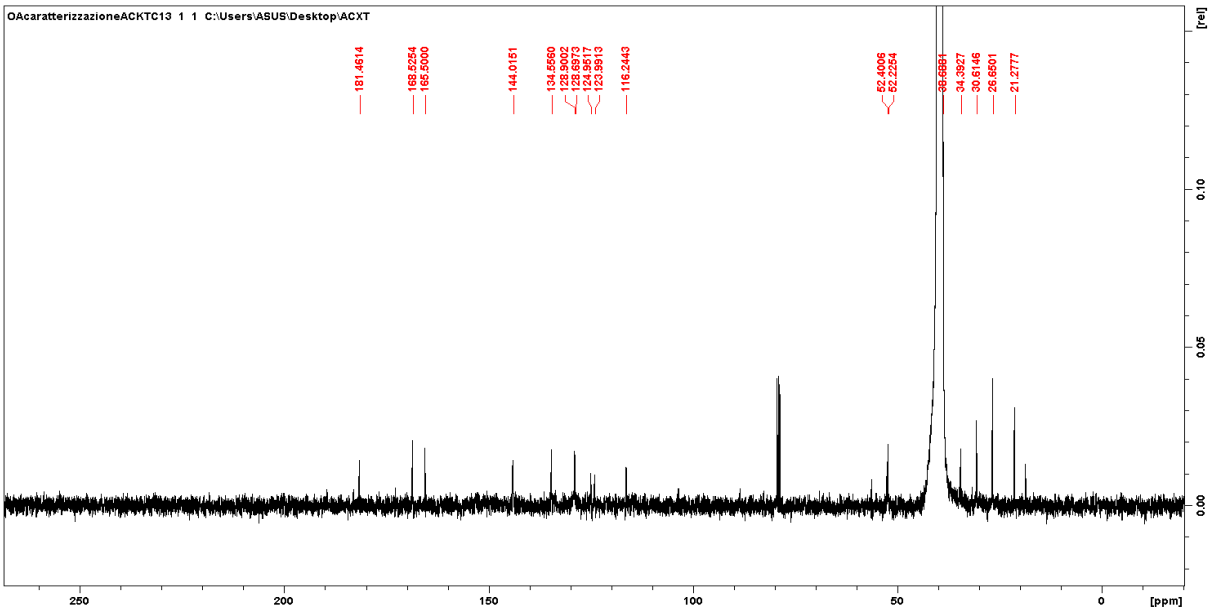


Figure S4.5. ^{13}C-NMR spectrum of compound **4**.

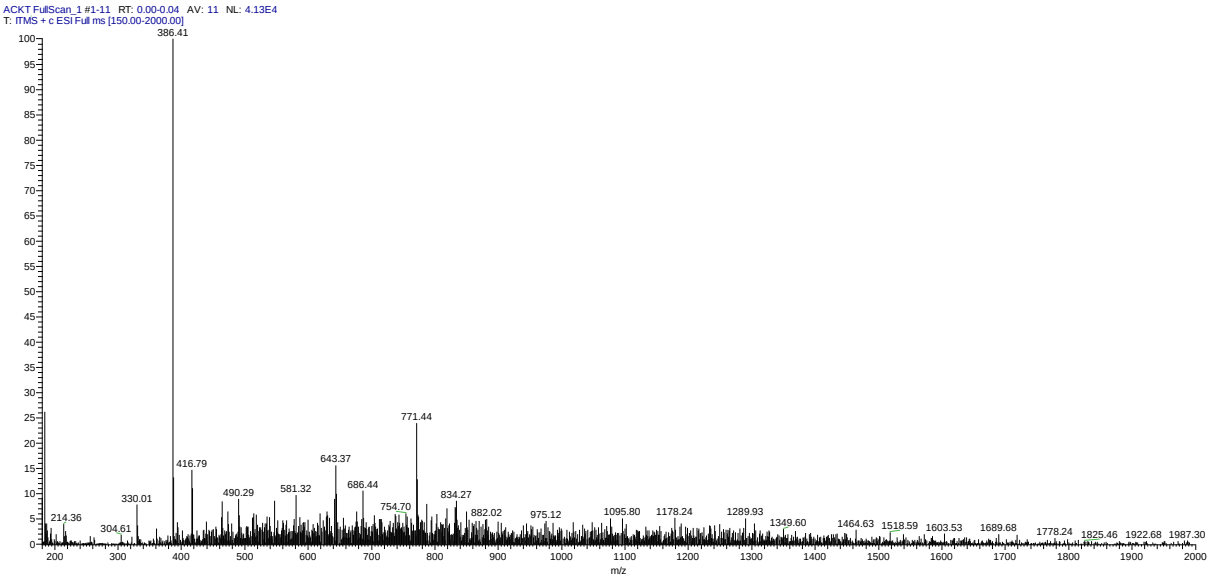


Figure S4.6. ESI-MS spectrum of compound **4**.

Compound **5**

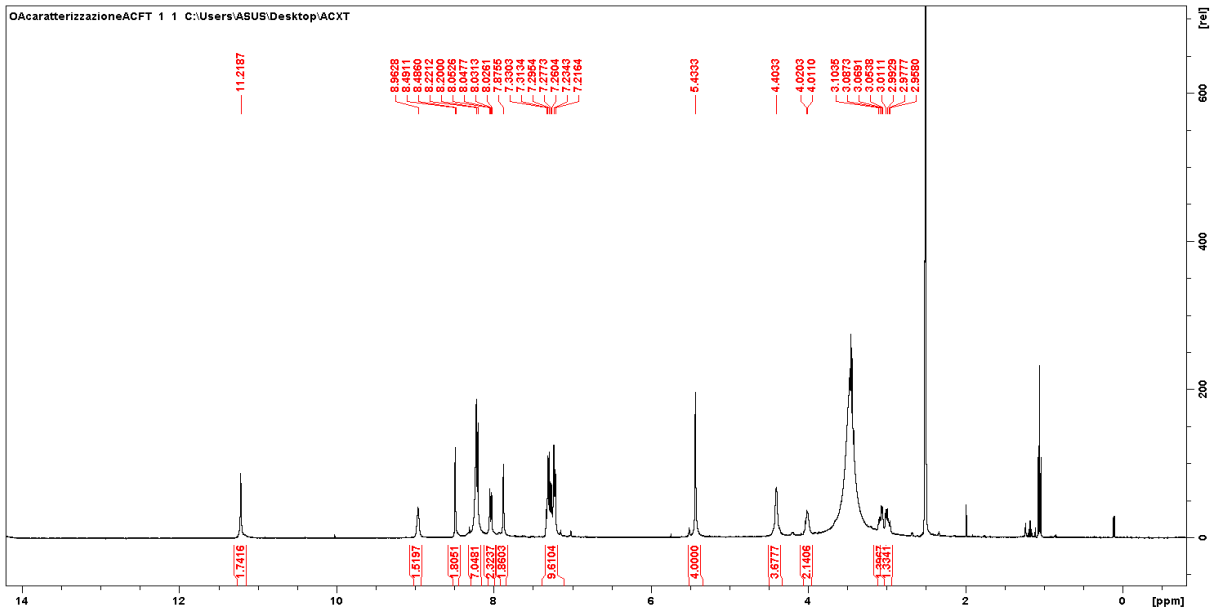


Figure S4.7. ^{1}H-NMR spectrum of compound **5**.

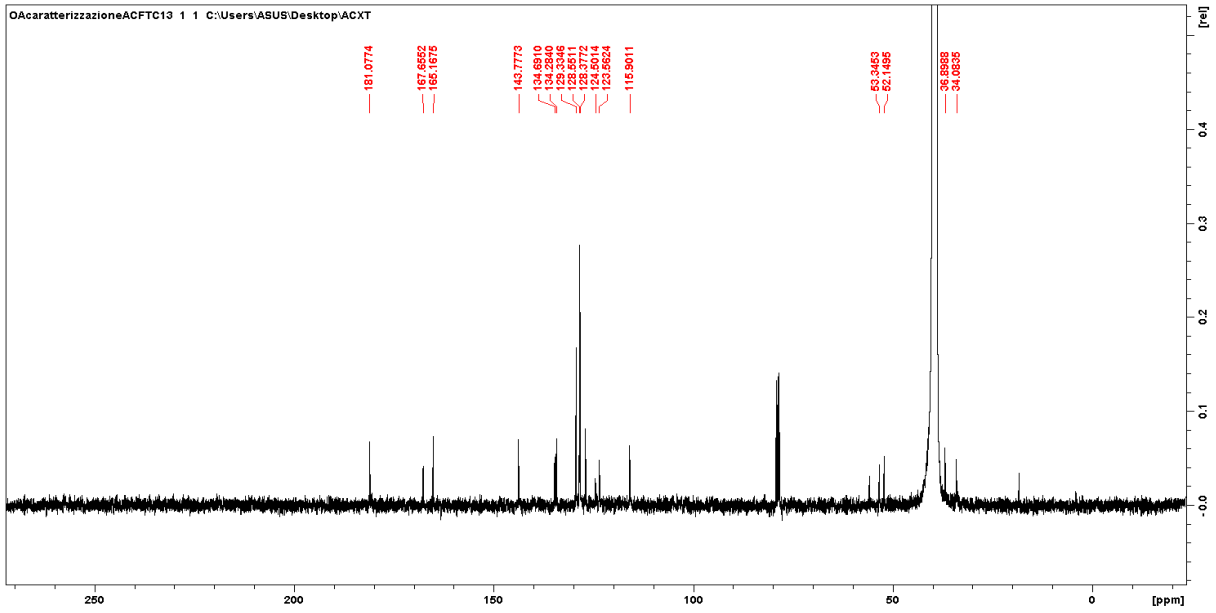


Figure S4.8. ^{13}C-NMR spectrum of compound **5**.

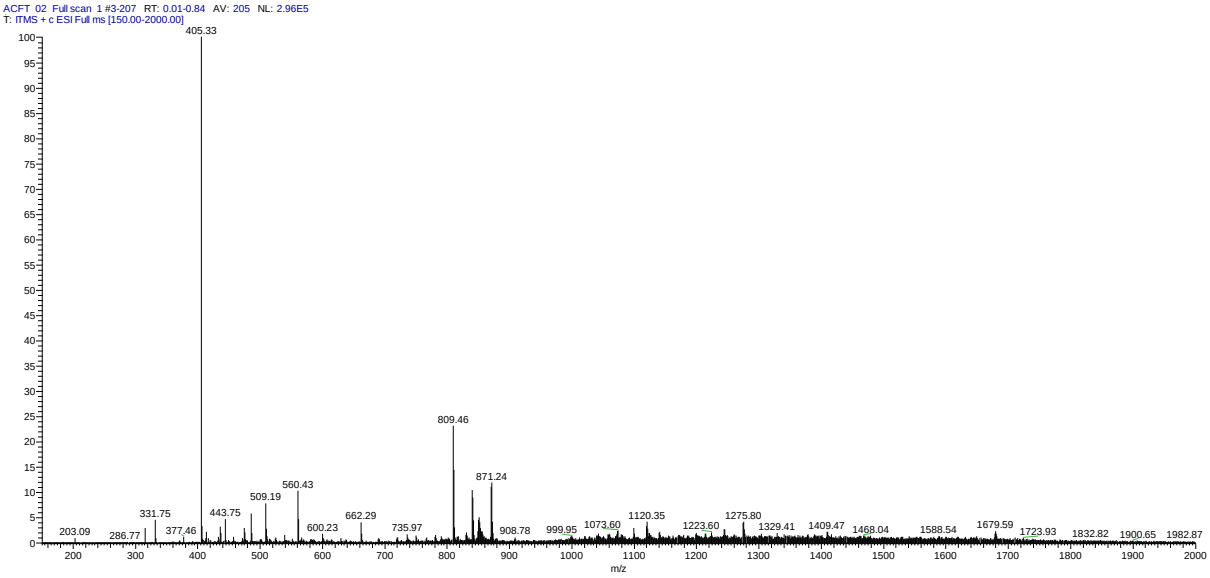


Figure S4.9. ESI-MS spectrum of compound **5**.

Compound **6**

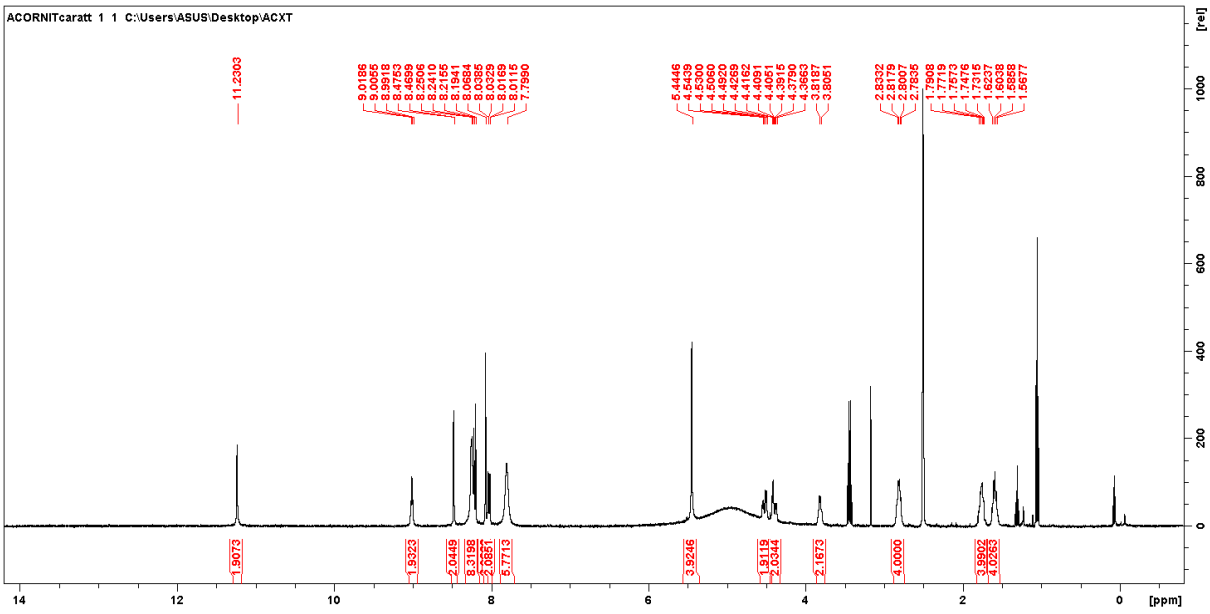


Figure S4.10. ^{1}H-NMR spectrum of compound **6**.

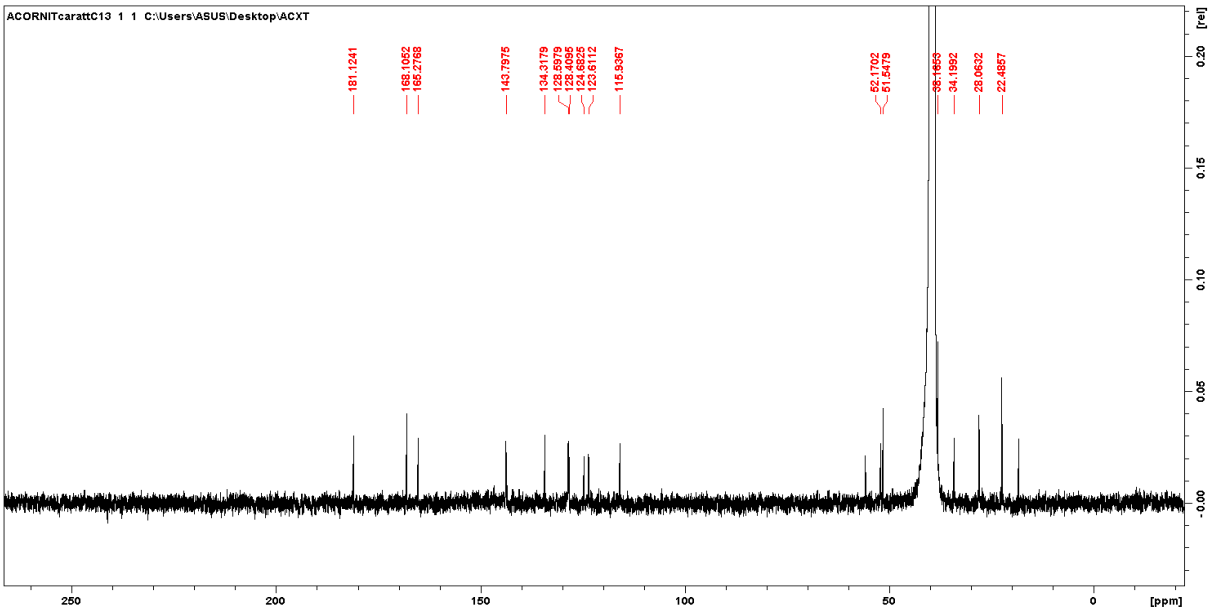


Figure S4.11. ^{13}C-NMR spectrum of compound **6**.

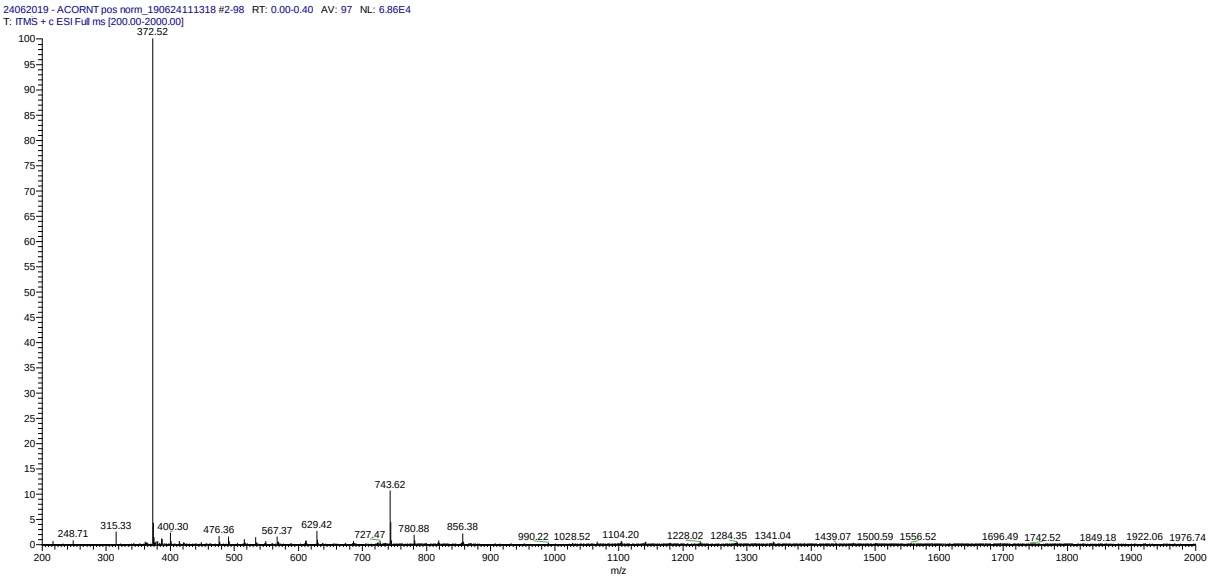


Figure S4.12. ESI-MS spectrum of compound **6**.

Compound **7**

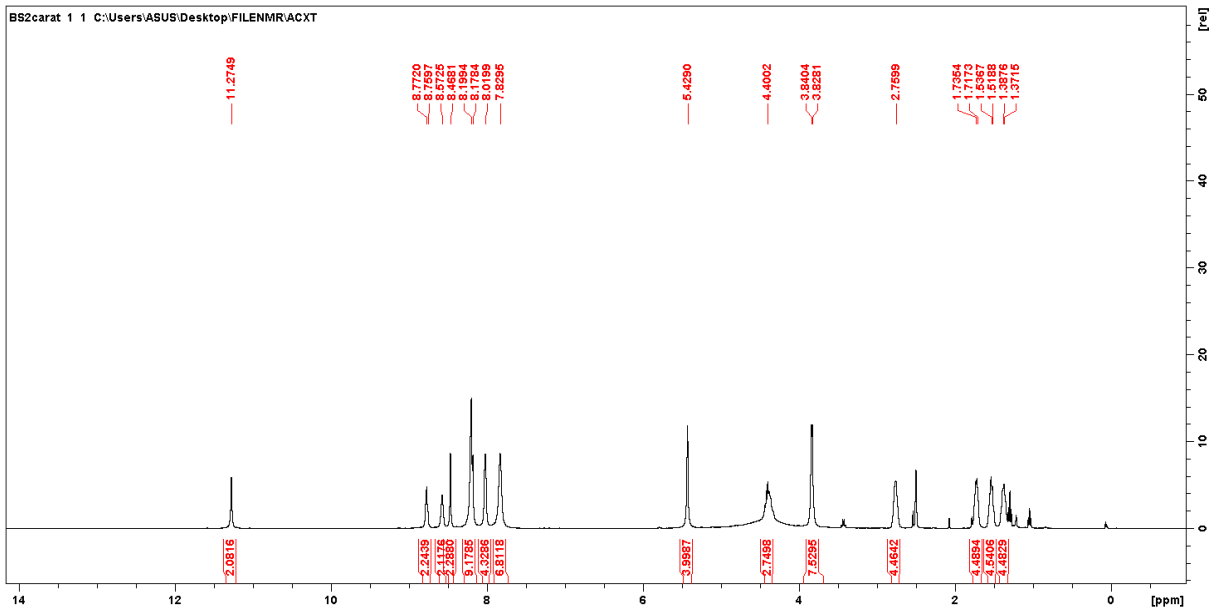


Figure S4.13. ^{1}H-NMR spectrum of compound **7**.

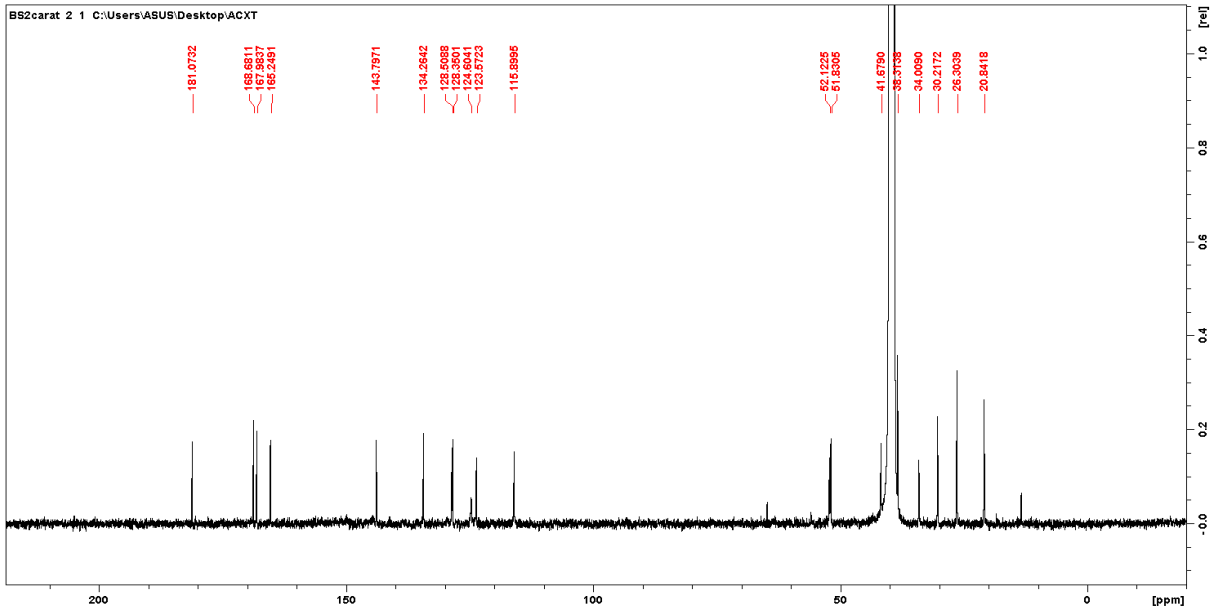


Figure S4.14. ^{13}C-NMR spectrum of compound **7**.

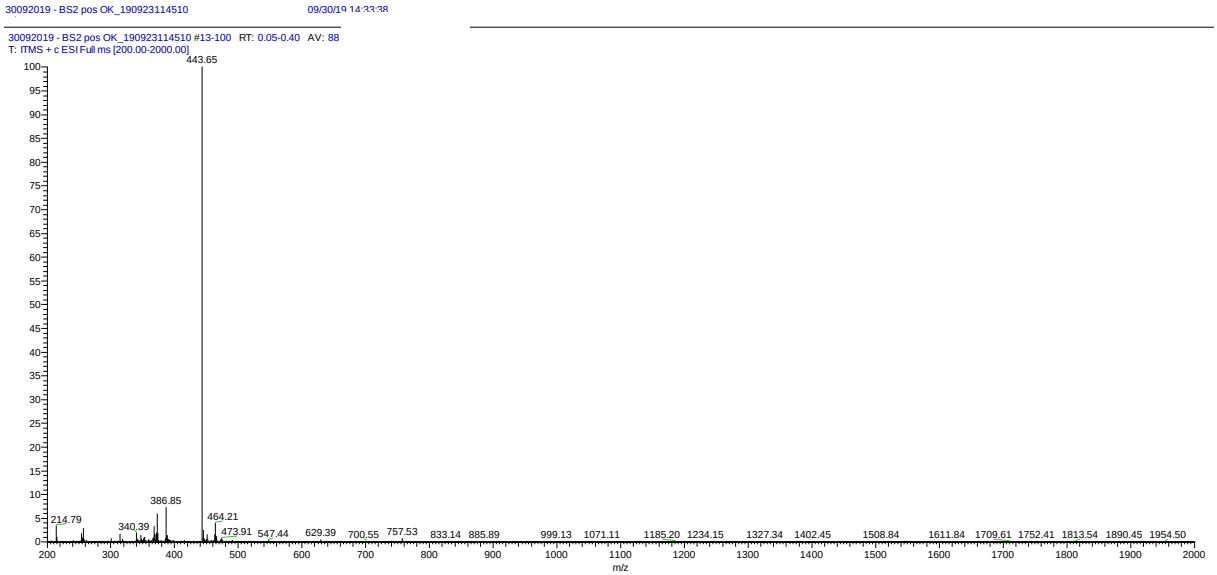


Figure S4.15. ESI-MS spectrum of compound **7**.

Compound **8**

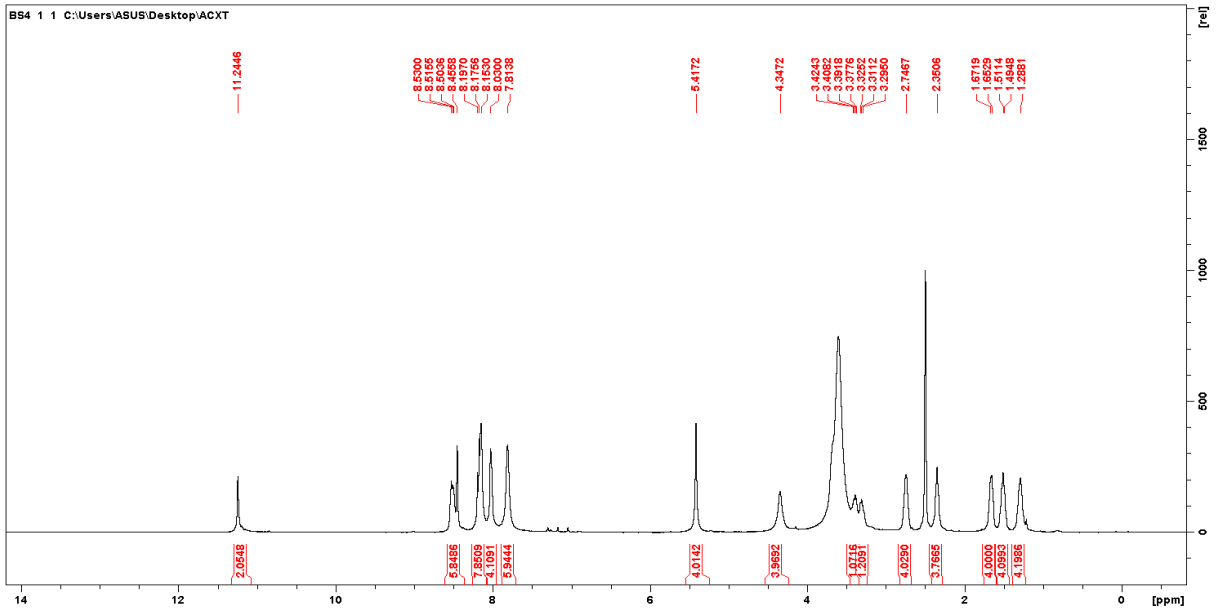


Figure S4.16. ^{1}H-NMR spectrum of compound **8**.

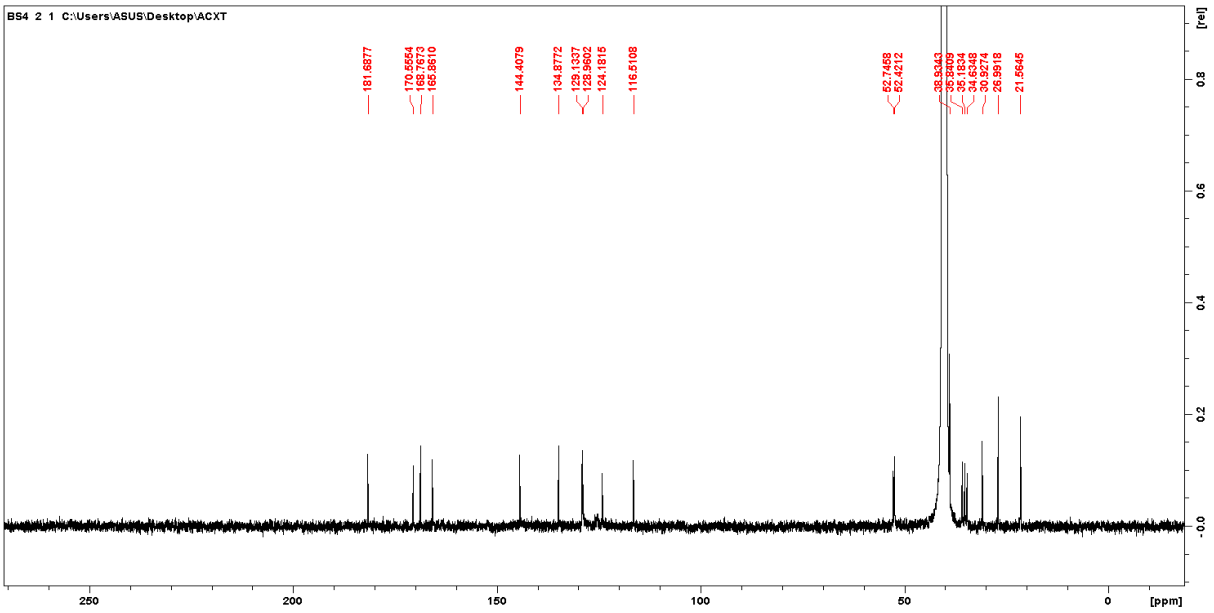


Figure S4.17. ^{13}C-NMR spectrum of compound **8**.

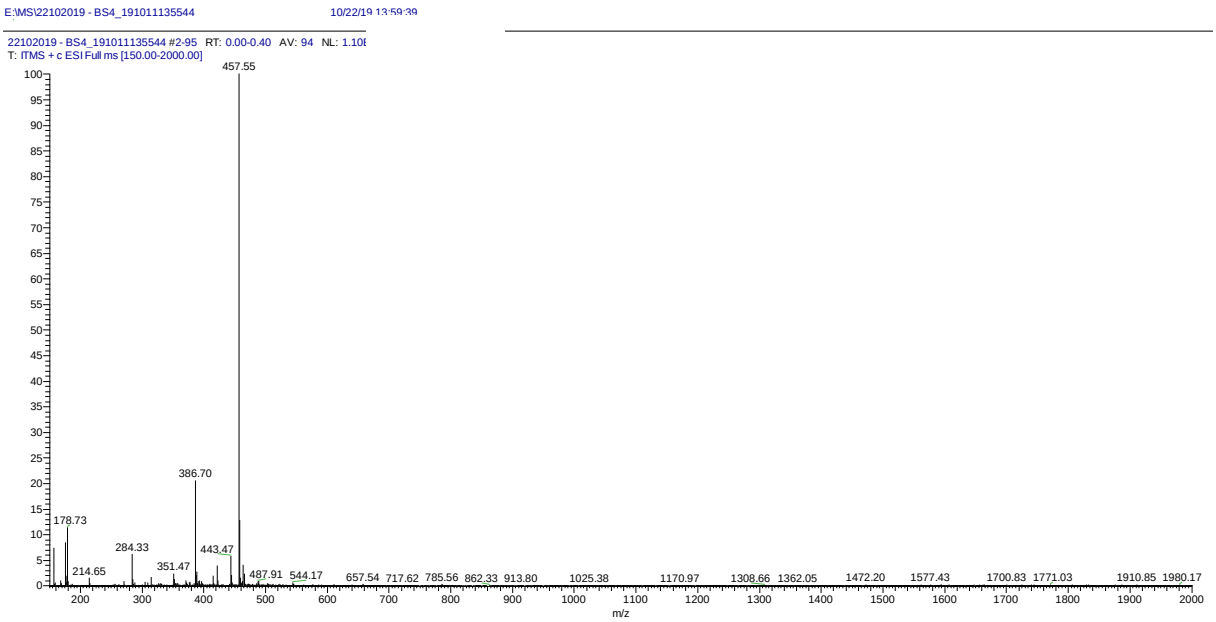


Figure S4.18. ESI-MS spectrum of compound **8**.

Compound **9**

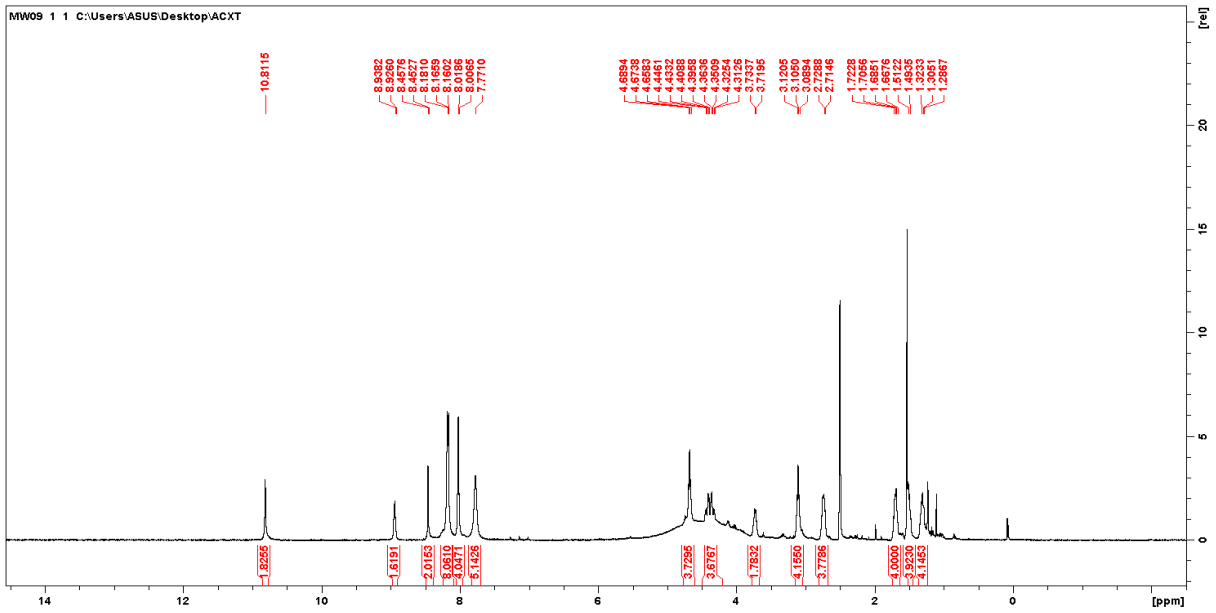


Figure S4.19. ^{1}H-NMR spectrum of compound **9**.

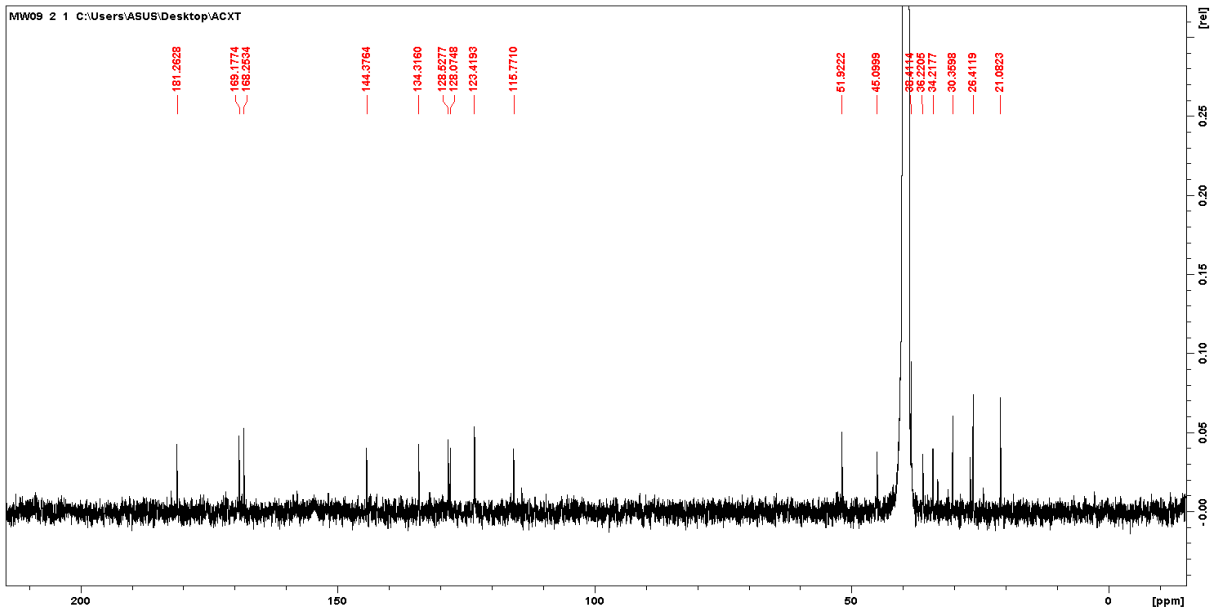


Figure S4.20. ^{13}C-NMR spectrum of compound **9**.

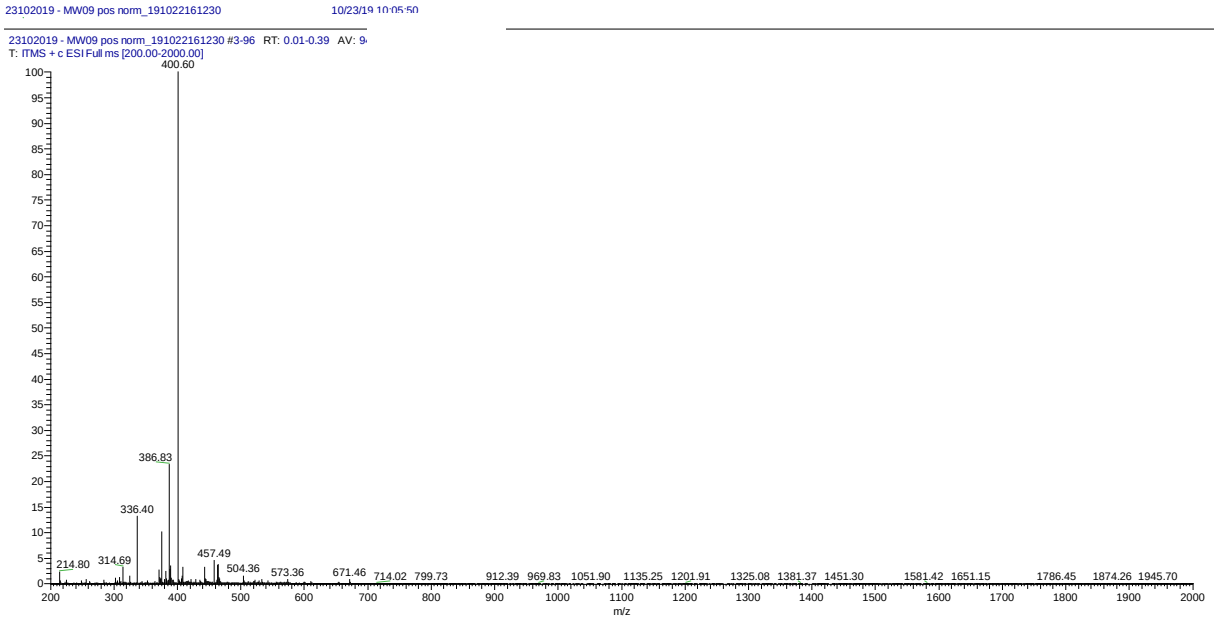


Figure S4.21. ESI-MS spectrum of compound **9**.

ESI-MS binding studies

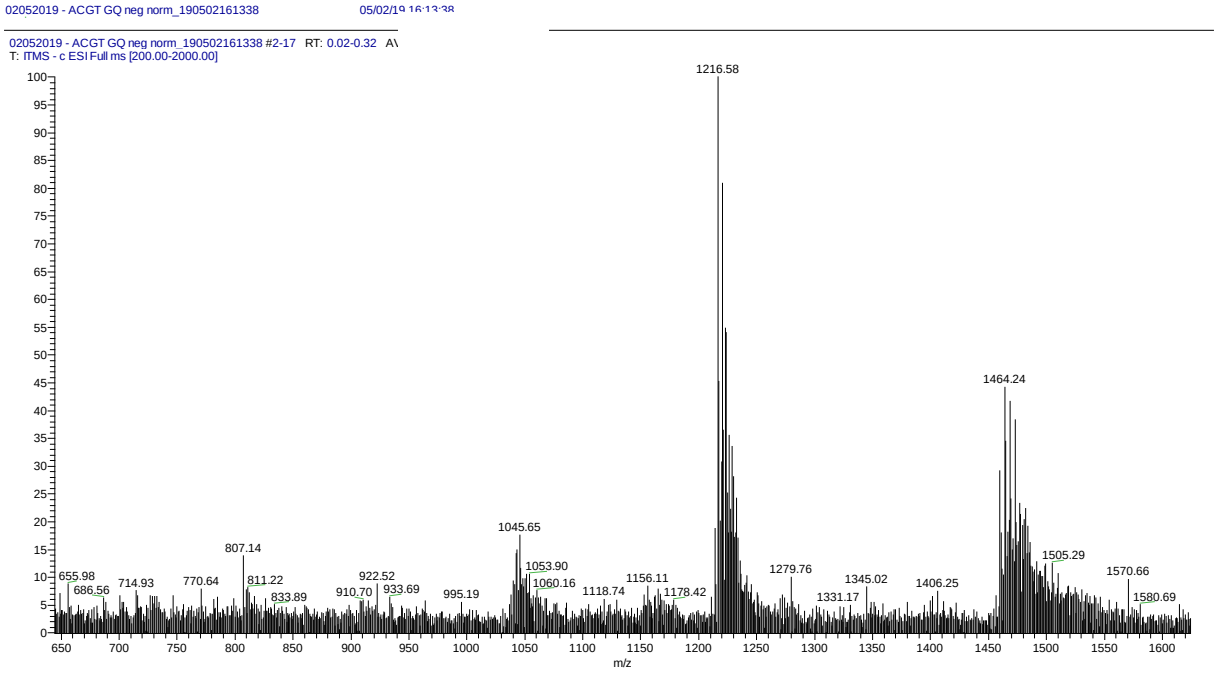


Figure S4.22. Mass spectrum of the G4/**3** binding experiment. The signals at m/z = 1045.65, 1216.58, and 1464.24 were attributed to the unbound DNA (z = -7, z = -6 and z = -5, respectively). No interaction peaks were detected.

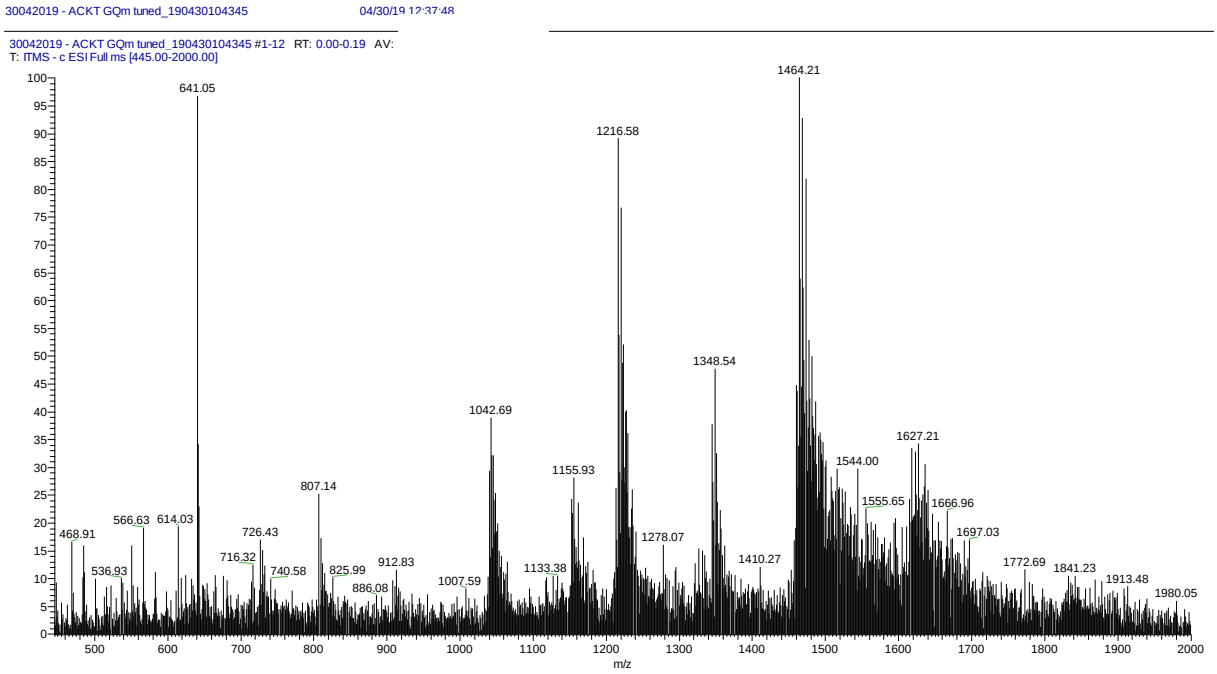


Figure S4.23. Mass spectrum of the G4/**4** binding experiment. The signals at m/z = 1042.69, 1216.58, and 1464.21 were attributed to the unbound DNA (z = -7, z = -6 and z = -5, respectively). The interaction peaks were detected at m/z = 1155.93, 1348.54 and 1627.21 (z = -7, z = -6 and z = -5, respectively).

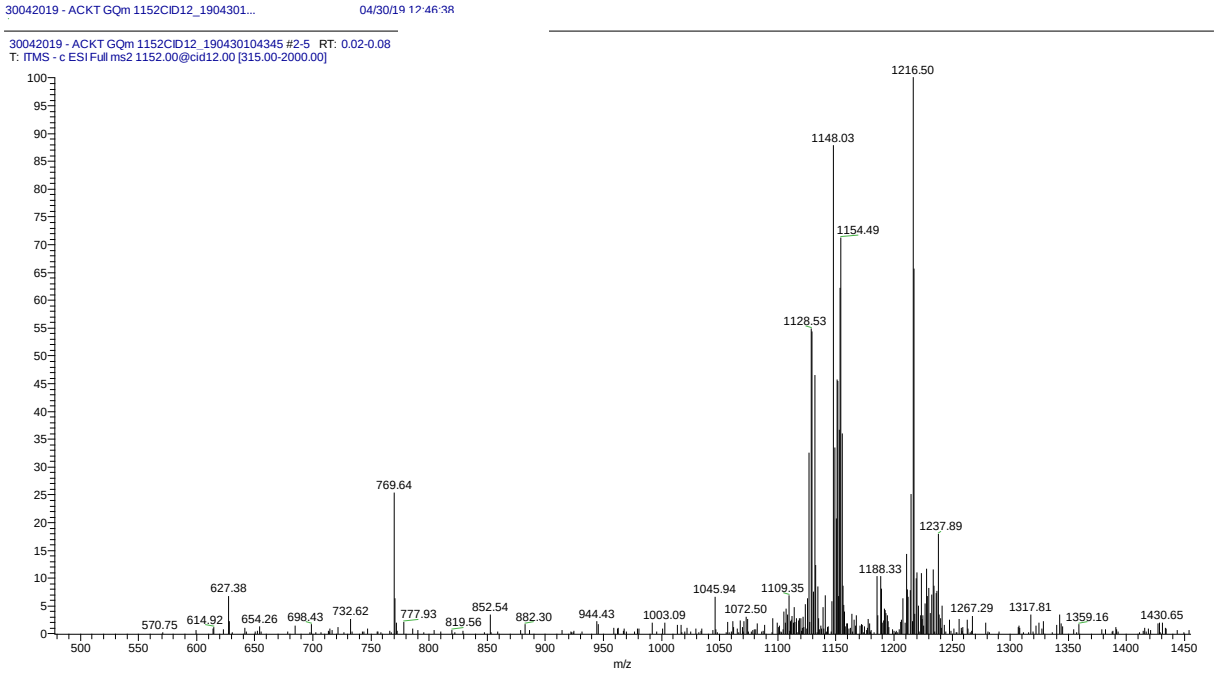


Figure S4.24. CID fragmentation spectrum of the G4/**4** complex (normalized collision energy = 12). The signal of the complex is still present (m/z = 1154.49, z = -7), and molecular ion of naked DNA (m/z = 1216.50, z = -6) and ligand (m/z = 769.64) start to appear.

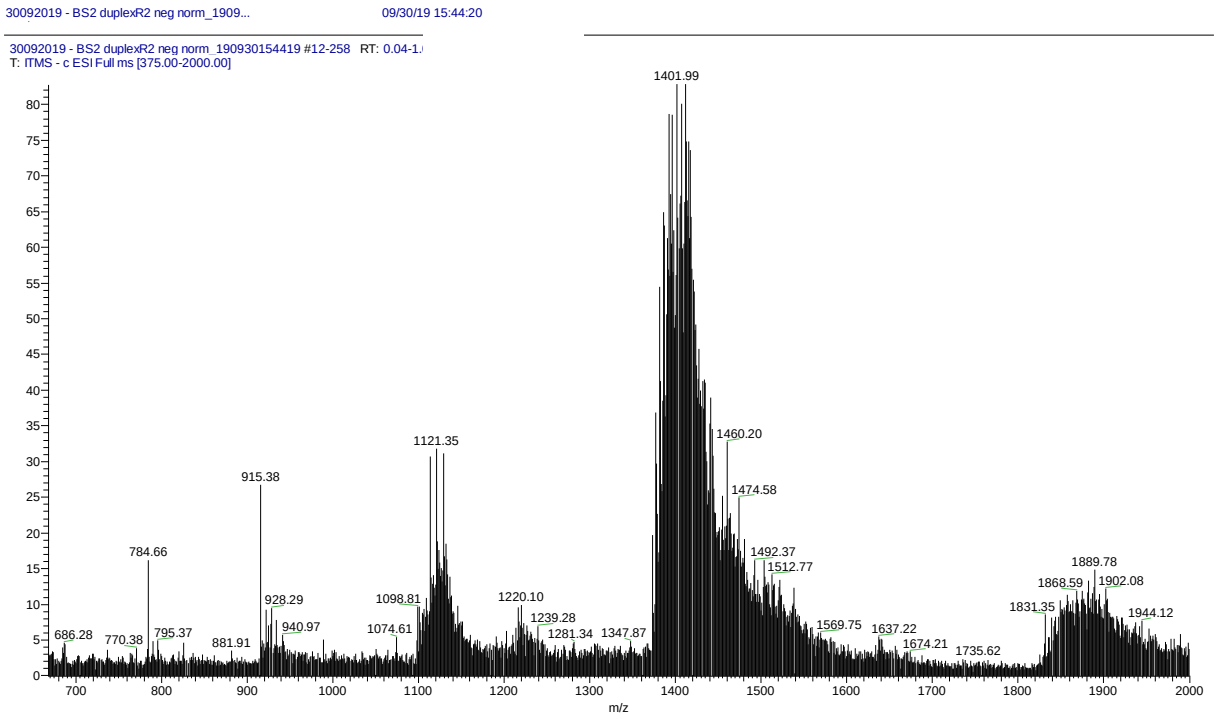


Figure S4.25. Mass spectrum of the dsDNA/**7** binding experiment. The signals at m/z = 915.38, 1121.35 and 1401.99 were attributed to the unbound DNA (z = -12, z = -10 and z = -8, respectively). No interaction peaks were detected.

Sequences used for fluorescence melting and circular dichroism

Table S4.1. Oligonucleotide sequences used in fluorescence melting study (marked with -L) and CD.

Oligonucleotide	***Sequence***	***Main G4 topology***
kit*-L	Dabcyl-GGCGAGGAGGGGCGTGGCCGGC-FAM	antiparallel
tel24-L	Dabcyl-TTAGGGTTAGGGTTAGGGTTAGGG-FAM	hybrid 1
tel23-L	Dabcyl-AGGGTTAGGGTTAGGGTTAGGGT-FAM	hybrid 2
bcl2-L	Dabcyl-GGGGTCGAGGGTGGGGTGCCGGGGG-FAM	parallel
scr_L	GGGTGTGAGTGTGAGTGTGAGG-FAM	-
co-scr_L	Dabcyl-CCTCACACTCACACTCACATCC	-
Tel24	TTAGGGTTAGGGTTAGGGTTAGGG	hybrid 1
Tel 23	AGGGTTAGGGTTAGGGTTAGGGT	hybrid 2

5 General conclusions and future perspectives

In this doctoral thesis it has been presented the research work that I have conducted during these three last years on the development of new ligands able to target the telomeric DNA, either in its double strand form and even more in its non-canonical G4 fold.

In a first study (Chapter 2) the use of the anthracene-propargylamine scaffold for the preparation of new ligands for both the dsDNA and the G4 was envisioned. A set of 6 final compounds were synthesized and then tested for their ability to bind to the telomeric nucleic acid. The results indicated a good ability for some of the compounds to bind to the dsDNA, even when compared to a potent intercalant such as mitoxantrone. An important affinity for the G4 was also observed in particular for one of the compounds; however, a selectivity for the non-canonical over the canonical forms of the nucleic acids was not evidenced.

In a second study (Chapter 3) we decided to optimize the anthracene-propargylamine scaffold for the binding to the G4 DNA, increasing therefore the selectivity for the latter over the dsDNA. In order to do that a new di-substituted ligand was designed taking the steps from the best ligand of the precedent work. A preliminary *in silico* study indicated the increased affinity for the telomeric G4 DNA of the new ligand respect to the previous one, then a set of binding experiments demonstrated a substantial increase in the affinity and selectivity of the ligand for the non-canonical form.

The third study (Chapter 4) involved the preparation of a set of 7 anthraquinone-amino acid conjugates as new ligands for the telomeric G4 DNA. The ligands demonstrated a good affinity and high selectivity for the G4 over the dsDNA and the ability to change, upon titration, the G4 topology from the hybrid to the antiparallel. Moreover, a high ability to stabilize the G4 arising from the blc2 sequence, which folds into the parallel topology, was also evidenced. A further computational study demonstrated the ability of the ligands to induce, in the simulation timeframe, variations on the G4 structure leading in most cases, and also for the hybrid and antiparallel topology, to the formation of stable complexes.

The ligands prepared in the three reported studies generally showed a marked the ability to bind to the nucleic acid; however, optimization is still possible for each set and other studies would surely increase the scientific values of the findings. In particular, for what concerns the mono-substituted derivatives, their role as dsDNA intercalants should be studied more deeply possibly by *in vitro* evaluations as antiproliferative agents against cancer cell lines. On the other hand, the di-substituted derivative demonstrated a greater affinity for the G4; however, the value appears still low and therefore a further optimization of the scaffold should be considered, moreover, in order to possibly evidence even higher affinity values, the ligand should be studied with other telomeric sequences, folding in different topologies. For what concerns the anthraquinone-amino acid conjugates, the best

ligands showed a great selectivity for the G4 and a significant ability to stabilize the G4 arrangement with high observed ΔT_m temperatures, for the different sequences. A possibility for a further evolution of these compounds could be represented by the side chains modification, for example decreasing the amount of H-bond donors by methylation of the amines therefore decreasing the interaction with the bulk solvent.

6 References

1. Chambers, V. S. *et al.* High-throughput sequencing of DNA G-quadruplex structures in the human genome. *Nat. Biotechnol.* **33**, 877–881 (2015).

2. Kolesnikova, S. and Curtis, E. A. Structure and Function of Multimeric G-Quadruplexes. *Mol.* **24**, 3074 (2019).

3. Fujioka, A. *et al.* Dynamics of the Ras/ERK MAPK Cascade as Monitored by Fluorescent Probes. *J. Biol. Chem.* **281**, 8917–8926 (2006).

4. Han, H. and Hurley, L. H. G-quadruplex DNA: a potential target for anti-cancer drug design. *Trends Pharmacol. Sci.* **21**, 136–142 (2000).

5. Francisco, A. P. and Paulo, A. Oncogene Expression Modulation in Cancer Cell Lines by DNA G-Quadruplex-Interactive Small Molecules. *Curr. Med. Chem.* **24**, (2016).

6. Müller, S. and Rodriguez, R. G-quadruplex interacting small molecules and drugs: from bench toward bedside. **7**, 663–679 (2014).

7. Siddiqui-Jain, A. *et al.* Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription. *Proc. Natl. Acad. Sci.* **99**, 11593–11598 (2002).

8. Sun, D. *et al.* Facilitation of a structural transition in the polypurine/polypyrimidine tract within the proximal promoter region of the human VEGF gene by the presence of potassium and G-quadruplex-interactive agents. *Nucleic Acids Res.* **33**, 6070–6080 (2005).

9. Dexheimer, T. S. *et al.* 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).

10. Cogoi, S. and Xodo, L. E. G-quadruplex formation within the promoter of the KRAS proto-oncogene and its effect on transcription. *Nucleic Acids Res.* **34**, 2536–2549 (2006).

11. Yang, D. and Hurley, L. H. Structure of the Biologically Relevant G-Quadruplex in The c-MYC Promoter. **25**, 951–968 (2007).

12. Rankin, S. *et al.* Putative DNA Quadruplex Formation within the Human c-kit Oncogene. (2005).

13. O'Hagan, M. P. *et al.* Binding and Beyond: What Else Can G-Quadruplex Ligands Do? *European J. Org. Chem.* **2019**, 4995–5017 (2019).

14. Shay, J. W. and Wright, W. E. Telomeres and telomerase: three decades of progress. *Nat. Rev. Genet. 2019 205* **20**, 299–309 (2019).

15. Hayflick, L. The limited in vitro lifetime of human diploid cell strains. *Exp. Cell Res.* **37**, 614–636 (1965).

16. Jafri, M. A. *et al.* Roles of telomeres and telomerase in cancer, and advances in telomerase-targeted therapies. *Genome Med.* **8**, (2016).

17. NW, K. *et al.* Specific association of human telomerase activity with immortal cells and cancer. *Science* **266**, 2011–2015 (1994).

18. Maciejowski, J. and de Lange, T. Telomeres in cancer: tumour suppression and genome instability. (2017).

19. 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).

20. Di Antonio, M. *et al.* Selective RNA versus DNA G-quadruplex targeting by situ click chemistry. *Angew.*

Chemie - Int. Ed. **51**, 11073–11078 (2012).

21. Fay, M. M. *et al.* P. RNA G-Quadruplexes in Biology: Principles and Molecular Mechanisms. *J. Mol. Biol.* **429**, 2127–2147 (2017).

22. Cusanelli, E. and Chartrand, P. Telomeric repeat-containing RNA TERRA: A noncoding RNA connecting telomere biology to genome integrity. *Front. Genet.* **6**, (2015).

23. Métifiot, M. *et al.* G-quadruplexes in viruses: function and potential therapeutic applications. *Nucleic Acids Res.* **42**, 12352–12366 (2014).

24. Cui, H. and Zhang, L. G-Quadruplexes Are Present in Human Coronaviruses Including SARS-CoV-2. *Front. Microbiol.* **0**, 2570 (2020).

25. Zhang, R. *et al.* Whole Genome Identification of Potential G-Quadruplexes and Analysis of the G-Quadruplex Binding Domain for SARS-CoV-2. *Front. Genet.* **0**, 1430 (2020).

26. Zhao, C. *et al.* Targeting RNA G-Quadruplex in SARS-CoV-2: A Promising Therapeutic Target for COVID-19? *Angew. Chemie Int. Ed.* **60**, 432–438 (2021).

27. Gueddouda, N. M. *et al.* Design, Synthesis, and Evaluation of 2,9-Bis[(substituted-aminomethyl)phenyl]-1,10-phenanthroline Derivatives as G-Quadruplex Ligands. *ChemMedChem* **12**, 146–160 (2017).

28. Hänsel-Hertsch, R. *et al.* S. DNA G-quadruplexes in the human genome: Detection, functions and therapeutic potential. *Nat. Rev. Mol. Cell Biol.* **18**, 279–284 (2017).

29. Shivalingam, A. *et al.* The interactions between a small molecule and G-quadruplexes are visualized by fluorescence lifetime imaging microscopy. *Nat. Commun.* **6**, (2015).

30. Qin, H. *et al.* Metallo-supramolecular Complexes Enantioselectively Eradicate Cancer Stem Cells in Vivo. *J. Am. Chem. Soc.* **139**, 16201–16209 (2017).

31. Abraham Punnoose, J. *et al.* Adaptive and Specific Recognition of Telomeric G-Quadruplexes via Polyvalency Induced Unstacking of Binding Units. *J. Am. Chem. Soc.* **139**, 7476–7484 (2017).

32. Neidle, S. Quadruplex nucleic acids as targets for anticancer therapeutics. *Nat. Rev. Chem.* **1**, (2017).

33. Xu, H. *et al.* CX-5461 is a DNA G-quadruplex stabilizer with selective lethality in BRCA1/2 deficient tumours. *Nat. Commun.* **8**, (2017).

34. Alberti, P. *et al.* DNA nanomachines and nanostructures involving quadruplexes. *Org. Biomol. Chem.* **4**, 3383–3391 (2006).

35. Olsen, C. M. *et al.* Unfolding of G-Quadruplexes: Energetic, and Ion and Water Contributions of G-Quartet Stacking. *J. Phys. Chem. B* **110**, 6962–6969 (2006).

36. Chen, Y. and Yang, D. Sequence, stability, and structure of G-Quadruplexes and their interactions with drugs. *Curr. Protoc. Nucleic Acid Chem.* (2012).

37. 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).

38. Maity, A. *et al.*Intra-locked G-quadruplex structures formed by irregular DNA G-rich motifs. *Nucleic Acids Res.* **48**, 3315–3327 (2020).

39. Cang, X. *et al.* Explaining the varied glycosidic conformational, G-tract length and sequence preferences for anti-parallel G-quadruplexes. *Nucleic Acids Res.* **39**, 4499–4512 (2011).

40. Largy, E. *et al.* Role of Alkali Metal Ions in G-Quadruplex Nucleic Acid Structure and Stability. *Met. Ions*

Life Sci. **16**, 203–258 (2016).

41. Lane, A. N. *et al.* Stability and kinetics of G-quadruplex structures. *Nucleic Acids Res.* **36**, 5482–5515 (2008).

42. Otero, R. *et al.* Guanine quartet networks stabilized by cooperative hydrogen bonds. *Angew. Chemie - Int. Ed.* **44**, 2270–2275 (2005).

43. Fonseca Guerra, C. *et al.* Telomere structure and stability: Covalency in hydrogen bonds, not resonance assistance, causes cooperativity in guanine quartets. *Chem. - A Eur. J.* **17**, 12612–12622 (2011).

44. Xu, W. *et al.* Atomic-scale structures and interactions between the guanine quartet and potassium. *Chem. Commun.* **49**, 7210–7212 (2013).

45. Venczel, E. A. and Sen, D. Parallel and Antiparallel G-DNA Structures from a Complex Telomeric Sequence. *Biochemistry* **32**, 6220–6228 (1993).

46. Włodarczyk, A. *et al.* Effect of ions on the polymorphism, effective charge, and stability of human telomeric DNA. Photon correlation spectroscopy and circular dichroism studies. *J. Phys. Chem. B* **109**, 3594–3605 (2005).

47. Clark, G. R. *et al.* Structure of the first parallel DNA quadruplex-drug complex. *J. Am. Chem. Soc.* **125**, 4066–4067 (2003).

48. Lin, C. and Yang, D. Human telomeric G-quadruplex structures and G-quadruplex-interactive compounds. *Methods Mol. Biol.* **1587**, 171–196 (2017).

49. Li, J. *et al.* Not so crystal clear: The structure of the human telomere G-quadruplex in solution differs from that present in a crystal. *Nucleic Acids Res.* **33**, 4649–4659 (2005).

50. Martino, L. *et al.* Shedding light on the interaction between TMPyP4 and human telomeric quadruplexes. *J. Phys. Chem. B* **113**, 14779–14786 (2009).

51. Zhou, J. *et al.* The structural transition and compaction of human telomeric G-quadruplex induced by excluded volume effect under cation-deficient conditions. *Biophys. Chem.* **136**, 124–127 (2008).

52. Zhou, J. *et al.* Human telomeric G-quadruplex formed from duplex under near physiological conditions: Spectroscopic evidence and kinetics. *Biochimie* **91**, 1104–1111 (2009).

53. Xue, Y. *et al.* Human Telomeric DNA Forms Parallel-Stranded Intramolecular G-Quadruplex in K+ Solution under Molecular Crowding Condition. *J. Am. Chem. Soc.* **129**, 11185–11191 (2007).

54. Zhou, J. *et al.* A “sugar-deficient” G-quadruplex: Incorporation of aTNA in G4 structures. *Chem. Sci.* **4**, 3693–3698 (2013).

55. Esposito, V. *et al.* A new class of DNA quadruplexes formed by oligodeoxyribonucleotides containing a 3’-3’ or 5’-5’ inversion of polarity site. *Chem. Commun.* 3953–3955 (2005).

56. Esposito, V. *et al.* Effects of the introduction of inversion of polarity sites in the quadruplex forming oligonucleotide TGGGT. *Bioorganic Med. Chem.* **17**, 1997–2001 (2009).

57. Webba da Silva, M. NMR methods for studying quadruplex nucleic acids. *Methods* **43**, 264–277 (2007).

58. Ma, Y. *et al.* Topologies of G-quadruplex: Biological functions and regulation by ligands. *Biochem. Biophys. Res. Commun.* **531**, 3–17 (2020).

59. Zakian, V. A. Telomeres: Beginning to Understand the End. *Science (80-.).* **270**, 1601–1607 (1995).

60. 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).

61. 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).

62. Dai, J. *et al.* Structure of the hybrid-2 type intramolecular human telomeric G-quadruplex in K+ solution: Insights into structure polymorphism of the human telomeric sequence. *Nucleic Acids Res.* **35**, 4927–4940 (2007).

63. Luu, K. N. *et al.* Structure of the human telomere in K+ solution: An intramolecular (3 + 1) G-quadruplex scaffold. *J. Am. Chem. Soc.* **128**, 9963–9970 (2006).

64. Lin, C. *et al.* Molecular Recognition of the Hybrid-2 Human Telomeric G-Quadruplex by Epiberberine: Insights into Conversion of Telomeric G-Quadruplex Structures. *Angew. Chemie - Int. Ed.* **57**, 10888–10893 (2018).

65. Chen, M. *et al.* Small-molecule selectively recognizes human telomeric G-quadruplex DNA and regulates its conformational switch. *Biophys. J.* **97**, 2014–2023 (2009).

66. Wang, Z. F. *et al.* A novel transition pathway of ligand-induced topological conversion from hybrid forms to parallel forms of human telomeric G-quadruplexes. *Nucleic Acids Res.* **44**, 3958–3968 (2016).

67. Wang, Y. and Patel, D. J. Solution structure of the human telomeric repeat d[AG3(T2AG3)3] G-tetraplex. *Structure* **1**, 263–282 (1993).

68. Mashimo, T. *et al.* Folding pathways of hybrid-1 and hybrid-2 G-quadruplex structures. *Nucleic Acids Symp. Ser. (Oxf).* 409–410 (2008).

69. Dai, J. *et al.* Polymorphism of human telomeric quadruplex structures. *Biochimie* **90**, 1172–1183 (2008).

70. Miller, M. C. *et al.* Hydration is a major determinant of the G-quadruplex stability and conformation of the human telomere 3' sequence of d(AG 3(TTAG3)3). *J. Am. Chem. Soc.* **132**, 17105–17107 (2010).

71. Renčiuk, D. *et al.* Arrangements of human telomere DNA quadruplex in physiologically relevant K+ solutions. *Nucleic Acids Res.* **37**, 6625–6634 (2009).

72. Pagano, B. *et al.* Looking for efficient G-quadruplex ligands: Evidence for selective stabilizing properties and telomere damage by drug-like molecules. *ChemMedChem* **10**, 640–649 (2015).

73. Phan, A. T. *et al.* Different loop arrangements of intramolecular human telomeric (3+1) G-quadruplexes in K+ solution. *Nucleic Acids Res.* **34**, 5715–5719 (2006).

74. Bao, H. *et al.* Hybrid-type and two-tetrad antiparallel telomere DNA G-quadruplex structures in living human cells. *Nucleic Acids Res.* **47**, 4940–4947 (2019).

75. Limongelli, V. *et al.* The G-triplex DNA. *Angew. Chemie - Int. Ed.* **52**, 2269–2273 (2013).

76. Jiang, H.-X. *et al.* Divalent cations and molecular crowding buffers stabilize G-triplex at physiologically relevant temperatures. *Sci. Reports 2015 51* **5**, 1–11 (2015).

77. Koirala, D. *et al.* Intramolecular folding in three tandem guanine repeats of human telomeric DNA. *Chem. Commun.* **48**, 2006–2008 (2012).

78. Cerofolini, L. *et al.* G-triplex structure and formation propensity. *Nucleic Acids Res.* **42**, 13393–13404 (2014).

79. Zhou, H. *et al.* Stable and Label-Free Fluorescent Probe Based on G-triplex DNA and Thioflavin T. *Anal. Chem.* **90**, 3220–3226 (2018).

80. Zhang, Q. *et al.* Characterization of intermolecular G-quadruplex formation over intramolecular G-triplex for DNA containing three G-tracts. *Analyst* **145**, 4254–4259 (2020).

81. Williamson, J. R. *et al.* Monovalent cation-induced structure of telomeric DNA: The G-quartet model. *Cell* **59**, 871–880 (1989).

82. Laughlan, G. *et al.* The high-resolution crystal structure of a parallel-stranded guanine tetraplex. *Science (80-.).* **265**, 520–524 (1994).

83. Phillips, K. *et al.* The crystal structure of a parallel-stranded guanine tetraplex at 0.95 Å resolution. *J. Mol. Biol.* **273**, 171–182 (1997).

84. Keniry, M. A. *et al.* Solution Structure of the Na+ form of the Dimeric Guanine Quadruplex [d(G3T4G3)]2. *Eur. J. Biochem.* **233**, 631–643 (1995).

85. Sundquist, W. I. and Klug, A. Telomeric DNA dimerizes by formation of guanine tetrads between hairpin loops. *Nature* **342**, 825–829 (1989).

86. Merkina, E. E. and Fox, K. R. Kinetic stability of intermolecular DNA quadruplexes. *Biophys. J.* **89**, 365–373 (2005).

87. Bryan, T. M. G-quadruplexes at telomeres: Friend or foe? *Mol.* **25** (2020).

88. Petraccone, L. Higher-Order Quadruplex Structures. 23–46 (2012).

89. Yu, H. Q. *et al.* Characterization of structure and stability of long telomeric DNA G-quadruplexes. *J. Am. Chem. Soc.* **128**, 15461–15468 (2006).

90. Zhao, J. and Zhai, Q. Recent advances in the development of ligands specifically targeting telomeric multimeric G-quadruplexes. *Bioorg. Chem.* **103**, (2020).

91. Abraham Punnoose, J. *et al.* Interaction of G-quadruplexes in the full-length 3 human telomeric overhang. *J. Am. Chem. Soc.* **136**, 18062–18069 (2014).

92. Zhou, C. Q. *et al.* Dinickel–Salphen Complexes as Binders of Human Telomeric Dimeric G-Quadruplexes. *Chem. - A Eur. J.* **23**, 4713–4722 (2017).

93. Collie, G. W. *et al.* Electrospray mass spectrometry of telomeric RNA (TERRA) Reveals the formation of stable multimeric G-quadruplex structures. *J. Am. Chem. Soc.* **132**, 9328–9334 (2010).

94. Xu, Y. *et al.* Consecutive formation of G-Quadruplexes in human telomerieoverhang DNA: A protective capping structure for telomere ends. *Angew. Chemie - Int. Ed.* **48**, 7833–7836 (2009).

95. Creze, C. *et al.* Structure of a d(TGGGGT) quadruplex crystallized in the presence of Li + ions. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **63**, 682–688 (2007).

96. Mal, P. K. *et al.* Racemic DNA crystallography. *Angew. Chemie - Int. Ed.* **53**, 14424–14427 (2014).

97. Adrian, M. *et al.* NMR spectroscopy of G-quadruplexes. *Methods* **57**, 11–24 (2012).

98. Bardin, C. and Leroy, J. L. The formation pathway of tetramolecular G-quadruplexes. *Nucleic Acids Res.* **36**, 477–488 (2008).

99. Matsugami, A. *et al.* Structure of a human telomeric DNA sequence stabilized by 8-bromoguanosine substitutions, as determined by NMR in a K+ solution. *FEBS J.* **274**, 3545–3556 (2007).

100. Zhang, N. *et al.* (3 + 1) Assembly of three human telomeric repeats into an asymmetric dimeric G-quadruplex. *J. Am. Chem. Soc.* **127**, 17277–17285 (2005).

101. Yue, D. J. E. *et al.* Formation of (3+1) G-quadruplexes with a long loop by human telomeric DNA spanning five or more repeats. *J. Am. Chem. Soc.* **133**, 11462–11465 (2011).

102. Phan, A. T. and Patel, D. J. Two-Repeat Human Telomeric d(TAGGGTTAGGGT) Sequence Forms Interconverting Parallel and Antiparallel G-Quadruplexes in Solution: Distinct Topologies, Thermodynamic Properties, and Folding/Unfolding Kinetics. *J. Am. Chem. Soc.* **125**, 15021–15027 (2003).

103. Wang, Z. F. *et al.* Structural basis of sodium-potassium exchange of a human telomeric DNA quadruplex without topological conversion. *Nucleic Acids Res.* **42**, 4723–4733 (2014).

104. Phan, A. T. Human telomeric G-quadruplex: Structures of DNA and RNA sequences. *FEBS J.* **277**, 1107–1117 (2010).

105. Phan, A. T. *et al.* Structure of two intramolecular G-quadruplexes formed by natural human telomere sequences in K+ solution. *Nucleic Acids Res.* **35**, 6517–6525 (2007).

106. Zhang, Z. *et al.* 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).

107. Lim, K. W. *et al.* Sequence variant (CTAGGG)n in the human telomere favors a G-quadruplex structure containing a G·C·G·C tetrad. *Nucleic Acids Res.* **37**, 6239–6248 (2009).

108. Martadinata, H. and Phan, A. T. Structure of human telomeric RNA (TERRA): Stacking of two G-quadruplex blocks in K+ solution. *Biochemistry* **52**, 2176–2183 (2013).

109. Xu, Y. *et al.* G-quadruplex formation by human telomeric repeats-containing RNA in Na + solution. *J. Am. Chem. Soc.* **130**, 11179–11184 (2008).

110. Martadinata, H. and Phan, A. T. Structure of propeller-type parallel-stranded RNA G-quadruplexes, formed by human telomeric RNA sequences in K+ solution. *J. Am. Chem. Soc.* **131**, 2570–2579 (2009).

111. Wang, Y. and Patel, D. J. Solution structure of the Tetrahymena telomeric repeat d(T2G4)4 G-tetraplex. *Structure* **2**, 1141–1156 (1994).

112. Liu, C. *et al.* A chair-type G-quadruplex structure formed by a human telomeric variant DNA in K+ solution. *Chem. Sci.* **10**, 218–226 (2019).

113. Schultze, P. *et al.* The effect of sodium, potassium and ammonium ions on the conformation of the dimeric quadruplex formed by the Oxytricha nova telomere repeat oligonucleotide d(G4T4G4). *Nucleic Acids Res.* **27**, 3018–3028 (1999).

114. Smith, F. W. and Feigon, J. Quadruplex structure of Oxytricha telomeric DNA oligonucleotides. *Nature* **356**, 164–168 (1992).

115. Schultze, P. *et al.* Refined solution structure of the dimeric quadruplex formed from the Oxytricha telomeric oligonucleotide d(GGGGTTTTGGGG). *Structure* **2**, 221–233 (1994).

116. Wang, Y. and Patel, D. J. Solution structure of the oxytricha telomeric repeat d[G4(T4G4)3] G-tetraplex. *J. Mol. Biol.* **251**, 76–94 (1995).

117. Phan, A. T. *et al.* Two-repeat Tetrahymena telomeric d(TGGGGTTGGGGT) sequence interconverts between asymmetric dimeric G-quadruplexes in solution. *J. Mol. Biol.* **338**, 93–102 (2004).

118. Phan, A. T. *et al.* Structure of an unprecedented G-quadruplex scaffold in the human c-kit promoter. *J. Am. Chem. Soc.* **129**, 4386–4392 (2007).

119. Kuryavyi, V. *et al.* 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).

120. Hsu, S. T. D. *et al.* A G-rich sequence within the c-kit oncogene promoter forms a parallel G-quadruplex having asymmetric G-tetrad dynamics. *J. Am. Chem. Soc.* **131**, 13399–13409 (2009).

121. Phan, A. T. *et al.* Small-molecule interaction with a five-guanine-tract g-quadruplex structure from the human myc promoter. *Nat. Chem. Biol.* **1**, 167–173 (2005).

122. Phan, A. T. and Patel, D. J. A site-specific low-enrichment ^{15}N, ^{13}C isotope-labeling approach to unambiguous NMR spectral assignments in nucleic acids. *J. Am. Chem. Soc.* **124**, 1160–1161 (2002).

123. Feigon, J. *et al.* ^{1}H NMR spectroscopy of DNA triplexes and quadruplexes. *Methods Enzymol.* **261**, 225–255 (1995).

124. Patel, D. J. and Tonelli, A. E. Assignment of the proton nmr chemical shifts of the T□N3H and G□N1H proton resonances in isolated AT and GC Watson-Crick base pairs in double-stranded deoxy oligonucleotides in aqueous solution. *Biopolymers* **13**, 1943–1964 (1974).

125. Wang, Y. and Patel, D. J. Guanine Residues in d(T2AG3) and d(T2G4) Form Parallel-Stranded Potassium Cation Stabilized G-Quadruplexes with Anti Glycosidic Torsion Angles in Solution. *Biochemistry* **31**, 8112–8119 (1992).

126. Amrane, S. *et al.* Formation of pearl-necklace monomorphic G-quadruplexes in the human CEB25 minisatellite. *J. Am. Chem. Soc.* **134**, 5807–5816 (2012).

127. Lin, C. *et al.* NMR Studies of G-Quadruplex Structures and G-Quadruplex-Interactive Compounds. *Methods Mol. Biol.* **2035**, 157–176 (2019).

128. Vorlíčková, M. *et al.* Circular dichroism and guanine quadruplexes. *Methods* **57**, 64–75 (2012).

129. Karsisiotis, A. I. *et al.* Topological characterization of nucleic acid G-quadruplexes by UV absorption and circular dichroism. *Angew. Chemie - Int. Ed.* **50**, 10645–10648 (2011).

130. Kypr, J. *et al.* Circular dichroism and conformational polymorphism of DNA. *Nucleic Acids Res.* **37**, 1713–1725 (2009).

131. Paramasivan, S. *et al.* Circular dichroism of quadruplex DNAs: Applications to structure, cation effects and ligand binding. *Methods* **43**, 324–331 (2007).

132. del Villar-Guerra, R. *et al.* G-Quadruplex Secondary Structure Obtained from Circular Dichroism Spectroscopy. *Angew. Chemie - Int. Ed.* **57**, 7171–7175 (2018).

133. Carvalho, J. *et al.* Circular dichroism of G-Quadruplex: A laboratory experiment for the study of topology and ligand binding. *J. Chem. Educ.* **94**, 1547–1551 (2017).

134. Mergny, J. L. *et al.* Thermal difference spectra: A specific signature for nucleic acid structures. *Nucleic Acids Res.* **33**, 1–6 (2005).

135. Gray, R. D. and Chaires, J. B. Isothermal folding of G-quadruplexes. *Methods* **57**, 47–55 (2012).

136. Mergny, J. L. and Lacroix, L. Analysis of Thermal Melting Curves. *Oligonucleotides* **13**, 515–537 (2003).

137. Mergny, J. L. and Lacroix, L. UV melting of G-quadruplexes. *Curr. Protoc. Nucleic Acid Chem.* (2009).

138. Kaur, A. *et al.* Förster resonance energy transfer (FRET) and applications thereof. *Anal. Methods* **12**, 5532–5550 (2020).

139. Risitano, A. and Fox, K. R. Stability of intramolecular DNA quadruplexes: Comparison with DNA duplexes. *Biochemistry* **42**, 6507–6513 (2003).

140. Kumar, N. and Maiti, S. Quadruplex to Watson-Crick duplex transition of the thrombin binding aptamer: A fluorescence resonance energy transfer study. *Biochem. Biophys. Res. Commun.* **319**, 759–767 (2004).

141. De Cian, A. *et al.* Fluorescence-based melting assays for studying quadruplex ligands. *Methods* **42**,

183–195 (2007).

142. Gabelica, V. *et al.* Internal energy build-up in matrix-assisted laser desorption/ionization. *J. Mass Spectrom.* **39**, 579–593 (2004).

143. Gabelica, V. and De Pauw, E. Internal energy and fragmentation of ions produced in electrospray sources. *Mass Spectrom. Rev.* **24**, 566–587 (2005).

144. Kebarle, P. and Verkcerk, U. H. Electrospray: From Ions in solution to Ions in the gas phase, what we know now. *Mass Spectrom. Rev.* **28**, 898–917 (2009).

145. Ribaudo, G. *et al.* Combining Electrospray Mass Spectrometry (ESI-MS) and Computational Techniques in the Assessment of G-Quadruplex Ligands: A Hybrid Approach to Optimize Hit Discovery. *J. Med. Chem.* **64**, 13174–13190 (2021).

146. Siegel, M. Early Discovery Drug Screening Using Mass Spectrometry. *Curr. Top. Med. Chem.* **2**, 13–33 (2005).

147. Guittat, L. *et al.* Ascididemin and meridine stabilise G-quadruplexes and inhibit telomerase in vitro. *Biochim. Biophys. Acta - Gen. Subj.* **1724**, 375–384 (2005).

148. Mazzitelli, C. L. *et al.* Evaluation of binding of perylene diimide and benzannulated perylene diimide ligands to DNA by electrospray ionization mass spectrometry. *J. Am. Soc. Mass Spectrom.* **17**, 593–604 (2006).

149. 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).

150. Rosu, F. *et al.* Positive and negative ion mode ESI-MS and MS/MS for studying drug-DNA complexes. *Int. J. Mass Spectrom.* **253**, 156–171 (2006).

151. Yuan, G. *et al.* Mass spectrometry of G-quadruplex DNA: Formation, recognition, property, conversion, and conformation. *Mass Spectrom. Rev.* **30**, 1121–1142 (2011).

152. Marchand, A. *et al.* 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).

153. Li, H. *et al.* Spectroscopy probing of the formation, recognition, and conversion of a G-quadruplex in the promoter region of the bcl-2 oncogene. *Chem. - A Eur. J.* **15**, 2445–2452 (2009).

154. Li, H. *et al.* Investigation of Formation, Recognition, Stabilization, and Conversion of Dimeric G-Quadruplexes of HIV-1 Integrase Inhibitors by Electrospray Ionization Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **19**, 550–559 (2008).

155. Kankia, B. I. *et al.* Folding of the thrombin aptamer into a G-quadruplex with Sr2+: Stability, heat, and hydration. *J. Am. Chem. Soc.* **123**, 10799–10804 (2001).

156. Petraccone, L. *et al.* Stability and Structure of Telomeric DNA Sequences Forming Quadruplexes Containing Four G-Tetrads with Different Topological Arrangements. *Biochemistry* **43**, 4877–4884 (2004).

157. Redman, J. E. Surface plasmon resonance for probing quadruplex folding and interactions with proteins and small molecules. *Methods* **43**, 302–312 (2007).

158. Čeru, S. *et al.* A new pathway of DNA G-quadruplex formation. *Angew. Chemie - Int. Ed.* **53**, 4881–4884 (2014).

159. Pagano, B. *et al.* Applications of isothermal titration calorimetry in biophysical studies of G-quadruplexes. *Int. J. Mol. Sci.* **10**, 2935–2957 (2009).

160. Pagano, B. *et al.* Differential scanning calorimetry to investigate G-quadruplexes structural stability. *Methods* **64**, 43–51 (2013).

161. Prado, E. *et al.* Influence of the SPR Experimental Conditions on the G-Quadruplex DNA Recognition by Porphyrin Derivatives. *Langmuir* **34**, 13057–13064 (2018).

162. Darby, R. A. J. *et al.* High throughput measurement of duplex, triplex and quadruplex melting curves using molecular beacons and a LightCycler. *Nucleic Acids Res.* **30**, e39–e39 (2002).

163. Risitano, A. and Fox, K. R. Influence of loop size on the stability of intramolecular DNA quadruplexes. *Nucleic Acids Res.* **32**, 2598–2606 (2004).

164. Okamoto, K. *et al.* G-quadruplex structures of human telomere DNA examined by single molecule FRET and BrG-substitution. *Bioorganic Med. Chem.* **16**, 6873–6879 (2008).

165. Rehm, C. *et al.* A bacterial DNA quadruplex with exceptional K+ selectivity and unique structural polymorphism. *Chem. Sci.* **5**, 2809–2818 (2014).

166. He, Y., Neumann, R. D. and Panyutin, I. G. Intramolecular quadruplex conformation of human telomeric DNA assessed with 125I-radioprobing. *Nucleic Acids Res.* **32**, 5359–5367 (2004).

167. Gaynutdinov, T. I. *et al.* Structural polymorphism of intramolecular quadruplex of human telomeric DNA: Effect of cations, quadruplex-binding drugs and flanking sequences. *Nucleic Acids Res.* **36**, 4079–4087 (2008).

168. Hänsel, R. *et al.* High-resolution insight into G-overhang architecture. *J. Am. Chem. Soc.* **135**, 2816–2824 (2013).

169. Abu-Ghazalah, R. M. and Macgregor, R. B. Structural polymorphism of the four-repeat Oxytricha nova telomeric DNA sequences. *Biophys. Chem.* **141**, 180–185 (2009).

170. Miura, T. *et al.* A phase diagram for sodium and potassium ion control of polymorphism in telomeric DNA. *J. Mol. Biol.* **248**, 233–238 (1995).

171. Abu-Ghazalah, R. M. *et al.* Concentration-dependent structural transitions of human telomeric DNA sequences. *Biochemistry* **51**, 7357–7366 (2012).

172. Brown, N. M. *et al.* Exceptionally slow kinetics of the intramolecular quadruplex formed by the Oxytricha telomeric repeat. *Org. Biomol. Chem.* **3**, 4153–4157 (2005).

173. Rachwal, P. A. *et al.* Intramolecular DNA quadruplexes with different arrangements of short and long loops. *Nucleic Acids Res.* **35**, 4214–4222 (2007).

174. Salmaso, V. and Moro, S. Bridging molecular docking to molecular dynamics in exploring ligand-protein recognition process: An overview. *Front. Pharmacol.* **9**, (2018).

175. Pagano, B. *et al.* State-of-the-Art Methodologies for the Discovery and Characterization of DNA G-Quadruplex Binders. *Curr. Pharm. Des.* **18**, 1880–1899 (2012).

176. Puig Lombardi, E. and Londoño-Vallejo, A. A guide to computational methods for G-quadruplex prediction. *Nucleic Acids Res.* **48**, 1–15 (2020).

177. Haider, S. and Neidle, S. Molecular modeling and simulation of G-quadruplexes and quadruplex-ligand complexes. *Methods Mol. Biol.* **608**, 17–37 (2010).

178. Cosconati, S. *et al.* Tandem application of virtual screening and NMR experiments in the discovery of brand new DNA quadruplex groove binders. *J. Am. Chem. Soc.* **131**, 16336–16337 (2009).

179. Chan, D. S. H. *et al.* Structure-based optimization of FDA-approved drug methylene blue as a c-myc G-quadruplex DNA stabilizer. *Biochimie* **93**, 1055–1064 (2011).

180. Monsen, R. C. and Trent, J. O. G-quadruplex virtual drug screening: A review. *Biochimie* **152**, 134–148 (2018).

181. Chen, S. Bin *et al.* Pharmacophore-based discovery of triaryl-substituted imidazole as new telomeric G-quadruplex ligand. *Bioorg. Med. Chem. Lett.* **21**, 1004–1009 (2011).

182. 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).

183. Li, Q. *et al.* Searching Drug-Like Anti-cancer Compound(s) Based on G-Quadruplex Ligands. *Curr. Pharm. Des.* **18**, 1973–1983 (2012).

184. Alcaro, S. *et al.* Conformational studies and solvent-accessible surface area analysis of known selective DNA G-Quadruplex binders. *Biochimie* **93**, 1267–1274 (2011).

185. Amaro, R. E. *et al.* Ensemble Docking in Drug Discovery. *Biophys. J.* **114**, 2271–2278 (2018).

186. Miller, E. B. *et al.* Reliable and Accurate Solution to the Induced Fit Docking Problem for Protein–Ligand Binding. *J. Chem. Theory Comput.* **17**, 2630–2639 (2021).

187. Honig, B. and Nicholls, A. Classical electrostatics in biology and chemistry. *Science (80-.).* **268**, 1144–1149 (1995).

188. Jayaram, B. *et al.* Solvation free energy of biomacromolecules: Parameters for a modified generalized born model consistent with the AMBER force field. *J. Phys. Chem. B* **102**, 9571–9576 (1998).

189. Sitkoff, D. *et al.* Calculation of electrostatic effects at the amino terminus of an alpha helix. *Biophys. J.* **67**, 2251–2260 (1994).

190. Genheden, S. and Ryde, U. The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert Opin. Drug Discov.* **10**, 449–461 (2015).

191. Morris, G. M. *et al.* Software news and updates AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J. Comput. Chem.* **30**, 2785–2791 (2009).

192. Trott, O. and Olson, A. J. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **31**, 455-461 (2009).

193. Alcaro, S. *et al.* Identification and characterization of new DNA G-quadruplex binders selected by a combination of ligand and structure-based virtual screening approaches. *J. Med. Chem.* **56**, 843–855 (2013).

194. Trotta, R. *et al.* A more detailed picture of the interactions between virtual screening-derived hits and the DNA G-quadruplex: NMR, molecular modelling and ITC studies. *Biochimie* **93**, 1280–1287 (2011).

195. Ranjan, N. *et al.* Surface Dependent Dual Recognition of a G-quadruplex DNA With Neomycin-Intercalator Conjugates. *Front. Chem.* **8**, (2020).

196. Jorgensen, W. L. *et al.* Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J. Am. Chem. Soc.* **118**, 11225–11236 (1996).

197. Friesner, R. A. *et al.* Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. *J. Med. Chem.* **47**, 1739–1749 (2004).

198. Artese, A. *et al.* Identification of new natural DNA G-quadruplex binders selected by a structure-based virtual screening approach. *Molecules* **18**, 12051–12070 (2013).

199. Kar, R. K. *et al.* Novel G-quadruplex stabilizing agents: In-silico approach and dynamics. *J. Biomol. Struct. Dyn.* **31**, 1497–1518 (2013).

200. Heddi, B. and Phan, A. T. Structure of Human Telomeric DNA in Crowded Solution. *J. Am. Chem. Soc.* **133**, 9824–9833 (2011).

201. Wang, K. B. *et al.* Indenoisoquinoline Topoisomerase Inhibitors Strongly Bind and Stabilize the MYC Promoter G-Quadruplex and Downregulate MYC. *J. Am. Chem. Soc.* **141**, 11059–11070 (2019).

202. Dai, J. *et al.* 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).

203. Kang, H. J. and Park, H. J. In silico identification of novel ligands for G-quadruplex in the c-MYC promoter. *J. Comput. Aided. Mol. Des.* **29**, 339–348 (2015).

204. 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).

205. MacKerell, A. D. *et al.* Development and Current Status of the CHARMM Force Field for Nucleic Acids. *Biopolymers*. **56**, 257-265 (2001).

206. Hart, K. *et al.* Optimization of the CHARMM additive force field for DNA: Improved treatment of the BI/BII conformational equilibrium. *J. Chem. Theory Comput.* **8**, 348–362 (2012).

207. Salsbury, A. M. *et al.* Polarizable Molecular Dynamics Simulations of Two c-kit Oncogene Promoter G-Quadruplexes: Effect of Primary and Secondary Structure on Loop and Ion Sampling. *J. Chem. Theory Comput.* **16**, 3430–3444 (2020).

208. Salsbury, A. M. and Lemkul, J. A. Molecular Dynamics Simulations of the c-kit1 Promoter G-Quadruplex: Importance of Electronic Polarization on Stability and Cooperative Ion Binding. *J. Phys. Chem. B* **123**, 148–159 (2019).

209. Ratnasinghe, B. D. *et al.* Ion Binding Properties and Dynamics of the bcl-2 G-Quadruplex Using a Polarizable Force Field. *J. Chem. Inf. Model.* **60**, 6476–6488 (2020).

210. Pérez, A. *et al.* Refinement of the AMBER force field for nucleic acids: Improving the description of α/γ conformers. *Biophys. J.* **92**, 3817–3829 (2007).

211. Galindo-Murillo, R. *et al.* Assessing the Current State of Amber Force Field Modifications for DNA. *J. Chem. Theory Comput.* **12**, 4114–4127 (2016).

212. Zgarbová, M. *et al.* Refinement of the Sugar-Phosphate Backbone Torsion Beta for AMBER Force Fields Improves the Description of Z- and B-DNA. *J. Chem. Theory Comput.* **11**, 5723–5736 (2015).

213. Machireddy, B. *et al.* Binding of BRACO19 to a telomeric G-quadruplex DNA probed by all-atom molecular dynamics simulations with explicit solvent. *Molecules* **24**, (2019).

214. Havrila, M. *et al.* Effect of Monovalent Ion Parameters on Molecular Dynamics Simulations of G-Quadruplexes. *J. Chem. Theory Comput.* **13**, 3911–3926 (2017).

215. Joung, I. S. and Cheatham, T. E. Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations. *J. Phys. Chem. B* **112**, 9020–9041 (2008).

216. Rebič, M. *et al.* Molecular Dynamics Simulation Study of Parallel Telomeric DNA Quadruplexes at Different Ionic Strengths: Evaluation of Water and Ion Models. *J. Phys. Chem. B* **120**, 7380–7391 (2016).

217. Largy, E. *et al.* Visualizing the Quadruplex: From Fluorescent Ligands to Light-Up Probes. 111–177 (2012).

218. Stasyuk, O. A. *et al.* Aromaticity of H-bonded and metal complexes of guanine tautomers. *Struct. Chem.* **27**, 111–118 (2016).

219. Martinez, C. R. and Iverson, B. L. Rethinking the term 'pi-stacking'. *Chem. Sci.* **3**, 2191–2201 (2012).

220. Hunter, C. A. and Sanders, J. K. M. The Nature of π-π Interactions. *J. Am. Chem. Soc.* **112**, 5525–5534 (1990).

221. Chakrabarti, P. and Bhattacharyya, R. Geometry of nonbonded interactions involving planar groups in proteins. *Prog. Biophys. Mol. Biol.* **95**, 83–137 (2007).

222. Hunter, C. A. Aromatic interactions in proteins, DNA and synthetic receptors. *Philos. Trans. R. Soc. London. Ser. A Phys. Eng. Sci.* **345**, 77–85 (1993).

223. McGaughey, G. B. *et al.* π-Stacking interactions. Alive and well in proteins. *J. Biol. Chem.* **273**, 15458–15463 (1998).

224. Spiegel, J. *et al.* The Structure and Function of DNA G-Quadruplexes. *Trends Chem.* **2**, 123–136 (2020).

225. Ribaudo, G. *et al.* Investigation of the molecular reactivity of bioactive oxiranylmethyloxy anthraquinones. *Arch. Pharm. (Weinheim).* **352**, 1900030 (2019).

226. Agudelo, D. *et al.* Intercalation of antitumor drug doxorubicin and its analogue by DNA duplex: Structural features and biological implications. *Int. J. Biol. Macromol.* **66**, 144–150 (2014).

227. Sun, D. *et al.* Inhibition of human telomerase by a G-Quadruplex-Interactive compound. *J. Med. Chem.* **40**, 2113–2116 (1997).

228. Perry, P. J. *et al.* 1,4- and 2,6-disubstituted amidoanthracene-9,10-dione derivatives as inhibitors of human telomerase. *J. Med. Chem.* **41**, 3253–3260 (1998).

229. Zagotto, G. *et al.* Amide bond direction modulates G-quadruplex recognition and telomerase inhibition by 2,6 and 2,7 bis-substituted anthracenedione derivatives. *Bioorganic Med. Chem.* **16**, 354–361 (2008).

230. Zagotto, G. *et al.* Aminoacyl-anthraquinone conjugates as telomerase inhibitors: Synthesis, biophysical and biological evaluation. *J. Med. Chem.* **51**, 5566–5574 (2008).

231. Zagotto, G. *et al.* Tuning G-quadruplex vs double-stranded DNA recognition in regioisomeric lysyl-peptidyl-anthraquinone conjugates. *Bioconjug. Chem.* **22**, 2126–2135 (2011).

232. Fedoroff, O.Y. *et al.* NMR-Based Model of a Telomerase-Inhibiting Compound Bound to G-Quadruplex DNA†. *Biochemistry* **37**, 12367–12374 (1998).

233. Liu, Z.-R. *et al.* Selective DNA Binding of (N-alkylamine)-Substituted Naphthalene Imides and Diimides to G+C-rich DNA. *J. Biomol. Struct. Dyn.* **14**, 331–339 (2012).

234. Han, F. X. *et al.* 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).

235. 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).

236. Gowan, S. M. *et al.* A G-quadruplex-interactive potent small-molecule inhibitor of telomerase exhibiting in vitro and in vivo antitumor activity. *Mol. Pharmacol.* **61**, 1154–1162 (2002).

237. Moore, M. J. B. *et al.* Trisubstituted acridines as G-quadruplex telomere targeting agents. Effects of extensions of the 3,6- and 9-side chains on quadruplex binding, telomerase activity, and cell proliferation. *J. Med. Chem.* **49**, 582–599 (2006).

238. 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).

239. Heald, R. A. *et al.* Antitumor polycyclic acridines. 8. Synthesis and telomerase-inhibitory activity of methylated pentacyclic acridinium salts. *J. Med. Chem.* **45**, 590–597 (2002).

240. Gowan, S. M. *et al.* Potent Inhibition of Telomerase by Small-Molecule Pentacyclic Acridines Capable of Interacting with G-Quadruplexes. *Mol. Pharmacol.* **60**, 981–988 (2001).

241. Oszczapowicz, J. *et al.* Formation of quinoacridinium system. A novel reaction of quinaldinium salts. *Tetrahedron.* **44**, 6645–6650 (1988).

242. Gavathiotis, E. *et al.* Drug recognition and stabilisation of the parallel-stranded DNA quadruplex $d(TTAGGGT)_4$ containing the human telomeric repeat. *J. Mol. Biol.* **334**, 25–36 (2003).

243. Mulholland, K. *et al.* Binding modes and pathway of RHPS4 to human telomeric G-quadruplex and duplex DNA probed by all-atom molecular dynamics simulations with explicit solvent. *Phys. Chem. Chem. Phys.* **19**, 18685–18694 (2017).

244. Shin-ya, K. *et al.* Telomestatin, a novel telomerase inhibitor from Streptomyces anulatus [17]. *J. Am. Chem. Soc.* **123**, 1262–1263 (2001).

245. Doi, T. *et al.* (S)-Stereoisomer of telomestatin as a potent G-quadruplex binder and telomerase inhibitor. *Org. Biomol. Chem.* **9**, 387–393 (2011).

246. De Cian, A. *et al.* Affinity and selectivity of G4 ligands measured by FRET. *Nucleic Acids Symp. Ser. (Oxf).* 235–236 (2005).

247. Pirota, V. *et al.* An overview of quadruplex ligands: Their common features and chemotype diversity. *Annu. Rep. Med. Chem.* **54**, 163–196 (2020).

248. Kim, M. Y. *et al.* Telomestatin, a potent telomerase inhibitor that interacts quite specifically with the human telomeric intramolecular G-quadruplex. *J. Am. Chem. Soc.* **124**, 2098–2099 (2002).

249. Perry, P. J. *et al.* 2,7-Disubstituted amidofluorenone derivatives as inhibitors of human telomerase. *J. Med. Chem.* **42**, 2679–2684 (1999).

250. Harrison, R. J. *et al.* Human telomerase inhibition by substituted acridine derivatives. *Bioorganic Med. Chem. Lett.* **9**, 2463–2468 (1999).

251. Riou, J. F. *et al.* Cell senescence and telomere shortening induced by a new series of specific G-quadruplex DNA ligands. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 2672–2677 (2002).

252. Nakatani, K. *et al.* Induction of a remarkable conformational change in a human telomeric sequence by the binding of naphthyridine dimer: Inhibition of the elongation of a telomeric repeat by telomerase. *J. Am. Chem. Soc.* **125**, 662–666 (2003).

253. Moorhouse, A. D. *et al.* Stabilization of G-quadruplex DNA by highly selective ligands via click chemistry. *J. Am. Chem. Soc.* **128**, 15972–15973 (2006).

254. Le Sann, C. *et al.* Synthesis and preliminary evaluation of novel analogues of quindolines as potential stabilisers of telomeric G-quadruplex DNA. *Tetrahedron* **63**, 12903–12911 (2007).

255. Martins, C. *et al.* Structure-based design of benzylamino-acridine compounds as G-quadruplex DNA telomere targeting agents. *Bioorganic Med. Chem. Lett.* **17**, 2293–2298 (2007).

256. Waller, Z. A. E. *et al.* Triarylpyridines: A versatile small molecule scaffold for G-quadruplex recognition. *Chem. Commun.* 1467–1469 (2008).

257. Dash, J. *et al.* G-quadruplex recognition by bis-indole carboxamides. *Chem. Commun.* 3055–3057 (2008).

258. 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).

259. Laronze-Cochard, M. *et al.* Synthesis and biological evaluation of novel 4,5-bis(dialkylaminoalkyl)-substituted acridines as potent telomeric G-quadruplex ligands. *Eur. J. Med. Chem.* **44**, 3880–3888 (2009).

260. Tan, J. H. *et al.* Isaindigotone derivatives: A new class of highly selective ligands for telomeric G-quadruplex DNA. *J. Med. Chem.* **52**, 2825–2835 (2009).

261. Bates, P. *et al.* Quartets in G-major. The First International Meeting on Quadruplex DNA. *EMBO Rep.* **8**, 1003–1010 (2007).

262. Franceschin, M. *et al.* Natural and synthetic G-quadruplex interactive berberine derivatives. *Bioorganic Med. Chem. Lett.* **16**, 1707–1711 (2006).

263. Ma, Y. *et al.* Synthesis and evaluation of 9-O-substituted berberine derivatives containing aza-aromatic terminal group as highly selective telomeric G-quadruplex stabilizing ligands. *Bioorganic Med. Chem. Lett.* **19**, 3414–3417 (2009).

264. Paul, A. *et al.* Binding of Gemini Bisbenzimidazole Drugs with Human Telomeric G-Quadruplex Dimers: Effect of the Spacer in the Design of Potent Telomerase Inhibitors. *PLoS One* **7**, e39467 (2012).

265. Maji, B. *et al.* Design and synthesis of new benzimidazole-carbazole conjugates for the stabilization of human telomeric DNA, telomerase inhibition, and their selective action on cancer cells. *J. Med. Chem.* **57**, 6973–6988 (2014).

266. Maji, B. *et al.* New dimeric carbazole-benzimidazole mixed ligands for the stabilization of human telomeric G-quadruplex DNA and as telomerase inhibitors. A remarkable influence of the spacer. *Org. Biomol. Chem.* **13**, 8335–8348 (2015).

267. Marchand, A. *et al.* Selective and Cooperative Ligand Binding to Antiparallel Human Telomeric DNA G-Quadruplexes. *Chem. - A Eur. J.* **22**, 9551–9555 (2016).

268. Shibata, T. *et al.* Characterization of the interaction between heme and a parallel G-quadruplex DNA formed from d(TTGAGG). *Biochim. Biophys. Acta - Gen. Subj.* **1861**, 1264–1270 (2017).

269. Arola-Arnal, A. *et al.* Effects of metal coordination geometry on stabilization of human telomeric quadruplex DNA by square-planar and square-pyramidal metal complexes. *Inorg. Chem.* **47**, 11910–11919 (2008).

270. Evans, S. E. *et al.* End-stacking of copper cationic porphyrins on parallel-stranded guanine quadruplexes. *J. Biol. Inorg. Chem.* **12**, 1235–1249 (2007).

271. Pan, J. and Zhang, S. Interaction between cationic zinc porphyrin and lead ion induced telomeric guanine quadruplexes: Evidence for end-stacking. *J. Biol. Inorg. Chem.* **14**, 401–407 (2009).

272. Bertrand, H. *et al.* The importance of metal geometry in the recognition of G-quadruplex-DNA by metal-terpyridine complexes. *Org. Biomol. Chem.* **5**, 2555–2559 (2007).

273. Campbell, N. H. *et al.* Molecular basis of structure-activity relationships between salphen metal complexes and human telomeric DNA quadruplexes. *J. Med. Chem.* **55**, 209–222 (2012).

274. Largy, E. *et al.* Tridentate N-donor palladium(II) complexes as efficient coordinating quadruplex DNA binders. *Chem. - A Eur. J.* **17**, 13274–13283 (2011).

275. Gray, R. D. *et al.* Energetics and kinetics of a conformational switch in G-quadruplex DNA. *J. Phys. Chem. B* **113**, 2676–2683 (2009).

276. Li, H. and Wei, C. Mechanism studies on anti-HepG2 cell proliferation of phenanthroline derivatives

as G-quadruplex DNA stabilizers. *Bioorganic Med. Chem. Lett.* **25**, 3798–3803 (2015).

277. Nielsen, M. C. *et al.* Design, synthesis and evaluation of 4,7-diamino-1,10-phenanthroline G-quadruplex ligands. *Bioorganic Med. Chem.* **17**, 8241–8246 (2009).

278. De Cian, A. *et al.* Highly efficient G-quadruplex recognition by bisquinolinium compounds. *J. Am. Chem. Soc.* **129**, 1856–1857 (2007).

279. Musetti, C. *et al.* Metal ion-mediated assembly of effective phenanthroline-based G-quadruplex ligands. *Dalt. Trans.* 3657–3660 (2009).

280. Nielsen, M. C. *et al.* Phenanthroline-2,9-bistriazoles as selective G-quadruplex ligands. *Eur. J. Med. Chem.* **72**, 119–126 (2014).

281. Larsen, A. F. and Ulven, T. Efficient synthesis of 4,7-diamino substituted 1,10-phenanthroline-2,9-dicarboxamides. *Org. Lett.* **13**, 3546–3548 (2011).

282. Larsen, A. F. *et al.* Tetrasubstituted phenanthrolines as highly potent, water-soluble, and selective G-quadruplex ligands. *Chem. - A Eur. J.* **18**, 10892–10902 (2012).

283. Amato, J. *et al.* Bis-indole derivatives with antitumor activity turn out to be specific ligands of human telomeric G-quadruplex. *Front. Chem.* **2**, (2014).

284. Wang, L. *et al.* Promoting the formation and stabilization of human telomeric G-quadruplex DNA, inhibition of telomerase and cytotoxicity by phenanthroline derivatives. *Org. Biomol. Chem.* **9**, 2648–2653 (2011).

285. Wang, L. and Wei, C. Spectroscopic and biological studies of phenanthroline compounds: Selective recognition of gene-promoter G-Quadruplex Dnas preferred over duplex DNA. *Chem. Biodivers.* **10**, 1154–1164 (2013).

286. Liao, S. *et al.* Microwave-assisted synthesis of phenanthroimidazole derivatives as stabilizer of c-myc G-quadruplex DNA. *Bioorganic Med. Chem.* **22**, 6503–6508 (2014).

287. Răsădean, D. M. *et al.* Amino-Acid-Derived Naphthalenediimides as Versatile G-Quadruplex Binders. *Chem. - A Eur. J.* **23**, 8491–8499 (2017).

288. Tikhomirov, A. S. *et al.* Tri-armed ligands of G-quadruplex on heteroarene-fused anthraquinone scaffolds: Design, synthesis and pre-screening of biological properties. *Eur. J. Med. Chem.* **159**, 59–73 (2018).

289. Saha, U. *et al.* Targeting human telomeric DNA quadruplex with novel berberrubine derivatives: insights from spectroscopic and docking studies. *J. Biomol. Struct. Dyn.* **37**, 1375–1389 (2019).

290. 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).

291. Gao, C. *et al.* BMPQ-1 binds selectively to (3+1) hybrid topologies in human telomeric G-quadruplex multimers. *Nucleic Acids Res.* **48**, 11259–11269 (2020).

292. Sasaki, S. *et al.* Linear consecutive hexaoxazoles as G4 ligands inducing chair-type anti-parallel topology of a telomeric G-quadruplex. *RSC Adv.* **10**, 43319–43323 (2020).

293. Craciun, A. M. *et al.* New 2,9-disubstituted-1,10-phenanthroline derivatives with anticancer activity by selective targeting of telomeric G-quadruplex DNA. *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* **249**, 119318 (2021).

294. Dervan, P. B. and Bürli, R. W. Sequence-specific DNA recognition by polyamides. *Curr. Opin. Chem. Biol.* **3**, 688–693 (1999).

295. Chen, A. Y. *et al.* DNA minor groove-binding ligands: A different class of mammalian DNA topoisomerase I inhibitors. *Proc. Natl. Acad. Sci. U. S. A.* **90**, 8131–8135 (1993).

296. Cinelli, M. A. Topoisomerase 1B poisons: Over a half-century of drug leads, clinical candidates, and serendipitous discoveries. *Med. Res. Rev.* **39**, 1294–1337 (2019).

297. Asamitsu, S. *et al.* Ligand Design to Acquire Specificity to Intended G-Quadruplex Structures. **25**, 417–430 (2019).

298. Mazerski, J. *et al.* The geometry of intercalation complex of antitumor mitoxantrone and ametantrone with DNA: molecular dynamics simulations. *Acta Biochim. Pol.* **45**, 1–11 (1998).

299. Pommier, Y. *et al.* DNA topoisomerases and their poisoning by anticancer and antibacterial drugs. *Chem. Biol.* **17**, 421–433 (2010).

300. Wilson, W. D. *et al.* The Search for Structure-Specific Nucleic Acid-Interactive Drugs: Effects of Compound Structure on RNA versus DNA Interaction Strength. *Biochemistry* **32**, 4098–4104 (1993).

301. Kumar, C. V. and Asuncion, E. H. DNA Binding Studies and Site Selective Fluorescence Sensitization of an Anthryl Probe. *J. Am. Chem. Soc.* **115**, 8547–8553 (1993).

302. Van Arman, S. A. and Czarnik, A. W. General Fluorescence Assay for Enzyme-Catalyzed Polyanion Hydrolysis Based on Template-Directed Excimer Formation. Application to Heparin and Polyglutamate. *J. Am. Chem. Soc.* **112**, 5376–5377 (1990).

303. Ostaszewski, R. *et al.* The synthesis of a new type of anthracene DNA intercalator. *Bioorg. Med. Chem. Lett.* **8**, 2995–2996 (1998).

304. Folini, M. *et al.* Remarkable interference with telomeric function by a G-quadruplex selective bisantrene regioisomer. *Biochem. Pharmacol.* **79**, 1781–1790 (2010).

305. Ribaudo, G. *et al.* Constrained bisantrene derivatives as G-quadruplex binders. *Arkivoc* **2016**, 145–160 (2016).

306. Gama, S. *et al.* Anthracene-terpyridine metal complexes as new G-quadruplex DNA binders. *J. Inorg. Biochem.* **160**, 275–286 (2016).

307. Liu, X. *et al.* Cu-NHC-TEMPO catalyzed aerobic oxidation of primary alcohols to aldehydes. *J. Org. Chem.* **78**, 8531–8536 (2013).

308. Le, C. M. *et al.* Synergistic steric effects in the development of a palladium-catalyzed alkyne carbohalogenation: Stereodivergent synthesis of vinyl halides. *Angew. Chemie - Int. Ed.* **54**, 254–257 (2015).

309. Nagarjuna, G., Ren, Y. and Moore, J. S. Synthesis and reactivity of anthracenyl-substituted arenediynes. *Tetrahedron Lett.* **56**, 3155–3159 (2015).

310. Schrödinger Release 2020-1: Glide; Schrödinger, LLC: New York, NY, 2020.

311. Roos, K. *et al.* OPLS3e: Extending Force Field Coverage for Drug-Like Small Molecules. *J. Chem. Theory Comput.* **15**, 1863–1874 (2019).

312. Sahu, B. *et al.* Synthetic and mechanistic investigations on the rearrangement of 2,3-unsaturated 1,4-bis(alkylidene)carbenes to enediynes. *European J. Org. Chem.* 2477–2489 (2007).

313. Thummala, Y. *et al.* DBU-Mediated Synthesis of Aryl Acetylenes or 1-Bromoethynylarenes from Aldehydes. *Adv. Synth. Catal.* **361**, 611–616 (2019).

314. Morri, A. K. *et al.* The Dual Role of 1,8-Diazabicyclo[5.4.0]undec-7-ene (DBU) in the Synthesis of Terminal Aryl- and Styryl-Acetylenes via Umpolung Reactivity. *Org. Lett.* **17**, 4640–4643 (2015).

315. Bieber, L. W. and Da Silva, M. F. Mild and efficient synthesis of propargylamines by copper-catalyzed Mannich reaction. *Tetrahedron Lett.* **45**, 8281–8283 (2004).

316. Peshkov, V. A. *et al.* A walk around the A3-coupling. *Chem. Soc. Rev.* **41**, 3790–3807 (2012).

317. https://www.molinspiration.com.

318. Petit, J. *et al.* Softening the Rule of Five - Where to draw the line? *Bioorganic Med. Chem.* **20**, 5343–5351 (2012).

319. Schrödinger Release 2021-3: Glide, Schrödinger, LLC, New York, NY, 2021.

320. Marchand, A. and Gabelica, V. Native electrospray mass spectrometry of DNA G-quadruplexes in potassium solution. *J. Am. Soc. Mass Spectrom.* **25**, 1146–1154 (2014).

321. Marchand, A. and Gabelica, V. Folding and misfolding pathways of G-quadruplex DNA. *Nucleic Acids Res.* **44**, 10999–11012 (2016).

322. Altieri, A. *et al.* Xanthene and xanthone derivatives as G-quadruplex stabilizing ligands. *Molecules* **18**, 13446–13470 (2013).

323. Ribaudo, G. *et al.* Evidence on selective binding to G-quadruplex DNA of isoflavones from Maclura pomifera by mass spectrometry and molecular docking. *Nat. Prod. Res.* **35**, 2583–2587.

324. Ribaudo, G. *et al.* Photoactivated semi-synthetic derivative of osajin selectively interacts with G-quadruplex DNA. *Nat. Prod. Res.* (2020).

325. Rigo, R. *et al.* G-quadruplexes in human promoters: A challenge for therapeutic applications. *Biochim. Biophys. Acta - Gen. Subj.* **1861**, 1399–1413 (2017).

326. Chang, Y. S. *et al.* SARS unique domain (SUD) of severe acute respiratory syndrome coronavirus induces NLRP3 inflammasome-dependent CXCL10-mediated pulmonary inflammation. *Int. J. Mol. Sci.* **21**, (2020).

327. Shen, L. W. *et al.* Inhibition of Influenza A virus propagation by benzoselenoxanthenes stabilizing TMPRSS2 Gene G-quadruplex and hence down-regulating TMPRSS2 expression. *Sci. Rep.* **10**, (2020).

328. Di Fonzo, S. *et al.* Ligand binding to G-quadruplex DNA: New insights from ultraviolet resonance Raman spectroscopy. *Phys. Chem. Chem. Phys.* **22**, 8128–8140 (2020).

329. Sengupta, A. *et al.* Promise of G-quadruplex structure binding ligands as epigenetic modifiers with anti-cancer effects. *Molecules* **24**, (2019).

330. Pirota, V. *et al.* Naphthalene diimides as multimodal G-quadruplex-selective ligands. *Molecules* **24**, (2019).

331. Bortolus, M. *et al.* Photo-induced spin switching in a modified anthraquinone modulated by DNA binding. *Photochem. Photobiol. Sci.* **18**, 2199–2207 (2019).

332. Ongaro, A. *et al.* Design and synthesis of a peptide derivative of ametantrone targeting the major groove of the $d(GGCGCC)_2$ palindromic sequence. *New J. Chem.* **44**, 3624–3631 (2020).

333. Chilka, P. *et al.* Small molecule fluorescent probes for G-quadruplex visualization as potential cancer theranostic agents. *Molecules* **24**, (2019).

334. Ongaro, A. *et al.* Synthesis via A3 Coupling Reaction of Anthracene-Propargylamine as a New Scaffold for the Interaction with DNA. *ChemistrySelect* **4**, 13138–13142 (2019).

335. Schrödinger Release 2020-1: Desmond Molecular Dynamics System, 2020, D. E. Shaw Research; Maestro-Desmond Interoperability Tools, Schrödinger: New York, NY, 2020.

336. Ferreira, R. *et al.* Mass spectrometry and ion mobility spectrometry of G-quadruplexes. A study of solvent effects on dimer formation and structural transitions in the telomeric DNA sequence d(TAGGGTTAGGGT). *Methods* **57**, 56–63 (2012).

337. Sonogashira, K. Development of Pd–Cu catalyzed cross-coupling of terminal acetylenes with sp2-carbon halides. *J. Organomet. Chem.* **653**, 46–49 (2002).

338. Xu, N. *et al.* Evaluation of alkaloids binding to the parallel quadruplex structure [d(TGGGGT)]4 by electrospray ionization mass spectrometry. *J. Mass Spectrom.* **47**, 694–700 (2012).

339. Kosiol, N. *et al.* G-quadruplexes: a promising target for cancer therapy. *Molecular Cancer* **20**, 1–18 (2021).

340. Zagotto, G. *et al.* Novel ametantrone-amsacrine related hybrids as topoisomerase IIβ poisons and cytotoxic agents. *Arch. Pharm. (Weinheim).* **347**, 728–737 (2014).

341. Ribaudo, G. *et al.* 9,10-Bis[(4-(2-hydroxyethyl)piperazine-1-yl)prop-2-yne-1-yl]anthracene: Synthesis and G-quadruplex Selectivity. *Molbank 2020, Vol. 2020, Page M1138* **2020**, M1138 (2020).

342. Ribaudo, G. *et al.* Enhanced G-quadruplex selectivity of flavonoid glycoside rutin over quercetin. (2020).

343. Percivalle, C. *et al.* Aryl ethynyl anthraquinones: A useful platform for targeting telomeric G-quadruplex structures. *Org. Biomol. Chem.* **12**, 3744–3754 (2014).

344. Appendino, G. *et al.* The 1,2,3-triazole ring as a peptido- and olefinomimetic element: Discovery of click yanilloids and cannabinoids. *Angew. Chemie - Int. Ed.* **46**, 9312–9315 (2007).

345. Bonandi, E. *et al.* The 1,2,3-triazole ring as a bioisostere in medicinal chemistry. *Drug Discovery Today* vol. **22**, 1572–1581 (2017).

346. Angell, Y. and Burgess, K. Ring Closure to β-Turn Mimics via Copper-Catalyzed Azide/Alkyne Cycloadditions. *J. Org. Chem.* **70**, 9595–9598 (2005).

347. Chau, N. T. T. *et al.* β-Cyclodextrins grafted with chiral amino acids: A promising supramolecular stabilizer of nanoparticles for asymmetric hydrogenation? *Appl. Catal. A Gen.* **467**, 497–503 (2013).

348. Soethoudt, M. *et al.* Selective Photoaffinity Probe That Enables Assessment of Cannabinoid CB2 Receptor Expression and Ligand Engagement in Human Cells. *J. Am. Chem. Soc.* **140**, 6067–6075 (2018).

349. Haridas, V. *et al.* Designer peptide dendrimers using click reaction. *Tetrahedron* **67**, 1873–1884 (2011).

350. Parkinson, G. N. *et al.* Crystal structure of parallel quadruplexes from human telomeric DNA. (2002).

351. Madhavi Sastry, G. *et al.* Protein and ligand preparation: Parameters, protocols, and influence on virtual screening enrichments. *J. Comput. Aided. Mol. Des.* **27**, 221–234 (2013).

352. Lu, C. *et al.* OPLS4: Improving Force Field Accuracy on Challenging Regimes of Chemical Space. *J. Chem. Theory Comput.* **17**, 4291–4300 (2021).

353. Rocard, L. *et al.* Templated Chromophore Assembly by Dynamic Covalent Bonds. *Angew. Chem. Int. Ed. Engl.* 54, 15739-43 (2015).

354. Vilaça, H. *et al.* New cyclic RGD peptides: Synthesis, characterization, and theoretical activity towards α v β 3 integrin. *Tetrahedron* **70**, 5420–5427 (2014).

355. Ling, T. *et al.* (+)-Dehydroabietylamine derivatives target triple-negative breast cancer. *Eur. J. Med. Chem.* **102**, 9–13 (2015).

356. Späth, A. and König, B. Binding of a hemoregulatory tetrapeptide by a bis-guanidinium crown ether. *Tetrahedron.* **66**, 6019–6025 (2010).

357. Hein, J. E. and Fokin, V. V. Copper-catalyzed azide-alkyne cycloaddition (CuAAC) and beyond: new reactivity of copper(I) acetylides. *Chem. Soc. Rev.* **39**, 1302–1315 (2010).

358. Worrell, B. T. *et al.* Direct evidence of a dinuclear copper intermediate in Cu(I)-catalyzed azide-alkyne cycloadditions. *Science (80-.).* **340**, 457–460 (2013).

359. Rosu, F. *et al.* Electrospray mass spectrometry to study drug-nucleic acids interactions. *Biochimie*. **90**, 1074–1087 (2008).

360. Mazzitelli, C. L. *et al.* Screening of Threading Bis-Intercalators Binding to Duplex DNA by Electrospray Ionization Tandem Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **18**, 311–321 (2007).

361. Darby, R. A. J. *et al.* High throughput measurement of duplex, triplex and quadruplex melting curves using molecular beacons and a LightCycler. *Nucleic Acids Res.* **30**, (2002).

362. De Cian, A. *et al.* Targeting telomeres and telomerase. *Biochimie.* **90**, 131–155 (2008).

363. RA, F. *et al.* Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. *J. Med. Chem.* **49**, 6177–6196 (2006).

7 Abbreviations

°C	celsius degrees
-mer	oligonucleotide length: 22-mer means an oligonucleotide composed of 22 nucleotides
AQ	anthraquinone
CD	circular dichroism
CuAAC	copper catalyzed azido alkyne cycloaddition
DCM	dichloromethane
DMF	dimethylformamide
DMSO	dimethylsulfoxide
DNA	deoxyribonucleic acid
dsDNA	double stranded DNA
E_{com}	center of mass energy
EDC	1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide
ESI-MS	electrospray mass spectrometry
Et_3N	triethylamine
EtOAc	ethyl acetate
EtOH	ethanol
eV	electron volt
F10T	5´- FAM- dTATAGCTATA-HEG-TATAGCTATA-TAMRA-3´
F21T	5´-FAM-G3[TTAGGG]-TAMRA-3'3
ff	force field
FRET	Förster or fluorescence resonance energy transfer
G4	G-quadruplex

HOBt	hydroxy benzo triazole
JC	Joung-Cheatham ions parameters
K_b	binding constant
kcal/mol	kilo calories on moles
K_d	dissociation constant
MD	molecular dynamics
MeOH	methanol
MM/GBSA	molecular mechanics/generalized Born and surface area continuum solvation
MM/PBSA	molecular mechanics/Poisson–Boltzmann and surface area continuum solvation
NaN3	sodium azide
nM	nano molar
nm	nano meters
NMR	nuclear magnetic resonance
NOE	nuclear Overhauser effect
PDB	protein data bank
pL	pico liters
ppm	parts per milions
RMSD	root mean square deviation, usually expressed in Å (ångström)
RNA	ribonucleic acid
Taq	Taq polymerase, thermostable DNA polymerase I
Tel23	telomeric sequence of 23 nucleotides
Tel24	telomeric sequence of 24 nucleotides
TFA	trifluoroacetic acid

$TIP4P_{ew}$	four sites water model used in MD
TMS	tetramethylsylane
TRAP	telomeric repeat amplification protocol
uM	micro molar
UV-Vis	ultraviolet-visible

Acknownledgements

Un sincero ringraziamento al mio supervisore, la Prof.ssa Alessandra Gianoncelli, al Dott. Giovanni Ribaudo ed alla dottoranda Erika Oselladore, che hanno reso possibile il lavoro svolto. Un ringraziamento anche al Prof. Giuseppe Zagotto per il supporto scientifico e logistico.

I would like to sincerely thank my supervisor, Prof. Alessandra Gianoncelli, Doctor Giovanni Ribaudo and the PhD student Erika Oselladore, which made possible the work done. I would also like to thank Prof. Giuseppe Zagotto for the scientific and logistic support.

Publications list (Alberto Ongaro, 28/10/2021, Scopus)

Ribaudo, G., Ongaro, A., Oselladore, E., Memo, M., Gianoncelli, A.
Combining Electrospray Mass Spectrometry (ESI-MS) and Computational Techniques in the Assessment of G-Quadruplex Ligands: A Hybrid Approach to Optimize Hit Discovery
(2021) Journal of Medicinal Chemistry, 64 (18), pp. 13174-13190.

Ongaro, A., Oselladore, E., Memo, M., Ribaudo, G., Gianoncelli, A.
Insight into the LFA-1/SARS-CoV-2 Orf7a Complex by Protein-Protein Docking, Molecular Dynamics, and MM-GBSA Calculations
(2021) Journal of Chemical Information and Modeling, 61 (6), pp. 2780-2787.

Ribaudo, G., Bortoli, M., Oselladore, E., Ongaro, A., Gianoncelli, A., Zagotto, G., Orian, L.
Selenoxide elimination triggers enamine hydrolysis to primary and secondary amines: A combined experimental and theoretical investigation
(2021) Molecules, 26 (9), art. no. 2770.

Mastinu, A., Ribaudo, G., Ongaro, A., Anna Bonini, S., Memo, M., Gianoncelli, A.
Critical review on the chemical aspects of cannabidiol (Cbd) and harmonization of computational bioactivity data
(2021) Current Medicinal Chemistry, 28 (2), pp. 213-237.

Ribaudo, G., Ongaro, A., Zagotto, G., Memo, M., Gianoncelli, A.
Evidence on selective binding to G-quadruplex DNA of isoflavones from Maclura pomifera by mass spectrometry and molecular docking
(2021) Natural Product Research, 35 (15), pp. 2583-2587.

Ongaro, A., Zagotto, G., Memo, M., Gianoncelli, A., Ribaudo, G.
Natural phosphodiesterase 5 (PDE5) inhibitors: a computational approach
(2021) Natural Product Research, 35 (10), pp. 1648-1653.

Oselladore, E., Ongaro, A., Zagotto, G., Memo, M., Ribaudo, G., Gianoncelli, A.
Combinatorial library generation, molecular docking and molecular dynamics simulations for enhancing the isoflavone scaffold in phosphodiesterase inhibition
(2020) New Journal of Chemistry, 44 (45), pp. 19472-19488.

Gianoncelli, A., Ongaro, A., Zagotto, G., Memo, M., Ribaudo, G.
2-(3,4-dihydroxyphenyl)-4-(2-(4-nitrophenyl)hydrazono)-4h-chromene-3,5,7-triol
(2020) MolBank, 2020 (3), art. no. M1144, pp. 1-6.

Ribaudo, G., Ongaro, A., Zagotto, G., Memo, M., Gianoncelli, A.
Therapeutic Potential of Phosphodiesterase Inhibitors against Neurodegeneration: The Perspective of the Medicinal Chemist
(2020) ACS Chemical Neuroscience, 11 (12), pp. 1726-1739.

Ribaudo, G., Ongaro, A., Oselladore, E., Zagotto, G., Memo, M., Gianoncelli, A.
9,10-Bis[(4-(2-hydroxyethyl)piperazine-1-yl)prop-2-yne-1-yl]anthracene: Synthesis and g-quadruplex selectivity
(2020) MolBank, 2020 (2), art. no. M1138.

Dalle Vedove, A., Zonta, F., Zanforlin, E., Demitri, N., Ribaudo, G., Cazzanelli, G., Ongaro, A., Sarno, S., Zagotto, G., Battistutta, R., Ruzzene, M., Lolli, G.
A novel class of selective CK2 inhibitors targeting its open hinge conformation
(2020) European Journal of Medicinal Chemistry, 195, art. no. 112267.

Ribaudo, G., Bortoli, M., Ongaro, A., Oselladore, E., Gianoncelli, A., Zagotto, G., Orian, L.
Fluoxetine scaffold to design tandem molecular antioxidants and green catalysts
(2020) RSC Advances, 10 (32), pp. 18583-18593.

Ongaro, A., Povolo, C., Zagotto, G., Ribaudo, G.

HPLC and NMR quantification of bioactive compounds in flowers and leaves of Brassica rapa: the influence of aging
(2020) Natural Product Research, 34 (9), pp. 1288-1291.

Ribaudo, G., Zagotto, G., Ongaro, A., Ricci, A., Caretta, A., Mucignat-Caretta, C.
A new sensitive and subunit-selective molecular tool for investigating protein kinase A in the brain
(2020) Archiv der Pharmazie, 353 (4), art. no. 1900326.

Ongaro, A., Ribaudo, G., Braud, E., Ethève-Quelquejeu, M., De Franco, M., Garbay, C., Demange, L., Gresh, N., Zagotto, G.
Design and synthesis of a peptide derivative of ametantrone targeting the major groove of the d(GGCGCC)2 palindromic sequence
(2020) New Journal of Chemistry, 44 (9), pp. 3624-3631.

Ribaudo, G., Oselladore, E., Ongaro, A., Zagotto, G., Memo, M., Gianoncelli, A.
Enhanced G-quadruplex selectivity of flavonoid glycoside rutin over quercetin
(2020) Natural Product Research.

Ribaudo, G., Ongaro, A., Oselladore, E., Zagotto, G., Memo, M., Gianoncelli, A.
A computational approach to drug repurposing against SARS-CoV-2 RNA dependent RNA polymerase (RdRp)
(2020) Journal of Biomolecular Structure and Dynamics.

Ribaudo, G., Ongaro, A., Zagotto, G., Memo, M., Gianoncelli, A.
Photoactivated semi-synthetic derivative of osajin selectively interacts with G-quadruplex DNA
(2020) Natural Product Research.

Ongaro, A., Ribaudo, G., Zagotto, G., Memo, M., Gianoncelli, A.
Synthesis via A3 Coupling Reaction of Anthracene-Propargylamine as a New Scaffold for the Interaction with DNA
(2019) ChemistrySelect, 4 (45), pp. 13138-13142.

Ribaudo, G., Ongaro, A., Zorzan, M., Pezzani, R., Redaelli, M., Zagotto, G., Memo, M., Gianoncelli, A.
Investigation of the molecular reactivity of bioactive oxiranylmethyloxy anthraquinones
(2019) Archiv der Pharmazie, 352 (5), art. no. 1900030.

Zorzan, M., Collazuol, D., Ribaudo, G., Ongaro, A., Scaroni, C., Zagotto, G., Armanini, D., Barollo, S., Galeotti, F., Volpi, N., Redaelli, M., Pezzani, R.
Biological effects and potential mechanisms of action of Pistacia lentiscus Chios mastic extract in Caco-2 cell model
(2019) Journal of Functional Foods, 54, pp. 92-97.

Ribaudo, G., Ongaro, A., Zagotto, G.
Natural compounds promoting weight loss: Mechanistic insights from the point of view of the medicinal chemist
(2019) Natural Products Journal, 9 (2), pp. 78-85.

Ribaudo, G., Ongaro, A., Zagotto, G.
56386043200;57203207633;6603713437;
5-Hydroxy-3-(4-hydroxyphenyl)-8,8-dimethyl-6- (3-methylbut-2-enyl)pyrano[2,3-h]chromen-4-one
(2018) MolBank, 2018 (3), art. no. M1004.